Copula Methods for Forecasting Multivariate Time Series*

Andrew J. Patton

Duke University

29 May 2012

Forthcoming in the Handbook of Economic Forecasting, Volume 2.

Abstract

Copula-based models provide a great deal of flexibility in modelling multivariate distributions, allowing the researcher to specify the models for the marginal distributions separately from the dependence structure (copula) that links them to form a joint distribution. In addition to flexibility, this often also facilitates estimation of the model in stages, reducing the computational burden. This chapter reviews the growing literature on copula-based models for economic and financial time series data, and discusses in detail methods for estimation, inference, goodness-of-fit testing, and model selection that are useful when working with these models. A representative data set of two daily equity index returns is used to illustrate all of the main results.

Keywords: dependence, correlation, tail risk, volatility, density forecasting.

J.E.L. codes: C32, C51, C53.

*I thank the Editors (Graham Elliott and Allan Timmermann), two anonymous referees, and Yanqin Fan, Dominique Guégan, Bruno Rémillard, and seminar participants at the Federal Reserve Bank of Saint Louis and HEC Montréal for helpful comments and suggestions, and Dong Hwan Oh for outstanding research assistance. Contact address: Department of Economics, Duke University, 213 Social Sciences Building, Box 90097, Durham NC 27708-0097. Email: andrew.patton@duke.edu. Matlab code to replicate the analysis in this chapter is available at `http://econ.duke.edu/~ap172/code.html`.

1 Introduction

This chapter reviews the growing literature on copula-based models for forecasting economic and financial time series data. Copula-based multivariate models allow the researcher to specify the models for the marginal distributions separately from the dependence structure (copula) that links these distributions to form the joint distribution. This frees the researcher from considering only existing multivariate distributions, and allows for a much greater degree of flexibility in specifying the model. In some applications estimation can also be done in stages, with the marginal distributions estimated separately from the dependence structure, facilitating the study of high-dimension multivariate problems.

All theoretical methods reviewed in this chapter are applied to a representative data set of daily returns on two equity indices, and detailed discussion of methods for estimation, inference, goodness-of-fit testing, and model selection that are useful when working with copula-based models is provided. While the main ideas in copula theory are not hard, they may initially appear foreign. One objective of this chapter is to lower the "entry costs" of understanding and applying copula methods for economic time series.

To fix ideas, let us first recall a key result in this literature due to Sklar (1959), which states that an n-dimensional joint distribution can be decomposed into its n univariate marginal distributions and an n-dimensional copula:

$$\begin{aligned} \text{Let} \quad \mathbf{Y} &\equiv [Y_1, ..., Y_n]' \sim \mathbf{F}, \ \text{ with } Y_i \sim F_i \\ \text{then} \quad \exists\, \mathbf{C} &: [0,1]^n \rightarrow [0,1] \\ \text{s.t.} \quad \mathbf{F}(\mathbf{y}) &= \mathbf{C}(F_1(y_1), ..., F_n(y_n)) \ \forall\, \mathbf{y} \in \mathbb{R}^n \end{aligned} \tag{1}$$

Thus the copula $\mathbf{C}$ of the variable $\mathbf{Y}$ is the function that maps the univariate marginal distributions F_i to the joint distribution $\mathbf{F}$. Another interpretation of a copula function is possible using the "probability integral transformation", $U_i \equiv F_i(Y_i)$. As Casella and Berger (1990) note, when F_i is continuous the variable U_i will have the $Unif(0,1)$ distribution regardless of the original distribution F_i:

$$U_i \equiv F_i(Y_i) \sim Unif(0,1), \quad i = 1, 2, .., n \tag{2}$$

The copula $\mathbf{C}$ of $\mathbf{Y} \equiv [Y_1, ..., Y_n]'$ can be interpreted as the joint distribution of the vector of probability integral transforms, $\mathbf{U} \equiv [U_1, ..., U_n]'$, and thus is a joint distribution function with $Unif(0,1)$

margins. Notice that, when the densities exist, the above representation of the joint *cdf* implies the following representation for the joint *pdf:*

$$\begin{aligned} \mathbf{f}\left(y_1,...,y_n\right) &= \mathbf{c}\left(F_1\left(y_1\right),...,F_n\left(y_n\right)\right) \times \prod_{i=1}^{n} f_i\left(y_i\right) \\ \text{where} \quad \mathbf{c}\left(u_1,...,u_n\right) &= \frac{\partial^n \mathbf{C}\left(u_1,...,u_n\right)}{\partial u_1 \cdot ... \cdot \partial u_n} \end{aligned} \tag{3}$$

What makes this representation particularly useful for empirical research is the converse of Sklar's theorem: given any set of n univariate distributions $(F_1,...,F_n)$ and any copula $\mathbf{C}$, the function $\mathbf{F}$ defined by equation (1) above defines a valid joint distribution with marginal distributions $(F_1,...,F_n)$. For example, one might combine a Normally distributed variable with an Exponentially distributed variable via a t copula, and obtain a strange but valid bivariate distribution. The ability to combine marginal distributions with a copula model allows the researcher to draw on the large body of research on modeling univariate distributions, leaving "only" the task of modelling the dependence structure.

This chapter will focus exclusively on multivariate forecasting problems using copula-based models, and exclude univariate copula-based models, such as those considered by Darsow, *et al.* (1992), Ibragimov (2009), Beare (2010), Chen and Fan (2006a) and Chen, *et al.* (2009) for example. While univariate copula-based time series models are indeed interesting, from a forecasting perspective they are essentially a particular type of nonlinear time series model, a topic covered in chapters by White (2006) and Teräsvirta (2006) in the first edition of this Handbook.

In multivariate forecasting problems we will be interested in a version of Sklar's theorem for conditional joint distributions presented in Patton (2006a), where we consider some information set $\mathcal{F}_{t-1}$, and decompose the conditional distribution of $\mathbf{Y}_t$ given $\mathcal{F}_{t-1}$ into its conditional marginal distributions and the conditional copula:

$$\begin{aligned} \text{Let} \quad \mathbf{Y}_t|\mathcal{F}_{t-1} &\sim \mathbf{F}\left(\cdot|\mathcal{F}_{t-1}\right) \\ \text{with} \quad Y_{it}|\mathcal{F}_{t-1} &\sim F_i\left(\cdot|\mathcal{F}_{t-1}\right), \quad i=1,2,...,n \\ \text{then} \quad \mathbf{F}\left(\mathbf{y}|\mathcal{F}_{t-1}\right) &= \mathbf{C}\left(F_1\left(y_1|\mathcal{F}_{t-1}\right),...,F_n\left(y_n|\mathcal{F}_{t-1}\right)|\mathcal{F}_{t-1}\right) \end{aligned} \tag{4}$$

If we define the (conditional) probability integral transform variables, $U_{it} = F_i\left(Y_{it}|\mathcal{F}_{t-1}\right)$, then the conditional copula of $\mathbf{Y}_t|\mathcal{F}_{t-1}$ is just the conditional distribution of $\mathbf{U}_t|\mathcal{F}_{t-1}$:

$$\mathbf{U}_t|\mathcal{F}_{t-1} \sim \mathbf{C}\left(\cdot|\mathcal{F}_{t-1}\right) \tag{5}$$

This highlights the potential for copula-based models to facilitate specification and estimation in stages: one can estimate models for each of the conditional marginal distributions, $F_i\left(\cdot|\mathcal{F}_{t-1}\right)$, construct the probability integral transform variables, and then consider copula models for the joint distribution of these variables. This results in a valid n-dimensional model, without the challenge of specifying and estimating it simultaneously.

Note in equation (4) that the *same* information appears in each of the marginals and the copula.[1] However, in empirical applications it may be the case that not every part of $\mathcal{F}_{t-1}$ is needed for every marginal distribution. For example, let $\mathcal{F}_{t-1}^{(i)}$ denote the information set generated by $\left(Y_{i,t-1}, Y_{i,t-2}, ...\right)$, and let $\mathcal{F}_{t-1}$ denote the information set generated by $\left(\mathbf{Y}_{t-1}, \mathbf{Y}_{t-2}, ...\right)$. For some processes we may find that $Y_{it}|\mathcal{F}_{t-1} \overset{d}{=} Y_{it}|\mathcal{F}_{t-1}^{(i)}$, i.e., processes where each variable depends only upon its own lags and not on lags of other variables. Thus it is possible to use models for marginal distributions that do not explicitly use the entire information set, but still satisfy the restriction that all margins and the copula use the same information set.

For inference on copula parameters and related quantities, an important distinction arises between fully parametric multivariate models (where the copula and the marginal distributions are all parametric) and semiparametric models (where the copula is parametric and the marginal distributions are nonparametric). The latter case has much empirical appeal, but slightly more involved methods for inference are required. We will review and implement methods for both parametric and semiparametric copula-based multivariate models.[2]

Several other surveys of copula theory and applications have appeared in the literature to date: Nelsen (2006) and Joe (1997) are two key text books on copula theory, providing clear and detailed introductions to copulas and dependence modelling, with an emphasis on statistical foundations. Frees and Valdez (1998) present an introduction to copulas for actuarial problems. Cherubini, *et al.* (2004) present an introduction to copulas using methods from mathematical finance, and McNeil, *et al.* (2005) present an overview of copula methods in the context of risk management. Genest and Favre (2007) present a description of semiparametric inference methods for *iid* data with a detailed

[1] When different information sets are used, the resulting function $\mathbf{F}\left(\cdot|\cdot\right)$ is *not* generally a joint distribution with the specified conditional marginal distributions, see Fermanian and Wegkamp (2012).

[2] Forecasts based on nonparametric estimation of copulas are not common in the economics literature, and we will not consider this case in this chapter. Related articles include Genest and Rivest (1993) and Capéraà, *et al.* (1997) for *iid* data, and Fermanian and Scaillet (2003), Fermanian, *et al.* (2004), Sancetta and Satchell (2004) and Ibragimov (2009) for time series data.

empirical illustration. Patton (2009a) presents a summary of applications of copulas to financial time series and an extensive list of references. Choros, *et al.* (2010) provide a concise survey of estimation methods, both parametric and nonparametric, for copulas for both *iid* and time series data. Manner and Reznikova (2011) present a survey specifically focused on time-varying copula models, and Patton (2012) provides a brief review of the literature on copula-based methods for univariate and multivariate time series.

This chapter will focus on the key steps in using a copula-based model for economic forecasting, and the outline of this chapter will follow these steps. In Section 2 we consider some dependence summary statistics, which are useful for describing the data and for making initial decisions on the types of copula models that may be useful for a given data set. In Section 3 we look at estimation and inference for copula models, covering both fully parametric and semiparametric models. In Section 4 we review model selection and goodness-of-fit tests that are applicable for copula-based models, and in Section 5 we look at some issues that arise in economic applications of copula based models, such as extracting linear correlation coefficients from a copula-based model and computing portfolio Value-at-Risk. Finally, in Section 6 we survey some of the many applications of copulas in economics and finance, and in Section 7 we discuss directions for future research in this area.

1.1 Empirical illustration: small cap and large cap equity indices

To illustrate the methods presented in this chapter, we consider the daily returns on two equity indices: the S&P 100 index of the largest U.S. firms (covering about 60% of total market capitalization), and the S&P 600 index of small firms (covering about 3% of market capitalization). The sample period is 17 August 1995 (the start date for the S&P 600 index) until 30 May 2011, which covers 3639 trading days. A time series plot of these two series over this sample period is presented in the upper panel of Figure 1, and a scatter plot of these returns is presented in the lower panel of Figure 1. Summary statistics for these returns are presented in Table 1.[3]

[INSERT FIGURE 1 AND TABLE 1 ABOUT HERE]

Before modelling the dependence structure between these two return series, we must first model

[3] Matlab code to replicate the analysis in this chapter is available at `http://econ.duke.edu/∼ap172/code.html`.

their conditional marginal distributions.[4] We will base our model on the following structure:

$$\begin{aligned} Y_{it} &= \mu_i\left(\mathbf{Z}_{t-1}\right) + \sigma_i\left(\mathbf{Z}_{t-1}\right)\varepsilon_{it}, \text{ for } i = 1,2, \text{ where } \mathbf{Z}_{t-1} \in \mathcal{F}_{t-1} \\ \varepsilon_{it}|\mathcal{F}_{t-1} &\sim F_i\left(0,1\right) \ \forall\ t \end{aligned} \tag{6}$$

That is, we will allow each series to have potentially time-varying conditional mean and variance, and we will assume that the standardized residual, ε_{it}, has a constant conditional distribution (with mean zero and variance one, for identification).[5]

Using the Bayesian Information Criterion (BIC) and considering ARMA models for the conditional mean up to order $(5,5)$, the optimal models were found to be an AR(2) for the S&P 100 and an AR(0) (i.e., just a constant) for the S&P 600. Testing for the significance of five lags of the "other" series, conditional on these models, yielded p-values of 0.13 and 0.34, indicating no evidence of significant cross-equation effects in the conditional mean. Again using the BIC and considering volatility models in the GJR-GARCH class, see Glosten, *et al.* (1993), of up to order (2,2), the optimal models for both series were of order (1,1). Using these models we construct the estimated standardized residuals as:

$$\hat{\varepsilon}_{it} \equiv \frac{Y_{it} - \mu_i\left(\mathbf{Z}_{t-1};\hat{\alpha}\right)}{\sigma_i\left(\mathbf{Z}_{t-1};\hat{\alpha}\right)}, \ \ i = 1,2 \tag{7}$$

where $\hat{\alpha}$ is the vector of estimated parameters for the models for the conditional mean and conditional variance.

We will consider both parametric and nonparametric models for F_i. Many choices are possible for the parametric model for F_i, including the Normal, the standardized Student's t (as in Bollerslev, 1987), the skewed t (as in Patton, 2004), and others. In this chapter we use the simple and flexible skewed t distribution of Hansen (1994), see Jondeau and Rockinger (2003) for further results on this distribution. This distribution has two "shape" parameters: a skewness parameter, $\lambda \in (-1,1)$, which controls the degree of asymmetry, and a degrees of freedom parameter $\nu \in (2,\infty]$ which controls the thickness of the tails. When $\lambda = 0$ we recover the standardized Student's t distribution,

[4] Modelling the dependence structure of the variables directly, using the *unconditional* probability transform variables, yields a model for the *unconditional* copula of the returns. This may be of interest in some applications, but in forecasting problems we almost certainly want to condition on the available information, and thus are lead to study the *conditional* copula, which requires specifying models for the conditional marginal distributions.

[5] When parametric models are considered for F_i it is possible to allow for this distribution to vary through time, see Patton (2004) for one example, but we will not consider this here for simplicity.

when $\nu \to \infty$ we obtain a skewed Normal distribution, and when $\nu \to \infty$ and $\lambda = 0$ we obtain the $N(0,1)$ distribution. For the nonparametric estimate of F_i we will use the empirical distribution function (EDF)[6]:

$$\hat{F}_i(\varepsilon) \equiv \frac{1}{T+1} \sum_{t=1}^{T} \mathbf{1}\{\hat{\varepsilon}_{it} \leq \varepsilon\} \tag{8}$$

Table 1 presents the estimated parameters of the skewed t distribution, and Figure 2 presents the fitted parametric estimates of this distribution. The upper panel shows that the fitted density appears to provide a reasonable fit to the empirical histogram. The lower panel presents a QQ plot, and reveals that a few extreme left tail observations are not captured by the models for each series.

[INSERT FIGURE 2 ABOUT HERE]

The lower rows of Table 1 report p-values from a test of the goodness-of-fit of the skewed t distribution using both the Kolmogorov-Smirnov (KS) and Cramer-von Mises (CvM) test statistics:

$$KS_i = \max_t \left| \hat{U}_{i,(t)} - \frac{t}{T} \right| \tag{9}$$

$$CvM_i = \sum_{t=1}^{T} \left(\hat{U}_{i,(t)} - \frac{t}{T} \right)^2 \tag{10}$$

where $\hat{U}_{i,(t)}$ is the t^{th} largest value of $\left\{\hat{U}_{i,j}\right\}_{j=1}^{T}$, i.e., the t^{th} order statistic of $\left\{\hat{U}_{i,j}\right\}_{j=1}^{T}$. Both of these test statistics are based on the estimated probability integral transformations:

$$\hat{U}_{it} \equiv F_{skew\ t}\left(\hat{\varepsilon}_{it}; \hat{\nu}_i, \hat{\lambda}_i\right) \tag{11}$$

In the absence of parameter estimation error, the KS and CvM test statistics have asymptotic distributions that are known, however the presence of estimated parameters in our model means that those distributions are not applicable here. To overcome this we exploit the fact that with parametric models for the mean, variance and error distribution we have completely characterized the conditional distribution, and thus can use a simple simulation-based method to obtain critical values (see Genest and Rémillard (2008) for example): (i) Simulate T observations for Y_{it} from

[6] Note that this definition of the EDF scales by $1/(T+1)$ rather than $1/T$, as is common in this literature. This has no effect asymptotically, and in finite samples is useful for keeping the estimated probability integral transforms away from the boundaries of the unit interval, where some copula models diverge.

this model using the estimated parameters, (ii) Estimate the models on the simulated data, (iii) Compute the KS and CvM statistics on the estimated probability integral transforms of the simulated data, (iv) Repeat steps (i)-(iii) S times (e.g., $S = 1000$), (v) Use the upper $1 - \alpha$ quantile of $\left\{\left(KS_{(s)}, CvM_{(s)}\right)\right\}_{s=1}^{S}$ as the critical value for these tests.

Implementing these tests on the S&P 100 and S&P 600 standardized residuals, we find p-values for the KS (CvM) tests of 0.12 and 0.09 (0.48 and 0.22), and thus fail to reject the null that the skew t model is well-specified for these two return series. This provides support for these models of the marginal distributions, allowing us to move on to modelling the copula.

2 Dependence summary statistics

When assuming Normality, the only relevant summary statistic for the dependence structure is the linear correlation coefficient, and this is routinely reported in empirical work on multivariate time series. However, when considering more flexible models for the dependence structure we need to also consider other measures of dependence, to provide some guidance on the types of models that might be suitable for the variables under analysis. This section describes some useful dependence measures and methods for conducting inference on estimates of these measures.

2.1 Measures of dependence

Numerous dependence measures exist in the literature, see Nelsen (2006, Chapter 5) and Joe (1997, Chapter 2) for detailed discussions. A key attribute of a dependence measure for providing guidance on the form of the copula is that it should be a "pure" measure of dependence (or "scale invariant", in the terminology of Nelsen 2006), and so should be unaffected by strictly increasing transformations of the data. This is equivalent to imposing that the measure can be obtained as a function of the *ranks* (or probability integral transforms) of the data only, which is in turn equivalent to it being a function solely of the copula, and not the marginal distributions. Linear correlation is *not* scale invariant (e.g., $Corr\left[X, Y\right] \neq Corr\left[\exp\left\{X\right\}, \exp\left\{Y\right\}\right]$) and is affected by the marginal distributions of the data. Given its familiarity in economics, it is still a useful measure to report, but we will augment it with other measures of dependence.

Firstly, we recall the definition of Spearman's rank correlation. We will denote the population rank correlation as ϱ and sample rank correlation as $\hat{\varrho}$:

$$\varrho = Corr\left[U_{1t}, U_{2t}\right] = 12E\left[U_{1t}U_{2t}\right] - 3 = 12\int_0^1\int_0^1 uv d\mathbf{C}\left(u,v\right) - 3 \tag{12}$$

$$\hat{\varrho} = \frac{12}{T}\sum_{t=1}^{T} U_{1t}U_{2t} - 3 \tag{13}$$

(Note that this formula exploits the fact that $E\left[U\right] = 1/2$ and $V\left[U\right] = 1/12$ for $U \sim Unif\left(0,1\right)$.) Rank correlation is constrained to lie in $\left[-1,1\right]$, with the bounds of this interval being attained only when one variable is a strictly increasing or decreasing function of the other. Rank correlation is useful for providing information on the *sign* of the dependence between two variables, which is important when considering copula models that can only accommodate dependence of a given sign (such as some Archimedean copulas).

We next consider "quantile dependence", which measures the strength of the dependence between two variables in the joint lower, or joint upper, tails of their support. It is defined as

$$\lambda^q = \begin{cases} \Pr\left[U_{1t} \le q | U_{2t} \le q\right], & 0 < q \le 1/2 \\ \Pr\left[U_{1t} > q | U_{2t} > q\right], & 1/2 < q < 1 \end{cases} \tag{14}$$

$$= \begin{cases} \frac{\mathbf{C}(q,q)}{q}, & 0 < q \le 1/2 \\ \frac{1-2q+\mathbf{C}(q,q)}{1-q}, & 1/2 < q < 1 \end{cases}$$

$$\hat{\lambda}^q = \begin{cases} \frac{1}{Tq}\sum_{t=1}^{T}\mathbf{1}\left\{U_{1t} \le q, U_{2t} \le q\right\}, & 0 < q \le 1/2 \\ \frac{1}{T(1-q)}\sum_{t=1}^{T}\mathbf{1}\left\{U_{1t} > q, U_{2t} > q\right\} & 1/2 < q < 1 \end{cases} \tag{15}$$

Quantile dependence provides a richer description of the dependence structure of two variables.[7] By estimating the strength of the dependence between the two variables as we move from the center $(q = 1/2)$ to the tails, and by comparing the left tail $(q < 1/2)$ to the right tail $(q > 1/2)$ we are provided with more detailed information about the dependence structure than can be provided by a scalar measure like linear correlation or rank correlation. Information on the importance of asymmetric dependence is useful as many copula models, such as the Normal and the Student's t copulas, impose symmetric dependence.

Tail dependence is a measure of the dependence between extreme events, and population tail

[7] The definition given here is tailored to positively dependent variables, as it traces out the copula along the main diagonal, $\mathbf{C}\left(q,q\right)$ for $q \in \left(0,1\right)$. It is easily modified to apply to negatively dependent variables, by considering $\mathbf{C}\left(q,1-q\right)$ and $\mathbf{C}\left(1-q,q\right)$.

dependence can be obtained as the limit of population quantile dependence as $q \to 0$ or $q \to 1$:

$$\begin{aligned} \lambda^L &= \lim_{q \to 0^+} \frac{\mathbf{C}(q,q)}{q} \\ \lambda^U &= \lim_{q \to 1^-} \frac{1-2q+\mathbf{C}(q,q)}{1-q} \end{aligned} \quad (16)$$

Sample tail dependence *cannot* simply be taken as $\hat{\lambda}^L = \lim_{q \to 0^+} \hat{\lambda}^q$, since if we set q close enough to zero we are assured that the estimate will be zero. (For example, if we use the EDF to estimate the marginal distributions, then any value of $q < 1/T$ or $q > 1 - 1/T$ will result in $\hat{\lambda}^q = 0$.) Thus estimating tail dependence from a finite sample of data must be done using an alternative approach.

Unlike the extreme tails of a univariate distribution, which under general conditions can be shown using extreme value theory to follow a functional form with just one or two free parameters, the tails of a bivariate distribution require the estimation of an unknown univariate function known as "Pickand's (1981) dependence function". It can be shown, see Frahm, *et al.* (2005), that estimating the upper and lower tail dependence coefficients is equivalent to estimating the value of the Pickand's dependence function at one-half. One simple nonparametric estimator of tail dependence considered in Frahm, *et al.* (2005) is the "log" estimator:

$$\begin{aligned} \hat{\lambda}^L &= 2 - \frac{\log\left(1-2(1-q^*)+T^{-1}\sum_{t=1}^{T}\mathbf{1}\left\{U_{1t} \le 1-q^*, U_{2t} \le 1-q^*\right\}\right)}{\log(1-q^*)} \quad \text{for } q^* \approx 0 \quad (17) \\ \hat{\lambda}^U &= 2 - \frac{\log\left(T^{-1}\sum_{t=1}^{T}\mathbf{1}\left\{U_{1t} \le 1-q^*, U_{2t} \le 1-q^*\right\}\right)}{\log(1-q^*)} \quad \text{for } q^* \approx 0 \end{aligned}$$

As usual for extreme value estimation, a threshold q^* needs to be chosen for estimation, and it can differ for the upper and lower tail. This choice involves trading off the variance in the estimator (for small values of q) against bias (for large values of q), and Frahm, *et al.* (2005) suggest a simple method for making this choice.[8] Information on the importance of tail dependence is useful as many copula models, such as the Normal and Frank copulas, impose zero tail dependence, and other copulas impose zero tail dependence in one of their tails (e.g., right for the Clayton copula and left for the Gumbel copula).

[8] Alternatively, one can specify and estimate parametric copulas for the joint upper and lower tails, and infer the tail dependence coefficients from the fitted models. This approach is discussed in Section 3.4.1 below.

2.2 Inference on measures of dependence

In addition to estimating dependence summary statistics, it is often of interest to obtain standard errors on these, either to provide an idea of the precision with which these parameters are estimated, or to conduct tests on these (we will consider tests for asymmetric dependence and for time-varying dependence below). If the data to be analyzed was known to already have $Unif\,(0,1)$ margins, then inference is straightforward, however in general this is not the case, and the data on which we compute the dependence summary statistics will usually depend on parameters estimated in an earlier part of the analysis. (For example, on ARMA models for the mean, GARCH models for the variance, and possibly shape parameters for the density of the standardized residuals.) The method for inference on the estimated dependence statistics is different depending on whether a parametric or a nonparametric model is used for the distribution of the standardized residuals. The methods described below are closely related to inference methods for estimated copula parameters, which are discussed in Section 3.

2.2.1 Parametric marginal distributions

Combining parametric marginal distributions for the standardized residuals with parametric models for the conditional means and variances yields a fully parametric model for the conditional marginal distributions. Inference on the estimated dependence statistics can be conducted in one of (at least) two ways. Firstly, one could treat this as multi-stage GMM, where the "moments" of all stages except for the estimation of the dependence statistics, are the scores of the marginal log-likelihoods (i.e., these are all maximum likelihood estimators), and the latter are the moments (or "estimating equations") that generate $\hat{\varrho}$, $\hat{\lambda}^q$, $\hat{\lambda}^L$ and $\hat{\lambda}^U$ as solutions. This is a minor adaptation of the methods in Patton (2006b), who considered multi-stage maximum likelihood estimation (MLE) for copula-based models of multivariate time series. We consider this method in detail in Section 3.1 below.

A second, simpler, approach based on a bootstrap may be desirable to avoid having to compute the moments outlined above: (i) Use the stationary bootstrap of Politis and Romano (1994), or another bootstrap method that preserves (at least asymptotically) the time series dependence in the data, to generate a bootstrap sample[9] of the data of length T. (ii) Estimate the model on

[9] It is important to maintain the cross-sectional dependence of the data, and so this shuffle should be done on entire rows of the matrix of standardized residuals, assuming that these are stored in a $T \times n$ matrix, and not separately for each series.

the simulated data, (iii) Compute the dependence measures on the estimated probability integral transformations, (iv) Repeat steps (i)-(iii) S times (e.g., $S = 1000$), (v) Use the $\alpha/2$ and $1 - \alpha/2$ quantiles of the simulated distribution of $\left\{ \left(\hat{\varrho}_i, \hat{\lambda}_i^q, \hat{\lambda}_i^L, \hat{\lambda}_i^U \right) \right\}_{i=1}^{S}$ to obtain a $1-\alpha$ confidence interval for these parameters. See Gonçalves and White (2004) for results on the bootstrap for nonlinear and serially dependent processes.

2.2.2 Nonparametric marginal distributions

Using the empirical distribution function (EDF), or some other nonparametric estimate, of the distributions for the standardized residuals with parametric models for the conditional means and variances makes the model semiparametric. As in the fully parametric case, inference on the estimated dependence statistics can be conducted either using the asymptotic distribution of the parameters of the model (including the infinite-dimensional marginal distributions) or using a bootstrap approach. Both of these approaches are based on the assumption that the underlying true conditional copula is *constant* through time.

Similar to the parametric case, in the first approach one treats this as multi-stage *semiparametric* GMM, where the "moments" of all stages except for the estimation of the dependence statistics, are the scores of the log-likelihood (i.e., these are all ML), and the latter are the moments that generate $\hat{\varrho}$, $\hat{\lambda}^q$, $\hat{\lambda}^L$ and $\hat{\lambda}^U$ as solutions. This is a minor adaptation of the methods in Chen and Fan (2006b), who considered multi-stage MLE for semiparametric copula-based models of multivariate time series. A key simplification of this approach, relative to the fully parametric case, is that the estimated parameters of the models for the conditional mean and variance do *not* affect the asymptotic distribution of the dependence statistics, see Rémillard (2010). This is a surprising result. Thus, in this semiparametric case and under the assumption of a constant conditional copula, one can ignore the estimation of the mean and variance models. The asymptotic distribution *does* depend on the estimation error coming from the use of the EDF, making the asymptotic variance different from standard MLE. We will discuss this method in detail in Section 3.2 below.

A second approach again exploits the bootstrap to obtain confidence intervals, and is simple to implement. Following Chen and Fan (2006b) and Rémillard (2010), we can treat the estimated standardized residuals as though they are the true standardized residuals (i.e., we can ignore the presence of estimation error in the parameters of the models for the conditional mean and variance), and under the assumption that the conditional copula is constant we can then use a simple *iid* bootstrap approach: (i) Randomly draw rows, with replacement, from the $T \times n$ matrix of standardized residuals until a bootstrap sample of length T is obtained, (ii) Estimate the dependence measures of the bootstrap sample, (iii) Repeat steps (i)-(ii) S times, (iv) Use the $\alpha/2$ and $1-\alpha/2$ quantiles of the simulated distribution of $\left\{\left(\hat{\varrho}_i, \hat{\lambda}_i^q, \hat{\lambda}_i^L, \hat{\lambda}_i^U\right)\right\}_{i=1}^S$ to obtain a $1-\alpha$ confidence interval for these parameters. Given how simple it is to compute the dependence statistics discussed above, this bootstrap approach is fast and convenient relative to one that relies on the asymptotic distribution of these statistics.

When the conditional copula is time-varying, the parameter estimation error from the models for the conditional mean and variance *cannot*, in general, be ignored, see Rémillard (2010), and so the above multi-stage GMM or *iid* bootstrap approaches are not applicable. Methods for conducting inference on the above parameters that are robust to time variation in the conditional copula are not yet available, to my knowledge. A potential method to overcome this is as follows: If the dynamics of the conditional copula (and conditional means and variances) are such that the serial dependence of the process can be replicated by a block bootstrap, then the approach used for fully parametric models may be suitable: (i) Use the a block bootstrap (e.g., that of Politis and Romano (1994)) to generate a bootstrap sample of the original data of length T, (ii) Estimate the conditional mean and variance models on the bootstrap sample, (iii) Compute the dependence measures on the estimated standardized residuals, (iv) Repeat steps (i)-(iii) S times, (v) Use the $\alpha/2$ and $1-\alpha/2$ quantiles of the simulated distribution of $\left\{\left(\hat{\varrho}_i, \hat{\lambda}_i^q, \hat{\lambda}_i^L, \hat{\lambda}_i^U\right)\right\}_{i=1}^S$ to obtain a $1-\alpha$ confidence interval for these parameters.[10]

2.3 Empirical illustration, continued

Using the small-cap and large-cap equity index return data and marginal distribution models described in Section 1.1, we now examine their dependence structure. The rank correlation between these two series is estimated at 0.782, and an 90% *iid* bootstrap confidence interval is $[0.769, 0.793]$. Thus the dependence between these two series is positive and relatively strong. The upper panel of Figure 3 presents the estimated quantile dependence plot, for $q \in [0.025, 0.975]$, along with 90%

[10] Gaißer, *et al.* (2010) suggest a block bootstrap to conduct inference on dependence measures for serially dependent data, and it is possible that this approach may be combined with the results of Rémillard (2010) to justify the inference method outlined here, however this has not been considered in the literature to date. Other work on related problems include Genest and Rémillard (2008) and Ruppert (2011).

(pointwise) *iid* bootstrap confidence intervals, and the lower panel presents the difference between the upper and lower portions of this plot, along with a pointwise confidence interval for this difference. As expected, the confidence intervals are narrower in the middle of the distribution (values of q close to $1/2$) and wider near the tails (values of q near 0 or 1).

This figure shows that observations in the lower tail are somewhat more dependent than observations in the upper tail, with the difference between corresponding quantile dependence probabilities being as high as 0.1. The confidence intervals show that these differences are borderline significant at the 0.10 level, with the upper bound of the confidence interval on the difference lying around zero for most values of q. We present a joint test for asymmetric dependence in the next section.

Figure 3 also presents estimates of the upper and lower tail dependence coefficients. These are based on the estimator in equation (17), using the method in Frahm, *et al.* (2005) to choose the threshold. The estimated lower tail dependence coefficient is 0.411 with a 90% bootstrap confidence interval of [0.112,0.664]. The upper tail dependence coefficient is 0.230 with confidence interval [0.022,0.529]. Thus we can reject the null of zero tail dependence for both the upper and lower tails.

[INSERT FIGURE 3 ABOUT HERE]

2.4 Asymmetric dependence

With an estimated quantile dependence function, and a method for obtaining standard errors, it is then possible to test for the presence of asymmetric dependence. This can provide useful guidance on the types of parametric copulas to consider in the modelling stage. A simple test for asymmetric dependence can be obtained by noting that under symmetric dependence we have:

$$\lambda^q = \lambda^{1-q} \ \forall \ q \in [0,1] \tag{18}$$

Testing this equality provides a test of a necessary but not sufficient condition for symmetric dependence. Rather than test each q separately, and run into the problem of interpreting a set of multiple correlated individual tests, it is desirable to test for asymmetry *jointly.* Stack the estimated quantile dependence measures into a vector of the form[11]:

$$\begin{aligned}\hat{\boldsymbol{\lambda}} &\equiv [\lambda^{q_1}, \lambda^{q_2}, ..., \lambda^{q_{2p}}]' \\ \text{where } \; q_{p+j} &= 1 - q_j, \;\; \text{for } j = 1, 2, ..., p \end{aligned} \tag{19}$$

and then test:

$$\begin{aligned} H_0 &: R\boldsymbol{\lambda} = 0 \quad \text{vs.} \quad H_a : R\boldsymbol{\lambda} \neq 0 \\ \text{where } \; R &\equiv \left[I_p \vdots - I_p \right] \end{aligned} \tag{20}$$

Using the fact that $\sqrt{T}\left(\hat{\boldsymbol{\lambda}} - \boldsymbol{\lambda}\right) \xrightarrow{d} N\left(0, V_\lambda\right)$ from Rémillard (2010), and a bootstrap estimate of V_λ, denoted $\hat{V}_{\lambda,S}$, we can use that under H_0 :

$$T\left(\hat{\boldsymbol{\lambda}} - \boldsymbol{\lambda}\right)' R' \left(R\hat{V}_{\lambda,S}R'\right)^{-1} R\left(\hat{\boldsymbol{\lambda}} - \boldsymbol{\lambda}\right) \xrightarrow{d} \chi^2_p \tag{21}$$

Implementing this test on the estimated quantile dependence function for the S&P 100 and S&P 600 standardized residuals, with $q \in \{0.025, 0.05, 0.10, \; 0.975, 0.95, 0.90\}$ yields a chi-squared statistic of 2.54, which corresponds to a p-value of 0.47, thus we fail to reject the null that the dependence structure is symmetric using this metric.

Of particular interest in many copula studies is whether the tail dependence coefficients (i.e., the limits of the quantile dependence functions) are equal. That is, a test of

$$H_0 : \lambda^L = \lambda^U \quad \text{vs.} \quad H_a : \lambda^L \neq \lambda^U \tag{22}$$

Using the estimates and bootstrap inference methods from the previous section this is simple to implement. As noted above, the estimated tail dependence coefficients are $\hat{\lambda}^L = 0.411$ and $\hat{\lambda}^U = 0.230$. The bootstrap p-value for this difference is 0.595, indicating no significant difference in the upper and lower tail dependence coefficients.

2.5 Time-varying dependence

There is an abundance of evidence that the conditional volatility of economic time series changes through time, see Andersen, *et al.* (2006) for example, and thus reason to think that the conditional

[11] An alternative to considering a finite number of values of q would be to consider λ as a function of all $q \in (0, 1)$. This is feasible, but with a more complicated limiting distribution, and we do not pursue this here.

dependence structure may also vary through time. For example, Figure 4 presents a time series plot of rolling 60-day rank correlation, along with pointwise bootstrap standard errors (correct only under the null that this correlation is not changing). This figure shows that the rank correlation hovered around 0.6-0.7 in the early part of the sample, rising to around 0.9 during the financial crisis of 2008-09.

[INSERT FIGURE 4 ABOUT HERE]

Before specifying a functional form for a time-varying conditional copula model, it is informative to test for the presence of time-varying dependence. The tests we will consider maintain a constant conditional copula under the null, and thus the results from Rémillard (2010) may be used here to obtain the limiting distribution of the test statistics we consider.

There are numerous ways to test for time-varying dependence. We will focus here on tests that look for changes in rank correlation, ϱ, both for the ease with which such tests can be implemented, and the guidance they provide for model specification.[12] The rank correlation measure associated with $\mathbf{C}_t$ will be denoted ϱ_t.

We will consider three types of tests for time-varying dependence. The first test is a simple test for a break in rank correlation at some specified point in the sample, t^*. Under the null, the dependence measure before and after this date will be equal, while under the alternative they will differ:

$$\begin{aligned} H_0 \quad &: \quad \varrho_1 = \varrho_2 \quad \text{vs.} \quad H_a : \varrho_1 \neq \varrho_2 \\ \text{where} \quad \varrho_t \quad &= \quad \begin{cases} \varrho_1, & t \leq t^* \\ \varrho_2, & t > t^* \end{cases} \end{aligned} \tag{23}$$

A critical value for $(\hat{\varrho}_1 - \hat{\varrho}_2)$ can be obtained by using the *iid* bootstrap described in Section 2.2.2, noting that by imposing *iid*-ness when generating the bootstrap samples we obtain draws that impose the null hypothesis. This test is simple to implement, but requires the researcher to have *a priori* knowledge of when a break in the dependence structure may have occurred. In some applications this is reasonable, see Patton (2006a) for one example, but in other cases the date of the break, if present, is not known.

[12] An alternative is to consider test statistics that look for changes any where in the copula, as in Rémillard (2010), which asymptotically will detect a greater variety of changes in the copula, but are harder to interpret and use in model specification, and may have lower power in finite samples.

A second test for time-varying dependence allows for a break in the rank correlation coefficient at some unknown date. As usual for these types of tests, we must assume that the break did not occur "too close" to the start or end of the sample period (so that we have sufficient observations to estimate the pre- and post-break parameter), and a common choice is to search for breaks in an interval $[t_L^*, t_U^*]$ where $t_L^* = \lceil 0.15T \rceil$ and $t_U^* = \lfloor 0.85T \rfloor$.[13] A variety of test statistics are available for these types of tests, see Andrews (1993), and a simple, popular statistic is the "sup" test

$$\hat{B}_{\text{sup}} = \max_{t^* \in [t_L^*, t_U^*]} \left| \hat{\varrho}_{1,t^*} - \hat{\varrho}_{2,t^*} \right| \tag{24}$$

$$\text{where} \quad \hat{\varrho}_{1,t^*} \equiv \frac{12}{t^*} \sum_{t=1}^{t^*} U_{1t} U_{2t} - 3 \tag{25}$$

$$\hat{\varrho}_{2,t^*} \equiv \frac{12}{T - t^*} \sum_{t=t^*+1}^{T} U_{1t} U_{2t} - 3$$

A critical value for $\hat{B}_{\text{sup}}$ can again be obtained by using the *iid* bootstrap described in Section 2.2.2.

A third test for time-varying dependence is based on the "ARCH LM" test for time-varying volatility proposed by Engle (1982). Rather than looking for discrete one-time breaks in the dependence structure, this test looks for autocorrelation in a measure of dependence, captured by an autoregressive-type model. For example, consider the following regression

$$U_{1t} U_{2t} = \alpha_0 + \sum_{i=1}^{p} \alpha_i U_{1,t-i} U_{2,t-i} + \epsilon_t \tag{26}$$

or a parsimonious version of this regression:

$$U_{1t} U_{2t} = \alpha_0 + \frac{\alpha_1}{p} \sum_{i=1}^{p} U_{1,t-i} U_{2,t-i} + \epsilon_t \tag{27}$$

Under the null of a constant conditional copula, we should find $\alpha_i = 0 \ \forall \ i \geq 1$, and this can be tested by forming the statistic

$$\hat{A}_p = \hat{\boldsymbol{\alpha}}' R' \left(R \hat{V}_\alpha R' \right)^{-1} R \hat{\boldsymbol{\alpha}}$$

$$\text{where} \quad \hat{\boldsymbol{\alpha}} \equiv [\alpha_0, ..., \alpha_p]'$$

$$R = \left[\mathbf{0}_{p \times 1} \vdots I_p \right]$$

[13] $\lceil a \rceil$ denotes the smallest integer greater than or equal to a, and $\lfloor b \rfloor$ denotes the largest integer smaller than or equal to b.

and using the usual OLS estimate of the covariance matrix for $\hat{V}_\alpha$. Critical values for this test statistic can again be obtained using the *iid* bootstrap described in Section 2.2.2.

Implementing these tests for time-varying dependence between the S&P 100 and S&P 600 standardized residuals yields results that are summarized in Table 2 below. Having no *a priori* dates to consider for the timing of a break, consider for illustration tests for a break at three points in the sample, at $t^*/T \in \{0.15, 0.50, 0.85\}$, which corresponds to the dates 23-Dec-1997, 7-July-2003, 8-Jan-2009. For the latter date evidence of a break in rank correlation is found, with a p-value of 0.045, while for the earlier two dates no evidence is present. Thus it appears that the rank correlation towards the end of the sample is different from that during the earlier part of the sample. However, given a lack of a reason for choosing a break date of 8-Jan-2009, a more appropriate test is one where the break date is estimated, and using that test the p-value is 0.269, indicating no evidence against a constant rank correlation in the direction of a one-time break.

The plot of rank correlation in Figure 4, and related evidence for relatively smoothly evolving conditional volatility of financial assets, suggests that if rank correlation is varying, it may be more in an autoregressive-type manner than as a discrete, one-time change. Using the AR specification for autocorrelation in $(U_{1t}U_{2t})$ described in equation (27), I find evidence of non-zero autocorrelation for lags 10 and 5, but no evidence at lag 1.

Thus, we can conclude that there is evidence against constant conditional rank correlation for the S&P100 and S&P 600 standardized residuals, and thus evidence against a constant conditional copula. Given the wealth of evidence that volatility changes through time, this is not overly surprising, but it provides a solid motivation for considering models of time-varying copulas.

[INSERT TABLE 2 ABOUT HERE]

3 Estimation and inference for copula models

This section covers inference on the parameters of copula-based multivariate models. A key motivation for obtaining the distribution of our parameter estimates is that the economic quantities of interest are functionals of the conditional distribution of $\mathbf{Y}_t$. For example, measures of dependence will be functions of the conditional copula (perhaps directly related to the copula parameters, perhaps not), and measures of risk will often be functions of both the copula and the marginal distributions. Understanding the estimation error in our model will enable us to derive the estimation error around the economic quantities of interest. Given their prevalence in the literature to date, we will focus on maximum likelihood estimation. Other estimation methods used in the literature are discussed in Section 3.3.

A majority of applications of copula models for multivariate time series build the model in stages, and that case is considered in detail here. We will assume that the conditional mean and variance are modelled using some parametric specification:

$$\begin{aligned} E\left[Y_{it}|\mathcal{F}_{t-1}\right] &\equiv \mu_i\left(\mathbf{Z}_{t-1}, \alpha^*\right), \quad \mathbf{Z}_{t-1} \in \mathcal{F}_{t-1} \\ V\left[Y_{it}|\mathcal{F}_{t-1}\right] &\equiv \sigma_i^2\left(\mathbf{Z}_{t-1}, \alpha^*\right) \end{aligned} \tag{28}$$

This assumption allows for a wide variety of models for the conditional mean: ARMA models, vector autoregressions, linear and nonlinear regressions, and others. It also allows for a variety of models for the conditional variance: ARCH and any of its numerous parametric extensions (GARCH, EGARCH, GJR-GARCH, etc., see Bollerslev, 2010), stochastic volatility models, and others. Note that $\mathcal{F}_{t-1}$ will in general include lags of all variables in $\mathbf{Y}_t$, not only lags of Y_{it}.

The standardized residuals are defined as:

$$\varepsilon_{it} \equiv \frac{Y_{it} - \mu_i\left(\mathbf{Z}_{t-1}, \alpha^*\right)}{\sigma_i\left(\mathbf{Z}_{t-1}, \alpha^*\right)} \tag{29}$$

The conditional distribution of ε_{it} is treated in one of two ways, either parametrically or nonparametrically. In the former case, this distribution may vary through time as a (parametric) function of $\mathcal{F}_{t-1}$-measurable variables (e.g., the time-varying skewed t distribution of Hansen, 1994), or may be constant. In the nonparametric case, we will follow the majority of the literature and assume that the conditional distribution is constant.

$$\varepsilon_{it}|\mathcal{F}_{t-1} \sim F_i\left(\cdot|\mathbf{Z}_{t-1}; \alpha^*\right) \tag{30}$$

$$\text{or} \quad \varepsilon_{it}|\mathcal{F}_{t-1} \sim iid\ F_i \tag{31}$$

For the identification of the parameters of the conditional mean and variance models, the distribution of ε_{it} must have zero mean and unit variance. The choice of a parametric or nonparametric model for the distribution of the standardized residuals leads to different inference procedures for the copula parameters, and we will treat these two cases separately below.

The conditional copula is the conditional distribution of the probability integral transforms of the standardized residuals. We will consider parametric copula models, and will consider both constant and time-varying cases:

$$\begin{aligned} U_{it} &\equiv F_i\left(\varepsilon_{it}\right), \quad i=1,2,...,n \\ \text{and }\ \mathbf{U}_t &\equiv \left[U_{1t},...,U_{nt}\right]' | \mathcal{F}_{t-1} \sim \begin{cases} iid\ \mathbf{C}\left(\gamma^*\right) \\ \mathbf{C}\left(\delta_t\left(\gamma^*\right)\right) \end{cases} \end{aligned} \tag{32}$$

where δ_t is the parameter of the copula $\mathbf{C}$, and its time series dynamics are governed by the parameter γ^*. In the constant case we have simply $\delta_t = \gamma^*\ \forall\ t$. The parameter for the entire model is $\theta^* \equiv \left[\alpha^{*\prime}, \gamma^{*\prime}\right]'$, with α^* containing all parameters related to the marginal distributions, and γ^* containing all parameters for the copula.

3.1 Parametric models

When all components of the multivariate model are parametric, the most natural estimation method is maximum likelihood: in writing down a fully parametric model for the conditional distribution of $\mathbf{Y}_t$, we have fully specified the likelihood.

$$\hat{\theta}_T = \arg\max_{\theta}\ \log \mathcal{L}_T\left(\theta\right) \tag{33}$$

$$\text{where} \quad \log \mathcal{L}_T\left(\theta\right) = \sum_{t=1}^{T} \log \mathbf{f}_t\left(\mathbf{Y}_t | \mathcal{F}_{t-1}; \theta\right) \tag{34}$$

$$\begin{aligned} \log \mathbf{f}_t\left(\mathbf{Y}_t | \mathcal{F}_{t-1}; \theta\right) &= \sum_{i=1}^{n} \log f_{it}\left(Y_{it} | \mathcal{F}_{t-1}; \alpha\right) \\ &\quad + \log \mathbf{c}\left(F_{1t}\left(Y_{1t} | \mathcal{F}_{t-1}; \alpha\right), ..., F_{nt}\left(Y_{nt} | \mathcal{F}_{t-1}; \alpha\right) | \mathcal{F}_{t-1}; \gamma\right) \end{aligned}$$

Under regularity conditions, see White (1994) for example[14], standard results for parametric time series models can be used to show that:

$$\sqrt{T}\left(\hat{\theta}_T - \theta^*\right) \xrightarrow{d} N\left(0, V_\theta^*\right) \quad \text{as} \quad T \to \infty \tag{35}$$

[14] For time-varying conditional copula models it can be difficult to establish sufficient conditions for stationarity, which is generally required for standard estimation methods to apply. Results for general classes of *univariate* nonlinear processes are presented in Carrasco and Chen (2002) and Meitz and Saikkonen (2008), however similar results for the multivariate case are not yet available. Researchers usually make these regularity conditions a high level assumption, and then use simulation results to provide some reassurance that these assumptions are plausible for the model(s) under consideration.

A consistent estimator of the asymptotic covariance matrix can also be obtained using standard methods:

$$\begin{aligned}\hat{V}_\theta &= \hat{A}_T^{-1}\hat{B}_T\hat{A}_T^{-1} \\ \text{where } \hat{B}_T &= \frac{1}{T}\sum_{t=1}^{T}\hat{\mathbf{s}}_t\hat{\mathbf{s}}_t' \quad \text{and} \quad \hat{A}_T = \frac{1}{T}\sum_{t=1}^{T}\hat{H}_t \\ \hat{\mathbf{s}}_t &= \frac{\partial}{\partial\theta}\log\mathbf{f}_t\left(\mathbf{Y}_t|\mathcal{F}_{t-1};\hat{\theta}_T\right) \\ \hat{H}_t &= \frac{\partial^2}{\partial\theta\partial\theta'}\log\mathbf{f}_t\left(\mathbf{Y}_t|\mathcal{F}_{t-1};\hat{\theta}_T\right)\end{aligned} \tag{36}$$

Under the assumption that the model is correctly specified, the "information matrix equality" holds, and so $B_0 = -A_0$, where $A_0 \equiv \lim_{T\to\infty}\hat{A}_T$ and $B_0 \equiv \lim_{T\to\infty}\hat{B}_T$. This means that we can alternatively estimate V_θ^* by $-\hat{A}_T^{-1}$ or by $\hat{B}_T^{-1}$. These estimators are all consistent for the true asymptotic covariance matrix:

$$\hat{V}_\theta - V_\theta^* \xrightarrow{p} 0 \quad \text{as} \quad T\to\infty \tag{37}$$

3.1.1 Multi-stage estimation of parametric copula-based models

In many applications the multivariate model is specified in such a way that the parameters can be estimated in separate stages. Such models require that the parameters that appear in the one marginal distribution do not also appear in another marginal distribution, and there are no cross-equation restrictions on these parameters. Standard models for the conditional mean (ARMA, VAR, etc.) satisfy this condition, as do most multivariate volatility models, with the notable exception of the BEKK model of Engle and Kroner (1995). If the parameters are indeed separable into parameters for the first margin, α_1, parameters for the second margin, α_2, etc., and parameters for the copula, γ, then the log-likelihood takes the form:

$$\begin{aligned}\sum_{t=1}^{T}\log\mathbf{f}_t\left(\mathbf{Y}_t;\theta\right) &= \sum_{t=1}^{T}\sum_{i=1}^{n}\log f_{it}\left(Y_{it};\alpha_i\right) \\ &\quad +\sum_{t=1}^{T}\log\mathbf{c}_t\left(F_{1t}\left(Y_{1t};\alpha_1\right),...,F_{nt}\left(Y_{nt};\alpha_n\right);\gamma\right)\end{aligned} \tag{38}$$

Maximizing the parameters separately for the margins and the copula is sometimes called "inference functions for margins", see Joe (1997) and Joe and Xu (1996), though more generally this is known as multi-stage maximum likelihood (MSML) estimation. Define the MSML estimator as

$$\begin{aligned}
\hat{\theta}_{T,MSML} &\equiv \left[\hat{\alpha}'_{1,T,MSML}, ..., \hat{\alpha}'_{n,T,MSML}, \hat{\gamma}'_{T,MSML}\right]' \\
\hat{\alpha}_{i,T,MSML} &\equiv \arg\max_{\alpha_i} \sum_{t=1}^{T} \log f_{it}\left(Y_{it};\alpha_i\right), \quad i=1,2,...,n \qquad (39)\\
\hat{\gamma}_{T,MSML} &\equiv \arg\max_{\gamma} \sum_{t=1}^{T} \log \mathbf{c}_t\left(F_{1t}\left(Y_{1t};\hat{\alpha}_{1,T,MSML}\right),...,F_{nt}\left(Y_{nt};\hat{\alpha}_{n,T,MSML}\right);\gamma\right)
\end{aligned}$$

Clearly, MSMLE is asymptotically less efficient than one-stage MLE. However, simulation studies in Joe (2005) and Patton (2006b) indicate that this loss is not great in many cases. The main appeal of MSMLE relative to (one-stage) MLE is the ease of estimation: by breaking the full parameter vector into parts the estimation problem is often greatly simplified.

As for one-stage MLE, under regularity conditions, see White (1994) or Patton (2006b), the MSMLE is asymptotically normal:

$$\sqrt{T}\left(\hat{\theta}_{T,MSML}-\theta^*\right) \xrightarrow{d} N\left(0,V^*_{MSML}\right) \quad \text{as} \quad T\to\infty \qquad (40)$$

While estimation is simplified by breaking up estimation in stages, the calculation of an estimator of the asymptotic covariance matrix is more complicated. A critical point to note is that one *cannot* simply take the inverse Hessian of the copula likelihood (the equivalent of $-\hat{A}_T$ in the previous section) as an estimator of the asymptotic covariance of the estimated copula parameters: that estimator ignores the estimation error that arises from the use of $\left[\hat{\alpha}'_{1,T,MSML},...,\hat{\alpha}'_{n,T,MSML}\right]'$ rather than $[\alpha_1^{*\prime},...,\alpha_n^{*\prime}]'$ in the copula estimation step. To capture that additional source of estimation error, the following estimator should be used:

$$\hat{V}_{MSML} = \hat{A}_T^{-1}\hat{B}_T\left(\hat{A}_T^{-1}\right)' \qquad (41)$$

Note that the information matrix equality does not hold for MSML, and so this "sandwich form" for the asymptotic covariance matrix estimator is required. The $\hat{B}_T$ matrix in this case is the analog of that in one-stage MLE:

$$\begin{aligned}
\hat{B}_T &= \frac{1}{T}\sum_{t=1}^{T}\hat{\mathbf{s}}_t\hat{\mathbf{s}}'_t \\
\text{where} \quad \hat{\mathbf{s}}_t &\equiv \left[\hat{\mathbf{s}}'_{1t},...,\hat{\mathbf{s}}'_{nt},\hat{\mathbf{s}}'_{ct}\right]' \qquad (42)\\
\hat{\mathbf{s}}_{it} &= \frac{\partial}{\partial\alpha_i}\log f_{it}\left(Y_{it};\hat{\alpha}_{i,T,MSML}\right), \quad i=1,2,...,n \\
\hat{\mathbf{s}}_{ct} &= \frac{\partial}{\partial\gamma}\log\mathbf{c}_t\left(F_{1t}\left(Y_{1t};\hat{\alpha}_{1,T,MSML}\right),...,F_{nt}\left(Y_{nt};\hat{\alpha}_{n,T,MSML}\right);\hat{\gamma}_{T,MSML}\right)
\end{aligned}$$

The $\hat{A}_T$ matrix takes a different form for MSML, reflecting the presence of estimated parameters in the copula log-likelihood:

$$\begin{aligned}
\hat{A}_T &= \frac{1}{T}\sum_{t=1}^{T}\hat{H}_t \\
\text{where} \quad \hat{H}_t &= \begin{bmatrix} \nabla^2_{11,t} & 0 & \cdots & 0 & 0 \\ 0 & \nabla^2_{22,t} & \cdots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & \nabla^2_{nn,t} & 0 \\ \nabla^2_{1c,t} & \nabla^2_{2c,t} & \cdots & \nabla^2_{nc,t} & \nabla^2_{cc,t} \end{bmatrix} \\
\nabla^2_{ii,t} &= \frac{\partial^2}{\partial\alpha_i\partial\alpha_i'}\log f_{it}\left(Y_{it};\hat{\alpha}_{i,T,MSML}\right), \quad i = 1,2,...,n \\
\nabla^2_{ic,t} &= \frac{\partial^2}{\partial\gamma\partial\alpha_i'}\log \mathbf{c}_t\left(F_{1t}\left(Y_{1t};\hat{\alpha}_{1,T,MSML}\right),...,F_{nt}\left(Y_{nt};\hat{\alpha}_{n,T,MSML}\right);\hat{\gamma}_{T,MSML}\right) \\
\nabla^2_{cc,t} &= \frac{\partial^2}{\partial\gamma\partial\gamma'}\log \mathbf{c}_t\left(F_{1t}\left(Y_{1t};\hat{\alpha}_{1,T,MSML}\right),...,F_{nt}\left(Y_{nt};\hat{\alpha}_{n,T,MSML}\right);\hat{\gamma}_{T,MSML}\right)
\end{aligned} \tag{43}$$

The above discussion shows that $\hat{V}_{MSML}$ is somewhat tedious to obtain, although each of the steps required is no more difficult that the usual steps required to estimate a "sandwich form" asymptotic covariance matrix.

An alternative to these calculations is to use a block bootstrap for inference, see Gonçalves and White (2004) for theoretical justification. This is done as follows: (i) Use a block bootstrap (e.g., the stationary bootstrap of Politis and Romano (1994)) to generate a bootstrap sample of the data of length T, (ii) Estimate the model using the same multi-stage approach as applied for the real data, (iii) Repeat steps (i)-(ii) S times (e.g., $S = 1000$), (iv) Use the $\alpha/2$ and $1-\alpha/2$ quantiles of the distribution of $\left\{\hat{\theta}_i\right\}_{i=1}^{S}$ to obtain a $1-\alpha$ confidence interval for these parameters.

3.2 Semiparametric models

Given the sample sizes that are commonly available in economics and finance, it is often possible to reliably estimate univariate distributions nonparametrically (for example, by using the empirical distribution function) but not enough to estimate higher-dimension distributions or copulas, necessitating the use of a parametric model. Semiparametric copula-based models marry these two estimation methods, using a nonparametric model for the marginal distributions, such as the empirical distribution function (EDF), and a parametric model for the copula. In such cases the estimation of the copula parameter is usually conducted via maximum likelihood, and in this literature this estimator is sometimes called the "canonical maximum likelihood" estimator.

$$\begin{aligned}
\hat{\gamma}_T &\equiv \arg\max_{\gamma} \sum_{t=1}^{T} \log \mathbf{c}\left(\hat{U}_{1t}, ..., \hat{U}_{nt}; \gamma\right) \\
\text{where } \hat{U}_{it} &\equiv \hat{F}_i\left(\hat{\varepsilon}_{it}\right), \quad i = 1, 2, ..., n \\
\hat{F}_i\left(\varepsilon\right) &\equiv \frac{1}{T+1} \sum_{t=1}^{T} \mathbf{1}\left\{\hat{\varepsilon}_{it} \leq \varepsilon\right\} \\
\hat{\varepsilon}_{it} &\equiv \frac{Y_{it} - \mu_i\left(\mathbf{Z}_{t-1}, \hat{\alpha}_i\right)}{\sigma_i\left(\mathbf{Z}_{t-1}, \hat{\alpha}_i\right)}
\end{aligned} \tag{44}$$

The asymptotic distribution of this estimator was studied by Genest, *et al.* (1995) for *iid* data and by Chen and Fan (2006a,b) for time series data.[15] The difficulty here, relative to the parametric case is that the copula likelihood now depends on the infinite-dimensional parameters F_i, as well as the marginal distribution parameters α. Standard maximum likelihood methods cannot be applied here. Chen and Fan (2006b) and Chan *et al.* (2009) provided conditions under which the following asymptotic normal distribution is obtained:

$$\sqrt{T}\left(\hat{\gamma}_T - \gamma^*\right) \xrightarrow{d} N\left(0, V_{SPML}^*\right) \quad \text{as} \quad T \to \infty \tag{45}$$

$$\text{where} \quad V_{SPML}^* = A_{CF}^{-1} \Sigma_{CF} A_{CF}^{-1}$$

The asymptotic covariance matrix, V_{SPML}, takes the "sandwich" form. The outer matrix, A_{CF}, is an inverse Hessian, and Chen and Fan (2006b) show that it can be estimated by:

$$\hat{A}_{CF,T} \equiv -\frac{1}{T} \sum_{t=1}^{T} \frac{\partial^2 \log \mathbf{c}_t\left(\hat{U}_{1t}, ..., \hat{U}_{nt}; \hat{\gamma}_T\right)}{\partial\gamma\partial\gamma'} \tag{46}$$

The inner matrix, Σ_{CF}, is a form of outer product of gradients, but for this semiparametric estimator it is not simply the scores of the log-likelihood; an additional term appears due to the presence of

[15] Chen, Fan and Tsyrennikov (2006) propose a one-stage estimator of this model, in contrast with the multi-stage estimator considered here, based on splines for the nonparametric marginal distribution functions which attains full efficiency.

the EDF in the objective function:

$$\hat{\Sigma}_{CF,T} = \frac{1}{T}\sum_{t=1}^{T}\mathbf{s}_t\mathbf{s}_t' \tag{47}$$

$$\text{where} \quad \mathbf{s}_t \equiv \frac{\partial}{\partial\gamma}\log\mathbf{c}_t\left(\hat{U}_{1t},...,\hat{U}_{nt};\hat{\gamma}_T\right)+\sum_{j=1}^{n}\hat{Q}_{jt} \tag{48}$$

$$\hat{Q}_{jt} \equiv \frac{1}{T}\sum_{s=1,s\neq t}^{T}\frac{\partial^2\log\mathbf{c}_t\left(\hat{U}_{1s},...,\hat{U}_{ns};\hat{\gamma}_T\right)}{\partial\gamma\partial U_j}\left(\mathbf{1}\left\{\hat{U}_{jt}\leq\hat{U}_{js}\right\}-\hat{U}_{js}\right) \tag{49}$$

The above result shows that the asymptotic variance of the MLE of the copula parameter depends on the estimation error in the EDF (through the terms $\hat{Q}_{jt}$) but surprisingly does *not* depend upon the estimated parameters in the marginal distributions ($\hat{\alpha}_j$). This is particularly surprising as all estimates in this framework are $\sqrt{T}$-consistent. Thus in this case the researcher can estimate ARMA-GARCH type models (or others) for the conditional mean and variance, compute the standardized residuals, and then *ignore*, for the purposes of copula estimation and inference, the estimation error from the ARMA-GARCH models. Two important caveats are worth noting here: Firstly, this only applies for *constant* conditional copula models; if the conditional copula is time-varying, the Rémillard (2010) shows that the estimation error from the models for the conditional mean and variance will affect the asymptotic distribution of the copula parameter estimate. Second, this only applies when the marginal distributions of the standardized residuals are estimated nonparametrically; as discussed in the previous section, with parametric marginal distribution models the estimation error from the models for the conditional mean and variance will, in general, affect the distribution of the copula parameter estimate.

Chen and Fan (2006b) and Rémillard (2010) also propose a simple bootstrap alternative to the above calculations for inference on the estimated copula parameters: (i) Use an *iid* bootstrap to generate a bootstrap sample of the estimated standardized residuals of length T, (ii) Transform each time series of bootstrap data using its empirical distribution function, (iii) Estimate the copula model on the transformed data (iv) Repeat steps (i)-(iii) S times (e.g., $S = 1000$), (v) Use the $\alpha/2$ and $1-\alpha/2$ quantiles of the distribution of $\left\{\hat{\theta}_i\right\}_{i=1}^{S}$ to obtain a $1-\alpha$ confidence interval for these parameters. Of course, the bootstrap distribution of the parameter estimates can also be used for conducting joint tests on the parameters.

Another alternative, proposed by Rémillard (2010), is to simulate from the estimated copula model, rather than bootstrap the standardized residuals: (i) Simulate a sample of length T using *iid* draws from the copula model using the estimated parameters, (ii) Transform each series using its EDF,[16], then follow steps (iii)-(v) of the bootstrap method above.

3.3 Other estimation methods

While maximum likelihood estimation is the most prevalent in the literature, other methods have been considered. Method of moments-type estimators, where the parameter of a given family of copulas has a known, invertible, mapping to a dependence measure (such as rank correlation or Kendall's tau) are considered in Genest (1987), Ghoudi and Rémillard (2004) and Rémillard (2010), among others. Generalized method of moments, where the number of dependence measures may be greater than the number of unknown parameters, and simulated method of moments are considered in Oh and Patton (2011a). Minimum distance estimation is considered by Tsukahara (2005). Bayesian estimation of copula models is considered in Min and Czado (2010), Smith, *et al.* (2010a,b), see Smith (2011) for a review.

3.4 Empirical illustration, continued

In this section we continue our study of daily returns on a large-cap equity index (the S&P 100) and a small-cap equity index (the S&P 600), over the period 1995-2011. In Section 1.1 we verified that simple AR-GARCH type models for the conditional mean and variance appeared to fit the data well, and we confirmed that the skewed t distribution of Hansen (1994) could not be rejected as a model for the conditional distribution of the standardized residuals using goodness-of-fit tests. In Sections 2.4 and 2.5 we found mild evidence of asymmetric dependence between these two series (with crashes being more strongly dependent than booms) and stronger evidence for time-varying dependence. We will now consider a variety of parametric models for the copula of these two series, along with several different approaches for computing standard errors on the estimated parameters. A summary of some common copula models and their properties is presented in Table 3.[17]

[16] Note that we estimate the marginal distributions of the simulated draws from the copula model using the EDF, even though the margins are known to be $Unif\,(0,1)$ in this case, so that the simulation approach incorporates the EDF estimation error faced in practice.

[17] Mixtures of copulas are also valid copulas, and thus by combining the simple copulas in Table 3 new models may be obtained, see Hu (2006) for example.

We will first discuss the use of parametric copula models for estimating tail dependence coefficients. Then we will consider models for the entire dependence function, first assuming that the conditional copula is constant, and then extend to time-varying conditional copulas.

[INSERT TABLE 3 ABOUT HERE]

3.4.1 Estimating tail dependence using parametric copula models

An alternative to the nonparametric estimation of tail dependence coefficients discussed in Section 2.1 is to specify and estimate parametric models for the tails of the joint distribution, see McNeil, *et al.* (2005) for example. For data sets with relatively few observations, the additional structure provided by a parametric model can lead to less variable estimates, though the use of a parametric model of course introduces the possibility of model misspecification.

This approach uses a parametric model on the bivariate tail and use the fitted model to obtain an estimate of the tail dependence coefficient. To allow for asymmetric dependence, this is done on the lower and upper tails separately.[18] To do this, note from Chen, *et al.* (2010), that if $(U, V) \sim \mathbf{C}(\theta)$, then the log-likelihood of (U, V) conditional on $\{U > q, V > q\}$ is

$$\begin{aligned}
\log L(\theta|q) &= \frac{1}{T}\sum_{t=1}^{T} l_t(\theta|q) \\
\text{where} \quad l_t(\theta|q) &= \delta_{1t}\delta_{2t}\log \mathbf{c}\left(\tilde{U}_t, \tilde{V}_t; \theta\right) + \delta_{1t}(1-\delta_{2t})\log \frac{\partial \mathbf{C}\left(\tilde{U}_t, \tilde{V}_t; \theta\right)}{\partial u} \\
&\quad + (1-\delta_{1t})\,\delta_{2t}\log \frac{\partial \mathbf{C}\left(\tilde{U}_t, \tilde{V}_t; \theta\right)}{\partial v} + (1-\delta_{1t})(1-\delta_{2t})\log \mathbf{C}\left(\tilde{U}_t, \tilde{V}_t; \theta\right) \\
\text{and} \quad \tilde{U}_t &= \max[U_t, q], \quad \tilde{V}_t = \max[V_t, q] \\
\delta_{1t} &= \mathbf{1}\{U_t > q\}, \quad \delta_{2t} = \mathbf{1}\{V_t > q\}
\end{aligned} \tag{50}$$

That is, we replace all values of (U_t, V_t) that are less than q by q, and we use the indicators δ_{1t} and δ_{2t} to record the values that are *not* censored. Maximizing the above likelihood yields an estimate of

[18] Note also that the parametric copula chosen must, obviously, be one that allows for non-zero tail dependence in the tail in which it is to be used. For example, using a Normal or Frank copula as a model for the tail copula guarantees that the estimated tail dependence coefficient is zero, as this is a feature of these copulas, see Table 3. Similarly, using the left tail of the Gumbel copula also ensures an estimated tail dependence of zero. Instead, one should use the right tail of a Gumbel copula, or a t copula, or the left tail of a Clayton copula, or one of many other copulas that allow for non-zero tail dependence. See de Haan, *et al.* (2008) for details on estimation and testing of parametric tail copulas.

the parameters of the upper tail copula. The lower tail copula can be modeled similarly. Estimation via MLE is generally simple, and all that is required beyond usual MLE is a function for the copula *cdf* (which is usually already known) and a function for $\partial\mathbf{C}/\partial u$ and $\partial\mathbf{C}/\partial v$. For many copulas these latter functions are easy to obtain. Given an estimate of the tail copula parameter for each of the tails, we obtain the estimated tail dependence coefficients as:

$$\begin{aligned} \hat{\lambda}^L &= \lim_{q\to 0^+} \frac{\mathbf{C}^L\left(q,q;\hat{\theta}^L\right)}{q} \\ \hat{\lambda}^U &= \lim_{q\to 1^-} \frac{1-2q+\mathbf{C}^U\left(q,q;\hat{\theta}^U\right)}{1-q} \end{aligned} \tag{51}$$

These coefficients are known in closed form for many commonly-used copulas (e.g., the Gumbel, Clayton, and Student's t), see Joe (1997), Nelsen (2006) and Demarta and McNeil (2005), and see Table 3 for a summary.

Table 4 presents four estimates of these coefficients, the first two are nonparametric (the expression for the "log" estimator is given in equation (17), and the "sec" estimator is given in Frahm, *et al.*, 2005), and the second two are parametric, based on the Gumbel and Student's t for the upper tail, and the "rotated Gumbel" and Student's t for the lower tail. The cutoffs used for determining the parametric tail copula are 0.025 and 0.975, which yields 49 (39) observations to estimate the lower (upper) tail copula.[19] The estimated tail copula parameters are $\hat{\theta}^L = 1.455$ and $\hat{\theta}^U = 1.263$, and using the expression for the tail dependence of a Gumbel copula presented in Table 3, the implied estimated tail dependence coefficients are $\hat{\lambda}^L = 0.390$ and $\hat{\lambda}^U = 0.269$. The Student's t tail copula parameters are $\left[\hat{\rho}^L, \hat{\nu}^L\right] = [0.592, 4.896]$ and $\left[\hat{\rho}^U, \hat{\nu}^U\right] = [0.446, 5.889]$, implying tail dependence coefficients of $\hat{\lambda}^L = 0.266$ and $\hat{\lambda}^U = 0.149$. An *iid* bootstrap was again used to obtain a 90% confidence interval on these estimates, reported in Table 4. As Table 4 reveals, the point estimates of the upper and lower tail dependence coefficients are very similar across three of the four methods, with the tail dependence implied by the Student's t copula being lower than the other three estimates. The precision of these estimates, however, varies greatly depending on whether a parametric or nonparametric approach is used.

[INSERT TABLE 4 ABOUT HERE]

[19] As usual in estimating "tail" quantities, the choice of cut-off is somewhat arbitrary. I experimented with cut-off values between 0.01 and 0.05.

3.4.2 Constant copula models

Next we consider copula models for the entire dependence structure, not just the tails. The estimation of constant copula models is straightforward and fast for the multi-stage estimation method we consider here, as the number of parameters in most (bivariate) copulas is just one or two. In higher dimensions the task is more challenging, see Oh and Patton (2011b) for an example of a 100-dimensional copula application. In Table 5 below we first present the estimated parameters and values of the log-likelihood for a variety of models. The left columns present results for the fully parametric case (where the parametric copulas are combined with parametric models for the marginal distributions) and the right columns contain results for the semiparametric models. The top three models in terms of log-likelihood are highlighted in bold.[20]

[INSERT TABLE 5 ABOUT HERE]

Table 5 reveals that of these nine specifications, the best copula model for both the parametric and semiparametric case is the Student's t copula, followed by the "rotated Gumbel" copula and then the Normal copula. By far the worst model is the "rotated Clayton" copula, which imposes zero lower tail dependence and allows only for upper tail dependence.

Next we focus on a subset of these models, and compute a range of different standard errors for the estimated parameters. For both the parametric and semiparametric cases, we consider (i) naïve standard errors, where the estimation error from the earlier stages of estimation (AR, GARCH and marginal distributions) is ignored, (ii) multi-stage MLE or multi-stage semiparametric MLE (MSML) standard errors, using the asymptotic distribution theory for these estimators in Patton (2006b) or Chen and Fan (2006b) respectively, (iii) bootstrap standard errors, using either a block bootstrap[21] of the original returns and estimation of all stages on the bootstrap sample (parametric case), based on Gonçalves and White (2004), or an *iid* bootstrap of the standardized residuals and estimation only of the EDF and the copula (semiparametric case), based on Chen and Fan (2006b) and Rémillard (2010), and (iv) a simulation-based standard error. For the parametric case the model for the entire joint distribution is simulated many times using the estimated parameters, and on each of the simulated samples the parameters are re-estimated, while for the semiparametric

[20] The inverse degrees of freedom parameter, ν^{-1}, is estimated to facilitate simple tests on this parameter below.

[21] Specifically, the stationary bootstrap of Politis and Romano (1994) with an average block length of 60 observations is used.

case only the copula model is simulated, the EDF of the simulated data is computed, and the copula parameters are re-estimated, as suggested by Rémillard (2010). In the parametric case this approach yields correct *finite-sample* standard errors, while the semiparametric case, and all the other methods of obtaining standard errors, rely on asymptotic theory. For the bootstrap and the simulation-based standard errors 1000 replications are used. The results are presented in Table 6.

[INSERT TABLE 6 ABOUT HERE]

Table 6 shows that the naïve standard errors are too small relative to the correct MSML standard errors, a predictable outcome given that naïve standard errors ignore the additional estimation error arising from the estimation of marginal distribution parameters. In the parametric case the naïve standard errors are on average about half as large as the MSML standard errors (average ratio is 0.54), while for the semiparametric case the ratio is 0.84. The relatively better performance in the semiparametric case is possibly attributable to the fact that the MSML standard errors in that case can, correctly, ignore the estimation error coming from the AR-GARCH models for the conditional mean and variance, with adjustment required only for estimation error coming from the EDF. In the fully parametric case, adjustments for estimation error from marginal distribution shape parameters *and* the parameters of the AR-GARCH models must be made.

In both the parametric and the semiparametric cases the bootstrap standard errors are very close to the MSML standard errors, with the ratio of the former to the latter being 0.98 and 0.97 respectively. This is what we expect asymptotically, and confirms that the researcher may use either 'analytical' MSML standard errors or more computationally-intensive bootstrap standard errors for inference on the estimated copula parameters. The simulation-based standard errors for the semiparametric case are also close to the MSML standard errors (with the average ratio being 1.07). In the parametric case, where correct finite-sample standard errors can be obtained, we see that these are smaller than the MSML and bootstrap standard errors, with the average ratio being around 0.7. Asymptotically we expect this ratio to go to 1, but in finite samples this value of ratio will depend on the particular model being used.

3.4.3 Time-varying copula models

Next we consider two time-varying models for the conditional copula of these standardized residuals. In both cases we will use the "GAS" model of Creal, *et al.* (2011), which specifies the time-varying copula parameter (δ_t) as evolving as a function of the lagged copula parameter and a "forcing variable" that is related to the standardized score of the copula log-likelihood. To deal with parameters that are constrained to lie in a particular range (e.g., a correlation parameter forced to take values only inside $(-1,1)$), this approach applies a strictly increasing transformation (e.g., log, logistic, arc tan) to the copula parameter, and models the evolution of the transformed parameter, denoted f_t :

$$\begin{aligned} f_t &= h(\delta_t) \Leftrightarrow \delta_t = h^{-1}(f_t) & (52) \\ \text{where} \quad f_{t+1} &= \omega + \beta f_t + \alpha I_t^{-1/2}\mathbf{s}_t & (53) \\ \mathbf{s}_t &\equiv \frac{\partial}{\partial \delta} \log \mathbf{c}(U_{1t}, U_{2t}; \delta_t) & (54) \\ I_t &\equiv E_{t-1}\left[\mathbf{s}_t \mathbf{s}_t'\right] = I(\delta_t) & (55) \end{aligned}$$

Thus the future value of the copula parameter is a function of a constant, the current value, and the score of the copula-likelihood, $I_t^{-1/2}\mathbf{s}_t$. We will consider a time-varying rotated Gumbel copula and time-varying Student's t copula. The Gumbel copula parameter is required to be greater than one, and the function $\delta_t = 1 + \exp(f_t)$ is used to ensure this. For the Student's t copula we will assume that the degrees of freedom parameter is constant and that only the correlation parameter is time-varying. As usual, this parameter must lie in $(-1,1)$, and the function $\delta_t = (1 - \exp\{-f_t\}) / (1 + \exp\{-f_t\})$ is used to ensure this.

The estimated parameters for these two models are presented in Table 7. For both the parametric and the semiparametric models we see that the Student's t specification has a higher value of the likelihood, perhaps reflecting its additional free parameter. Consistent with what one might expect given results in the volatility literature, the estimated degrees of freedom parameter is higher for the time-varying Student's t copula model than for the constant version (11.2 compared with 6.9). Thus time-varying dependence may explain some (but not all) of the tail dependence estimated via the constant Student's t copula, see Manner and Segers (2011) on stochastic copulas and tail dependence.

When the time-varying conditional copula model is combined with parametric marginal distributions the resulting joint distribution is fully parametric, and all of the inference methods reviewed for the constant copula case may be applied here. The left panel of Table 7 presents four different estimates of the standard errors of these models. As in the constant case, we again observe that the naïve standard errors, which ignore the estimation error contributed from the marginal distributions, are too small relative to the MSML standard errors, and the MSML and bootstrap standard errors are generally similar.

When the marginal distributions are estimated using the EDF, the resulting joint distribution is semiparametric. Unlike the constant copula case, the true standardized residuals in this case are not *jointly iid*, even though they are individually *iid*, which means that the theoretical results of Chen and Fan (2006a) and Rémillard (2010) cannot be applied. Moreover this implies (see Rémillard, 2010) that the estimation error coming from the parametric models for the marginal dynamics *will*, in general, affect the asymptotic distribution of the estimated copula parameters. Inference methods for these models have not yet been considered in the econometrics or statistics literature, to the best of my knowledge. One intuitive inference method, which still needs formal justification, is to use a block bootstrap technique similar to the parametric case, where the original data are bootstrapped (in blocks, to preserve the temporal dependence structure) and then the semiparametric model is estimated on the bootstrap data[22]. Standard errors using such a technique are presented in the right panel of Table 7, along with naïve standard errors that ignore the estimation error in the marginal distributions altogether.

[INSERT TABLE 7 ABOUT HERE]

4 Model selection and goodness-of-fit testing

In this section we consider the problems of model selection and goodness-of-fit (GoF) testing. The latter problem is the traditional specification testing problem, and seeks to determine whether the proposed copula model is different from the (unknown) true copula. The former testing problem seeks to determine which model in a given set of competing copula models is the "best", according to some measure.

In economic applications GoF tests and model selection tests are complementary: In some applications a GoF test is too weak a criterion, as limited data may mean that several, non-overlapping, models are not rejected. In other applications a GoF test may be too strict a criterion, as in economics we generally do not expect *any* of our models to be correctly specified, and a rejection does not necessarily mean that the model should be discarded. Model selection tests, on the other hand, allow the researcher to identify the best model from the set, however they do

[22] See footnote 10 for discussion.

not usually provide any information on whether the best model is close to being true (which is a question for a GoF test) or whether it is the "best of a bad bunch" of models. These caveats noted, GoF tests and model selection tests are useful ways of summarizing model performance.

4.1 Tests of goodness of fit

Inference for tests of goodness-of-fit (GoF) differ depending on whether the model under analysis is parametric or semiparametric, and we will consider these two cases separately. We will focus on in-sample (full sample) tests of GoF, see Chen (2011) for analysis of out-of-sample GoF tests.

Two tests that are widely used for GoF tests of copula models are the Kolmogorov-Smirnov and the Cramer-von Mises tests[23]. The test statistics for these tests in univariate applications are presented in equations (9) and (10); the multivariate versions of these statistics are presented below. These tests use the empirical copula, denoted $\hat{\mathbf{C}}_T$, which is also defined below.

$$\hat{\mathbf{C}}_T(\mathbf{u}) \equiv \frac{1}{T}\sum_{t=1}^{T}\prod_{i=1}^{n}\mathbf{1}\left\{\hat{U}_{it} \leq u_i\right\} \tag{56}$$

$$KS_C = \max_t \left|\mathbf{C}\left(\mathbf{U}_t;\hat{\theta}_T\right) - \hat{\mathbf{C}}_T(\mathbf{U}_t)\right| \tag{57}$$

$$CvM_C = \sum_{t=1}^{T}\left\{\mathbf{C}\left(\mathbf{U}_t;\hat{\theta}_T\right) - \hat{\mathbf{C}}_T(\mathbf{U}_t)\right\}^2 \tag{58}$$

Note that approaches based on a comparison with the empirical copula, such as those above, only work for *constant* copula models, as they rely on the empirical copula serving as a nonparametric estimate of the true conditional copula. When the true conditional copula is time-varying, the empirical copula can no longer be used for that purpose. One way of overcoming this problem is to use the fitted copula model to obtain the "Rosenblatt" transform of the data, which is a multivariate version of the probability integral transformation, and was used in Diebold, *et al.* (1999) and further studied in Rémillard (2010). In the bivariate case, the transform is

$$\begin{aligned} V_{1t} &= U_{1t} \;\forall\; t \\ V_{2t} &= \mathbf{C}_{2|1,t}\left(U_{2t}|U_{1t};\theta\right) \end{aligned} \tag{59}$$

[23] Genest, *et al.* (2009) provide a comprehensive review of the many copula GoF tests available in the literature, and compare these tests via a simulation study. Across a range of data generating processes, they conclude that a Cramer-von Mises test (applied to the empirical copula or to the Rosenblatt transform of the original data) is the most powerful, a finding that is supported by Berg (2009) who considers some further tests.

where $\mathbf{C}_{2|1,t}$ is the conditional distribution of $U_{2t}|U_{1t}$. In general multivariate applications, the transformation is:

$$V_{it} = \frac{\partial^{i-1}\mathbf{C}\left(U_{1t},...,U_{it},1,...,1\right)}{\partial u_1\cdots\partial u_{i-1}} \Bigg/ \frac{\partial^{i-1}\mathbf{C}\left(U_{1t},...,U_{i-1,t},1,...,1\right)}{\partial u_1\cdots\partial u_{i-1}}, \quad i=2,...,n \tag{60}$$
$$\equiv \frac{\mathbf{C}_{i|i-1,...,1}\left(U_{it}|U_{i-1,t},...,U_1\right)}{\mathbf{c}_{1,2,...,i-1}\left(U_{i-1,t},...,U_1\right)}$$

i.e., the numerator is the conditional distribution of U_{it} given $[U_{1t},...,U_{i-1,t}]$, and the denominator is the conditional density of $[U_{1t},...,U_{i-1,t}]$.

The usefulness of this transformation lies in the result that if the specified conditional copula model is correct, then

$$\mathbf{V}_t \equiv [V_{1t},...,V_{nt}]' \sim iid\ \mathbf{C}_{indep} \tag{61}$$

That is, the Rosenblatt transformation of the original data returns a vector of *iid* and mutually independent $Unif\,(0,1)$ variables. With this result in hand, we can again use the KS or CvM test statistics to test whether the empirical copula of the estimated Rosenblatt transforms is significantly different from the independence copula.[24]

$$\hat{\mathbf{C}}_T^v(\mathbf{v}) \equiv \frac{1}{T}\sum_{t=1}^{T}\prod_{i=1}^{n}\mathbf{1}\left\{V_{it}\le v_i\right\} \tag{62}$$

$$\mathbf{C}^v\left(\mathbf{V}_t;\hat{\theta}_T\right) = \prod_{i=1}^{n}V_{it} \tag{63}$$

$$KS_R = \max_t\left|\mathbf{C}^v\left(\mathbf{V}_t;\hat{\theta}_T\right)-\hat{\mathbf{C}}_T^v(\mathbf{V}_t)\right| \tag{64}$$

$$CvM_R = \sum_{t=1}^{T}\left\{\mathbf{C}^v\left(\mathbf{V}_t;\hat{\theta}_T\right)-\hat{\mathbf{C}}_T^v(\mathbf{V}_t)\right\}^2 \tag{65}$$

4.1.1 Fully parametric

For fully parametric copula-based models, GoF testing is a relatively standard problem, as these models are simply non-linear time series models. See Corradi and Swanson's (2006) review article on evaluating predictive densities, Bontemps, *et al.* (2011) and Chen (2011) on GoF tests for multivariate distributions via moment conditions, Chen (2007) for moment-based tests directly on

[24] Note that the order of the variables affects the Rosenblatt transformation. In most economic applications the ordering of the variables is arbitrary. One way to overcome this is to conduct the test on all possible orderings and then define a new test statistic as the maximum of all the test statistics. The simulation-based methods for obtaining critical values described below could also be applied to this "combination" test statistic.

the copula, and Diebold, *et al.* (1999) on GoF tests via Rosenblatt's transform, discussed below (although the latter paper ignores estimation error in the model parameters).

A difficulty in obtaining critical values for GoF test statistics, such as the KS and CvM test statistics, is that they depend on estimated parameters, both in the copula and also in marginal distributions. As discussed in the context of obtaining standard errors on estimated copula parameters, the parameter estimation error coming from the marginal distributions cannot in general be ignored.

GoF tests can be implemented in various ways, but for fully parametric models a simple simulation-based procedure is always available: (i) Estimate the margins and copula model parameters on the actual data to obtain the parameter estimate, $\hat{\theta}_T$ (ii) Compute the GoF test statistic (for example, the Kolmogorov-Smirnov or Cramer-von Mises test statistics) on the actual data, $\hat{G}_T$ (iii) Simulate a time series of length T from the model using the estimated parameter $\hat{\theta}_T$ (iv) Estimate the model on the simulated data to obtain $\hat{\theta}_T^{(s)}$ (v) Compute the GoF statistic on the simulated data, $\hat{G}_T^{(s)}$, (vi) Repeat steps (iii)-(v) S times, (vii) compute the simulation-based p-value for this test as:

$$p_{T,S} = \frac{1}{S} \sum\nolimits_{s=1}^{S} \mathbf{1}\left\{ \hat{G}_T^{(s)} \geq \hat{G}_T \right\} \tag{66}$$

4.1.2 Semiparametric

Rémillard (2010) considers GoF tests for semiparametric copula-based models for time series, and shows the surprising and useful result that the asymptotic distributions of GoF copula tests are unaffected by the estimation of marginal distribution parameters (as was the case for the asymptotic distribution of the estimated copula parameters). The estimation error coming from the use of the empirical distribution functions *does* matter, and he proposes a simple simulation-based method to capture this: (i) estimate the margins and copula model parameters on the actual data to obtain the parameter estimate, $\hat{\theta}_T$ (ii) Compute the GoF test statistic (for example, the Kolmogorov-Smirnov or Cramer-von Mises test statistics) on the actual data, $\hat{G}_T$ (iii) Simulate a time series of length T from the copula model using the estimated parameter $\hat{\theta}_T$ (iv) Transform each time series of simulated data using its empirical distribution function (v) Estimate the copula model on the transformed simulated data to obtain $\hat{\theta}_T^{(s)}$ (vi) Compute the GoF statistic on the simulated data, $\hat{G}_T^{(s)}$, (vi) Repeat steps (iii)-(vi) S times, (viii) compute the simulation-based p-value for this test as in the parametric case.

The case of nonparametric margins combined with a time-varying conditional copula has not yet been considered in the literature. In the empirical example below I obtain a simulation-based p-value using the same approach as the parametric case considered in the previous section, using the EDF in place of the estimated parametric marginal distribution. Theoretical support for this approach is still required.

4.1.3 Empirical illustration, continued

Table 8 presents the results of four GoF tests for the copula models considered in Section 3.4. The top panel considers fully parametric models, and the lower panel semiparametric models. Both KS and CvM tests are applied, either to the empirical copula of the standardized residuals (KS_C and CvM_C) or to the Rosenblatt transformation of the standardized residuals (KS_R and CvM_R). For the two time-varying copula models only the tests based on the Rosenblatt transformation are applicable.

The left panel presents the p-values from an implementation of these tests that ignores the estimation error from the marginal distributions, though it does take into account the estimation error from the copula parameters. The right panel presents p-values from tests that appropriately account for estimation error from the marginal distributions. We observe in Table 8 that ignoring estimation error leads almost uniformly to p-values that are larger than when this estimation error is taken into account. Thus in addition to providing a false estimate of high precision of estimated parameters, as observed in Tables 6 and 7, ignoring estimation error from the marginal distributions also provides a false indication of a good fit to the data.

Using the correct p-values, we observe that the constant conditional copula models are all rejected, particularly so when combined with nonparametric marginal distributions. The time-varying (GAS) copula models both pass the GoF tests in the parametric case, however only the rotated Gumbel specification passes the CvM test in the semiparametric case. Thus we have substantial evidence against the constant copula assumption, and moderate evidence that the two time-varying copula models described in Section 3.4 are also rejected.

[INSERT TABLE 8 ABOUT HERE]

4.2 Model selection tests

The problem of finding the model that is best, according to some criterion, among a set of competing models (i.e., the problem of "model selection") may be undertaken either using the full sample (in-sample) of data, or using an out-of-sample (OOS) period. The treatment of these two cases differs, as does the treatment of parametric and semiparametric models, and we will consider all four combinations. The problem also differs on whether the competing models are nested or non-nested. Below we will focus on pair-wise comparisons of models; for comparisons of large collections of models see White (2000), Romano and Wolf (2005) and Hansen, *et al.* (2011), for example.

4.2.1 In-sample, nested model comparison via parameter restrictions

In-sample model selection tests are generally straightforward if the competing models are *nested*, as a likelihood ratio test can generally be used.[25] In this case the smaller model is held as the true model under the null hypothesis, and under the alternative the larger model is correct. For example, comparing a Normal copula with a Student's t copula can be done via a test on the inverse degree of freedom parameter:

$$H_0 : \nu^{-1} = 0 \quad \text{vs.} \quad H_a : \nu^{-1} > 0 \tag{67}$$

Notice that the parameter, ν^{-1}, is on the boundary under the null, and so the usual t-statistic will not have the usual $N(0,1)$ limited distribution, however the right-tail critical values (which are the ones that are relevant for testing against this alternative) are the same, e.g., 90% and 95% critical values for the t statistic are 1.28 and 1.64. These tests can be used in both fully parametric and semiparametric applications.

4.2.2 Fully parametric, in-sample

Rivers and Vuong (2002) consider model selection for general parametric nonlinear dynamic models. They allow for many $\sqrt{T}$-consistent estimators (e.g., ML, GMM, minimum distance), they consider nested and non-nested models, and they allow one or both models to be misspecified. This latter feature is particularly attractive in economic applications. Rivers and Vuong (2002) consider a

[25] The problem becomes more complicated if the smaller model lies on the boundary of the parameter space of the larger model, or if some of the parameters of the larger model are unidentified under the null that the smaller model is correct. See Andrews (2001) and Andrews and Ploberger (1994) for discussion of these issues.

range of different applications, but for copula applications their results simplify greatly if (i) the models are non-nested, (ii) we estimate the marginals and the copula by ML (one-stage *or* multi-stage) *and* (iii) we compare models using their joint log-likelihood. In this case, the null and alternative hypotheses are:

$$\begin{aligned} & H_0 & : & \quad E\left[L_{1t}\left(\theta_1^*\right) - L_{2t}\left(\theta_2^*\right)\right] = 0 & (68) \\ \text{vs.} \quad & H_1 & : & \quad E\left[L_{1t}\left(\theta_1^*\right) - L_{2t}\left(\theta_2^*\right)\right] > 0 & \\ & H_2 & : & \quad E\left[L_{1t}\left(\theta_1^*\right) - L_{2t}\left(\theta_2^*\right)\right] < 0 & \\ \text{where} \quad & L_{it}\left(\theta_i^*\right) & \equiv & \quad \log \mathbf{f}_{it}\left(\mathbf{Y}_t; \theta_i^*\right) & (69) \end{aligned}$$

(Note that if the same marginal distributions are used for both models, then the difference between the joint log-likelihoods reduces to the difference between the copula likelihoods.) Rivers and Vuong (2002) show that a simple t-statistic on the difference between the sample averages of the log-likelihoods has the standard Normal distribution under the null hypothesis:

$$\frac{\sqrt{T}\left\{\bar{L}_{1T}\left(\hat{\theta}_{1T}\right) - \bar{L}_{2T}\left(\hat{\theta}_{2T}\right)\right\}}{\hat{\sigma}_T} \xrightarrow{d} N(0,1) \quad \text{under } H_0 \tag{70}$$

$$\text{where} \quad \bar{L}_{iT}\left(\hat{\theta}_{iT}\right) \equiv \frac{1}{T}\sum_{t=1}^{T} L_{it}\left(\hat{\theta}_{iT}\right), \quad i = 1, 2$$

and $\hat{\sigma}_T^2$ is some consistent estimator of $V\left[\sqrt{T}\left\{\bar{L}_{1T}\left(\hat{\theta}_{1T}\right) - \bar{L}_{2T}\left(\hat{\theta}_{2T}\right)\right\}\right]$, such as the Newey-West (1987) HAC estimator. This is a particularly nice result as it shows that we can ignore estimation error in $\hat{\theta}_{1T}$ and $\hat{\theta}_{2T}$, and do not need to compute asymptotic variances of these quantities or use simulations to get critical values. Note that the Rivers and Vuong (2002) test may be applied to both constant *and* time-varying conditional copula models.

Rivers and Vuong (2002) show that their test can also be applied when some metric other than the joint likelihood is used for measuring goodness of fit. In this case the variance, $\hat{\sigma}_T^2$, needs to be adjusted to take into account the estimation error from the parameters.

4.2.3 Semiparametric, in-sample

Chen and Fan (2006b) consider a similar case to Rivers and Vuong (2002), but for semiparametric copula-based models, under the assumption that the conditional copula is constant. Chen and

Fan (2006b) show that when the models are "generalized non nested"[26] the likelihood ratio t test statistic is again Normally distributed under the null hypothesis:

$$\frac{\sqrt{T}\left\{\bar{L}_{1T}\left(\hat{\theta}_{1T}\right)-\bar{L}_{2T}\left(\hat{\theta}_{2T}\right)\right\}}{\hat{\sigma}_T}\rightarrow N\left(0,1\right)\quad \text{under } H_0$$

$$\begin{aligned}
\text{where }\ \hat{\sigma}_T^2 &= \frac{1}{T}\sum_{t=1}^{T}\left(\tilde{d}_t+\sum_{j=1}^{n}\left\{\hat{Q}_{2jt}\left(\hat{\gamma}_{2T}\right)-\hat{Q}_{1jt}\left(\hat{\gamma}_{1T}\right)\right\}\right)^2 && (71)\\
d_t &\equiv \log \mathbf{c}_1\left(\hat{\mathbf{U}}_t;\hat{\gamma}_{1T}\right)-\log \mathbf{c}_2\left(\hat{\mathbf{U}}_t;\hat{\gamma}_{2T}\right)\\
\tilde{d}_t &= d_t-\bar{d}_T\\
\hat{Q}_{ijt}\left(\hat{\gamma}_{iT}\right) &\equiv \frac{1}{T}\sum_{s=1,s\neq t}^{T}\left\{\frac{\partial \log \mathbf{c}_i\left(\hat{\mathbf{U}}_s;\hat{\gamma}_{iT}\right)}{\partial u_j}\left(\mathbf{1}\left\{\hat{U}_{jt}\leq \hat{U}_{js}\right\}-\hat{U}_{js}\right)\right\} && (72)
\end{aligned}$$

Note that the asymptotic variance is more complicated in one sense, as the estimation error coming from the use of the EDF must be incorporated, which is accomplished through the terms $\hat{Q}_{1j}$ and $\hat{Q}_{2j}$. It is simpler in another sense, as the authors exploit the constant conditional copula assumption and avoid the need for a HAC estimator of the variance of $\bar{d}_T$.

The Chen and Fan (2006b) test for comparing copula models is derived under the assumption that the conditional copula is constant, and corresponding results for the time-varying case are not available in the literature, to my knowledge.

4.2.4 Empirical illustration, continued

The upper panel of Table 9 presents the results of Rivers and Vuong (2002) pair-wise comparison tests of the parametric copula-based multivariate models introduced in Section 3.4 above. These results show that the Clayton copula is significantly beaten by all three other models, while the Student's t copula significantly outperforms all three other models. (The comparison of the Student's t copula with the Normal copula is done as a one-sided t test on the significance of the inverse degrees of freedom parameter, as in equation (67) above). The rotated Gumbel copula is better but not significantly better than the Normal copula. The lower panel of Table 9 presents

[26] Chen and Fan (2006b) define two copula models to be generalized non-nested if the set $\{\mathbf{u}:\mathbf{c}_1\left(\mathbf{u};\alpha_1^*\right)\neq \mathbf{c}_2\left(\mathbf{u};\alpha_2^*\right)\}$ has positive Lebesgue measure, where α_i^* is the limiting parameter of copula model i, i.e., if the models, evaluated at their limiting parameters, differ somewhere in their support.

the corresponding Chen and Fan (2006b) tests for the semiparametric copula-based multivariate models, and the same conclusions are obtained.

With parametric marginal distributions we can also use the Rivers and Vuong test (2002) to compare the time-varying rotated Gumbel and Student's t copulas. The t-statistic from that test is 4.27, very strongly in favor of the time-varying Student's copula.

Comparisons of time-varying and constant conditional copulas are usually complicated by the presence of a parameter that is unidentified under the null hypothesis. When using the GAS model, see equation (53), a constant copula is obtained when $\alpha = 0$, but this leaves β unidentified. Tests to accommodate this may be obtained by combining the results of Rivers and Vuong (2002) with those of Andrews (2001) and Andrews and Ploberger (1994).[27]

[INSERT TABLE 9 ABOUT HERE]

4.2.5 Out-of-sample model comparisons

We now consider out-of-sample methods for evaluating copula-based multivariate models. This is an important aspect of the evaluation of economic forecasts, see West (2006) for motivation and discussion. In this analysis, we estimate the model using an in-sample period (of length $R < T$) and evaluate it on the remaining $P = T - R$ observations (the "out-of-sample", OOS, period). Estimation of the model as we progress through the OOS period can be done in one of three ways. First, using "recursive" or "expanding window" estimation, where the forecast for observation t is based on data in the interval $[1, t-1]$. Alternatively, one can estimate the model using a "rolling" window, using data only in the interval $[t-R, t-1]$. This method is thought to provide some robustness against structural breaks in the data generating process, but involves "throwing away" observations from the start of the in-sample period. Finally, one can use "fixed window" estimation, where the model is estimated just once, using data from $[1, R]$. This latter method is useful when the model is computationally intensive to estimate. Let $\hat{\theta}_t$ denote the parameter vector of the

[27] Theoretically, the problem of an identified parameter under the null only appears when comparing constant and time-varying versions of the same copula (e.g., constant and time-varying Gumbel copulas), and does not arise when comparing copulas from different families (e.g., a constant Normal and a time-varying Gumbel). However, comparisons of constant and time-varying versions of the same copula are the most natural ones to consider, and thus this problem cannot be so easily avoided.

multivariate density obtained for a forecast of $\mathbf{Y}_t$ using one of these three estimation methods.[28]

A useful way to compare multivariate (or univariate) density forecasts, is to compare their OOS log-likelihood values, see Diks, *et al.* (2010) for example. Averaging across the OOS period, this can be interpreted as measuring the (negative of the) Kullback-Leibler distance of the density forecast from the true, unknown, conditional density, and so a model with a higher OOS log-likelihood is interpreted as being closer to the truth.[29]

$$\bar{L}_{OOS} \equiv \frac{1}{P}\sum_{t=R+1}^{T} \log \mathbf{f}_t\left(Y_{1t},...,Y_{nt};\hat{\theta}_t\right) \tag{73}$$

Using the fact that a multivariate log-likelihood can be decomposed into the marginal log-likelihoods and the copula, note that the difference between two multivariate log-likelihoods with the same marginal distributions is equal to the difference solely between their copula log-likelihoods:

$$\begin{aligned}\log \mathbf{f}_t^{(a)}\left(Y_{1t},...,Y_{nt}\right) - \log \mathbf{f}_t^{(b)}\left(Y_{1t},...,Y_{nt}\right) &= \log \mathbf{c}_t^{(a)}\left(F_{1t}\left(Y_{1t}\right),...,F_{nt}\left(Y_{nt}\right)\right) \\ &\quad - \log \mathbf{c}_t^{(b)}\left(F_{1t}\left(Y_{1t}\right),...,F_{nt}\left(Y_{nt}\right)\right)\end{aligned} \tag{74}$$

This is particularly useful for semiparametric multivariate models using the EDF for the marginal distributions: without further assumptions that model does not provide marginal densities and so the marginal log-likelihoods are not available.

The OOS evaluation of predictive models differs not only according to whether the models are fully parametric or semiparametric (as we have observed in numerous instances above), but also in the treatment of the parameter estimation error in the forecasts. Giacomini and White (2006) consider OOS forecasting models that are based on an estimation window of finite length (i.e., a fixed or rolling estimation scheme), and consider the forecast performance of two competing models conditional on their estimated parameters:

$$\begin{aligned} H_0 &: E\left[\log \mathbf{c}_1\left(\hat{\mathbf{U}}_t;\hat{\gamma}_{1t}\right) - \log \mathbf{c}_2\left(\hat{\mathbf{U}}_t;\hat{\gamma}_{2t}\right)\right] = 0 \\ \text{vs} \quad H_1 &: E\left[\log \mathbf{c}_1\left(\hat{\mathbf{U}}_t;\hat{\gamma}_{1t}\right) - \log \mathbf{c}_2\left(\hat{\mathbf{U}}_t;\hat{\gamma}_{2t}\right)\right] > 0 \\ H_2 &: E\left[\log \mathbf{c}_1\left(\hat{\mathbf{U}}_t;\hat{\gamma}_{1t}\right) - \log \mathbf{c}_2\left(\hat{\mathbf{U}}_t;\hat{\gamma}_{2t}\right)\right] < 0 \end{aligned} \tag{75}$$

[28] Note that although $\hat{\theta}_t$ has a subscript "t", it uses data only up until $t-1$ (recursive or rolling window) or until $R < t$ (fixed window). The subscript refers to the realization of the target variable, $\mathbf{Y}_t$.

[29] One could also consider weighted likelihoods, placing more emphasis on particular regions of the support, such as the tails versus the center or the left tail versus the right tail, see Amisano and Giacomini (2007) Gneiting and Ranjan (2011) and Diks, *et al.* (2011).

Importantly, the estimated parameters appear in the null, so a good model that is badly estimated will be punished. This has some particularly useful features for evaluating copula-based models: Firstly, we can compare both nested and non-nested models. In fact, we can even compare the *same* model estimated in two different ways (e.g., using one-stage MLE or MSMLE). Secondly, we do not need to pay special attention to whether the model is fully parametric or semiparametric. The asymptotic framework of Giacomini and White (2006) requires no adjustments for the estimated parameters of the models being compared, and the limiting distribution of the test statistic is $N(0,1)$. The only complication is that a HAC estimate of the asymptotic variance is required, as the differences in log-likelihoods may be serially correlated and heteroskedastic.

When the estimation window is expanding *and* the model is fully parametric, one can instead use the framework of West (1996). In this case the null and alternative hypotheses relate to the probability limit of the estimated parameters, denoted γ_1^* and γ_2^*.

$$H_0 : E\left[\log \mathbf{c}_1\left(\mathbf{U}_t; \gamma_1^*\right) - \log \mathbf{c}_2\left(\mathbf{U}_t; \gamma_2^*\right)\right] = 0 \tag{76}$$

In West's (1996) framework, the estimation error in $\hat{\theta}_t$ will affect the asymptotic variance of the t-statistic, and he provides a consistent estimator of the extra terms that need to be estimated. He notes that this estimation error can be ignored if $P/R \to 0$ as $P, R \to \infty$ (i.e., the estimation window is "large" relative to the OOS period), or if the comparison of model accuracy is done using the same loss function as used in estimation, and so if we estimate the marginals and the copula by ML (one-stage *or* multi-stage) *and* we compare models using their joint log-likelihood, then West's test is numerically identical to the Giacomini and White (2006) test, although the tests differ in their statement of the null hypothesis and thus in the interpretation of the result. It is important to note that West's (1996) approach can only be applied to non-nested models[30], and only to fully parametric models; the extension to consider semiparametric multivariate density models has not been treated in the literature, to my knowledge.

4.2.6 Empirical illustration, continued

We now consider out-of-sample comparisons of the various copula-based multivariate models applied to S&P 100 and S&P 600 index returns. These comparisons will be done using the joint log-

[30] McCracken (2007) considers nested models in this framework, but only linear specifications, and so cannot generally be used in multivariate density forecasting applications.

likelihood, and within the parametric and semiparametric groups of models this simplifies to a comparison of the copula log-likelihoods. In all cases we will use the Giacomini and White (2006) test, with the in-sample period being the first ten years of the sample period (August 17, 1995 to August 17, 2005, so $R = 2519$ observations) and the OOS period being the remainder (August 18, 2005 to May 20, 2011, so $P = 1450$ observations). To simplify the problem, we will consider a fixed estimation window, and only estimate the models once, using the first R observations.

The top panel of Table 10 reports the t-statistics of pair-wise comparisons. We find that all but one pair-wise comparison is significant, indicating good power to differentiate between these models, and the best model turns out to be the Student's t-GAS model, followed by the Rotated Gumbel-GAS model. Both of these models beat all of the constant copula models, consistent with our earlier findings of significant evidence of time-varying dependence, and with the GoF test results discussed in Section 4.1.3. The same conclusions are found for pair-wise comparisons of semiparametric models, presented in the middle panel of Table 10.

The bottom row of Table 10 presents results from tests to compare multivariate models with the same copula but different models for the marginal distributions, either the parametric skew t distribution, or a nonparametric estimate. As noted above, the nonparametric estimate we use is the EDF, and does not have an unique counterpart for the density, which is needed to compute the log-likelihood. To overcome this for this test, one can use a nonparametric density estimate, such as one based on a Gaussian kernel with Silverman's bandwidth.[31] The results in Table 10 indicate that for all choices of copula model the parametric density estimate is preferred to the nonparametric estimate in terms of OOS fit (the t-statistics are all positive), however only for the time-varying copulas are these differences (borderline) significant, with t-statistics of 1.99 and 1.80.

[INSERT TABLE 10 ABOUT HERE]

5 Other issues in applications

In this section we will discuss two examples of estimation and computation issues that arise when applying copulas to multivariate time series. We will consider the general case that the conditional

[31] This is the default kernel density estimate using Matlab's "pltdens.m" function. The bandwidth is $1.06\hat{\sigma}T^{-1/5}$, where $\hat{\sigma}^2$ is the sample variance of the standardized residuals (which is 1.00 for both series).

copula is time-varying, which of course nests the constant conditional copula case. Let

$$\begin{aligned} (U_{1t}, U_{2t})\,|\mathcal{F}_{t-1} &\sim \mathbf{C}(\delta_t) \\ \text{where} \quad \delta_t &= \delta(\mathbf{Z}_{t-1}, \gamma^*), \quad \text{for } \mathbf{Z}_{t-1} \in \mathcal{F}_{t-1} \end{aligned} \tag{77}$$

We will assume below that the marginal distributions are estimated using the EDF, but all of the methods also apply for parametric marginal models.

5.1 Linear correlation in copula-based multivariate models

The upper and lower tail dependence implied by many well-known copulas are known in closed form, see Table 4 for example. The tail dependence implied by the time-varying rotated Gumbel and Student's t GAS copula models are presented in Figure 5.

[INSERT FIGURE 5 ABOUT HERE]

Corresponding formulas for rank correlation are often not available, and formulas for the more familiar *linear* correlation are never available, as the linear correlation depends both upon on the copula model and the marginal distribution specification. While linear correlation has its drawbacks as a measure of dependence, it is still the most widely-known in economics and it is often useful to present as a summary of the linear dependence implied by a given model. Given the specification for our multivariate time series model in equation (6), the conditional correlation of the two variables can be expressed as:

$$\begin{aligned} \rho_t &\equiv Corr_{t-1}[Y_{1t}, Y_{2t}] = Corr_{t-1}[\varepsilon_{1t}, \varepsilon_{2t}] \\ &= E_{t-1}[\varepsilon_{1t}\varepsilon_{2t}], \quad \text{since } \varepsilon_{it}|\mathcal{F}_{t-1} \sim F_i(0,1) \\ &= E_{t-1}\left[F_1^{-1}(U_{1t})\,F_2^{-1}(U_{2t})\right], \quad \text{since } U_{it} \equiv F_i(\varepsilon_{it}) \end{aligned} \tag{78}$$

The last expression cannot usually be obtained analytically, however two numerical approaches are available. The first is to use two-dimensional numerical integration:[32]

$$E_{t-1}\left[F_1^{-1}(U_{1t})\,F_2^{-1}(U_{2t})\right] \equiv \int_0^1 \int_0^1 F_1^{-1}(u_1)\,F_2^{-1}(u_2)\,\mathbf{c}(u_1, u_2; \delta_t(\gamma))\,du_1 du_2 \tag{79}$$

[32] For example, via the built-in function "dblquad.m" in Matlab.

An alternative approach is to use simulation:

$$E_{t-1}\left[F_1^{-1}(U_{1t})\,F_2^{-1}(U_{2t})\right] \approx \frac{1}{S}\sum_{s=1}^{S} F_1^{-1}\left(u_1^{(s)}\right)F_2^{-1}\left(u_2^{(s)}\right) \tag{80}$$
$$\text{where } \left(u_1^{(s)},u_2^{(s)}\right) \sim iid\ \mathbf{C}\left(\delta_t(\gamma)\right),\quad s=1,2,...,S$$

where S is the number of simulations (e.g., $S = 1000$). When the copula is time-varying, these simulations need to be done for each day in the sample, as each day will have a different value for the copula parameter. When the sample size is large this can be quite a computational burden (although the problem is parallelizable).

One way to reduce the number of computations across days in the sample is to exploit the fact that for many copulas the mapping from copula parameter to correlation is smooth, and so one can compute this mapping for a reduced number of values of the copula parameter (its minimum and maximum value over the sample period, and, e.g., 10 evenly-spaced values in between) and then use interpolation to obtain the correlation.[33] Note that this grid must cover *all* time-varying parameters in the copula and the distributions of the standardized residuals. For example, if we allowed both the correlation and the degrees of freedom parameter to change in the Student's t copula then we need a grid of, say, 10×10 values.[34] The spline approach is particularly useful when there are few varying parameters in the copula and marginal distributions; when this gets even moderately large, it may be faster to simply do the simulation for each day of the sample.

Before relying on interpolation it is of course important to check that the function is indeed smooth. Figure 6 presents the interpolated mapping from the Gumbel parameter and t copula correlation parameter (the degrees of freedom parameter was held fixed at the value reported in Table 6) to the linear correlation that is obtained, using the EDF for the marginal distributions. This mapping was estimated using 10 equally spaced nodes and 100,000 simulations, and is shown to be a good approximation from a comparison with the mapping using 20 equally spaced nodes. With this mapping it is fast to get the implied linear correlation for the entire time series (3969 dates), and this is plotted in Figure 7.

[33] Given a fixed amount of computing time there is often a trade-off between the number of nodes at which to compute the correlation, and the precision of the estimate at each node. Since the interpolation step takes the values at the nodes as the *true* values, it is very important to make sure that these are as accurate as possible. Thus it is usually better to have fewer nodes estimated very precisely than many nodes estimated imprecisely.

[34] Further, if the marginal distributions of the standardized residuals were allowed to vary through time (e.g., with time-varying skewness and kurtosis) then a grid would need to cover variations in these parameters too.

[INSERT FIGURES 6 AND 7 ABOUT HERE]

5.2 Value-at-Risk and Expected Shortfall in copula-based multivariate models

Multivariate models of financial time series are often used in risk management, and two key measures of risk are Value-at-Risk and Expected Shortfall. (See the chapter by Komunjer in this Handbook for a review of methods for VaR forecasting.) For a portfolio return Y_t, with conditional distribution F_t, these measures are defined as

$$\begin{aligned} VaR_t^q &\equiv F_t^{-1}(q), \ \text{ for } q \in (0,1) \\ ES_t^q &\equiv E\left[Y_t | \mathcal{F}_{t-1}, Y_t \leq VaR_t^q\right], \ \text{ for } q \in (0,1) \end{aligned} \tag{81}$$

That is, the $q\%$ VaR is the q^{th} percentile of the conditional distribution, and the corresponding ES is the expected value of Y_t conditional on it lying below its VaR. When the joint distribution of the variables of interest is elliptical (e.g., Normal or Student's t) the distribution of any linear combination of these variables (such as a portfolio return) is known in closed form. When more flexible models are used for the marginal distributions and the copula the distribution of linear combinations of the variables is generally *not* known in closed form, and obtaining these risk measures requires a different approach.

One simple means of obtaining the VaR and ES of a portfolio of variables whose distribution is modelled using a copula-based approach is via simulation. At each point in the sample, we generate S observations from the multivariate model, then form the portfolio return, and then use the empirical distribution of those simulated portfolio returns to estimate the VaR and ES measures. For values of q closer to zero or one, larger values of S are required.

Figure 8 presents results for an equal-weighted portfolio with $q = 0.01$, and use $S = 5000$ simulations on each date. We can see that the VaR ranges from around -2% at the start of the sample, to -14% at the height of the financial crisis. Expected Shortfall ranges from around -3% and is as low as -17%. The risk estimates implied by the Rotated Gumbel model are below those from the Student's t model on around 70% of the days, consistent with the much greater lower tail dependence implied by this copula.

To better see the differences in the VaR and ES estimates implied by the two copulas, Figure 9 presents the values of these measures for an equal-weighted portfolio of returns with mean zero and variance one, using the empirical distribution of the standardized residuals for the S&P 100 and S&P 600 for the marginal distributions, for rank correlation ranging from zero to 0.99. A spline is used to map rank correlation to the Gumbel and Student's t copula parameters, analogous to that for linear correlation discussed in the previous section. To estimate the VaR and ES measures for each level of rank correlation 1 million simulations are used. This figure yields two main insights. First, the differences between the predicted VaR and ES from the various models are greatest for more extreme quantiles: the 0.1% VaR and ES measures vary more across copulas than the corresponding measures at the 1% level. This is consistent with the observation that these copulas have broadly similar implications for the middle of the distribution, but can differ more substantially in the joint tails. Second, the differences between these copulas are greatest for rank correlation around 0.3 to 0.7. This is intuitive, given that for rank correlation 1 (implying perfect positive dependence, or "comonotonicity") these models are identical, and for rank correlation of zero both the Gumbel and Normal copulas imply independence, while not so for the Student's t copula if $\nu < \infty$. We can see from Figure 9 that all three copula models yield identical results for rank correlation equal to one, and that the rotated Gumbel and Normal copulas yield the same risk estimates when rank correlation is zero, while the Student's t copula indicates slightly more risk (for this figure I used the estimated degrees of freedom from the time-varying t copula, which was 15.4). Thus the range of rank correlations where there is the greatest possibility of different estimates of risk and dependence is around 0.3 to 0.7, which happens to be around the values observed for many financial asset returns.

[INSERT FIGURES 8 AND 9 ABOUT HERE]

6 Applications of copulas in economics and finance

In this section we review some of the many applications of copulas in economics and finance, broadly categorized into the areas of application.

6.1 Risk management

One of the first areas of application of copulas in economics and finance was risk management. The focus of risk managers on Value-at-Risk (VaR), and other measures designed to estimate the probability of large losses, leads to a demand for flexible models of the dependence between sources of risk. See Komunjer (2011) for a recent review of VaR methods. Hull and White (1998), Cherubini and Luciano (2001), Embrechts, *et al.* (2002, 2003) and Embrechts and Höing (2006) study the VaR of portfolios. Rosenberg and Schuermann (2006) use copulas to consider 'integrated' risk management problems, where market, credit and operational risks must be considered jointly. McNeil, *et al.* (2005) and Alexander (2008) provide clear and detailed textbook treatments of copulas and risk management.

6.2 Derivative contracts

Another early application of copulas was to the pricing of credit derivatives (credit default swaps and collateralized debt obligations, for example), as these contracts routinely involve multiple underlying sources of risk. Li (2000) was first to use copulas in a credit risk application, see also Frey and McNeil (2001), Schönbucher and Schubert (2001), Giesecke (2004) Hofert and Scherer (2011) and Duffie (2011) for applications to default risk. Applications of copulas in other derivatives markets include Rosenberg (2003), Bennett and Kennedy (2004), Cherubini, *et al.* (2004), van den Goorbergh, *et al.* (2005), Salmon and Schleicher (2006), Grégoire, *et al.* (2008), Taylor and Wang (2010), and Cherubini, *et al.* (2012).

6.3 Portfolio decision problems

Considering portfolio decision problems in their most general form involves finding portfolio weights that maximize the investor's expected utility, and thus requires a predictive multivariate distribution for the assets being considered. Applications of copulas in portfolio problems include Patton (2004), who considers a bivariate equity portfolio problem using time-varying copulas; Hong, *et al.* (2007) consider an investment decision involving eleven equity portfolios under "disappointment aversion" preferences; Christoffersen and Langlois (2011) consider portfolio decisions involving four common equity market factors; Garcia and Tsafack (2011) consider portfolio decisions involving stocks and bonds in two countries; and Christoffersen, *et al.* (2011) consider a time-varying copula model for 33 developed and emerging equity market indices.

6.4 Time-varying copula models

The econometrics literature contains a wealth of evidence that the conditional volatility of economic time series changes through time, motivating the consideration of models that also allow the conditional copula to vary through time. Various models have been proposed in the literature to date. Patton (2002, 2004, 2006a), Jondeau and Rockinger (2006), Christoffersen, *et al.* (2011) and Creal, *et al.* (2011) consider models of time-varying copulas where the copula functional form is fixed and its parameter is allowed to vary through time as a function of lagged information, similar to the famous ARCH model for volatility, see Engle (1982) and Bollerslev (1986). "Stochastic copula" models, analogous to stochastic volatility models, see Shephard (2005), were proposed by Hafner and Manner (2010) and further studied in Manner and Segers (2011). "Locally constant" copula models are considered by Giacomini, Härdle and Spokoiny (2009), Guégan and Zhang (2009), Dias and Embrechts (2010), Harvey (2010), Rémillard (2010) and Busetti and Harvey (2011). Regime switching models, as in Hamilton (1989), for the conditional copula allow the functional form of the copula to vary through time and are considered by Rodriguez (2007), Okimoto (2008), Chollete, *et al.* (2009), Markwat, *et al.* (2009), Garcia and Tsafack (2011).

6.5 Other applications

There are several other noteworthy economic applications of copulas that do not neatly fit into one of the above categorizations. Breymann, *et al.* (2003) and Dias and Embrechts (2010) study the copulas of financial assets using intra-daily data sampled at different frequencies; Granger, *et al.* (2006) use copulas to provide a definition of a 'common factor in distribution'; Bartram, *et al.* (2007) use a time-varying conditional copula model to study financial market integration between seventeen European stock market indices; Heinen and Rengifo (2007) use copulas to model multivariate time series of count data; Rodriguez (2007) uses copulas to study financial contagion; Dearden, *et al.* (2008) and Bonhomme and Robin (2009) use copulas to model the dynamics in a panel of earnings data; Lee and Long (2009) use copulas to flexibly model the uncorrelated residuals of a multivariate GARCH model; Patton (2009b), Dudley and Nimalendran (2011) and Kang, *et al.* (2010) apply copulas to study dependence between hedge funds and other assets; and Zimmer (2011) studies the how simplified copula models relate to the recent U.S. housing crisis.

7 Conclusions and directions for further research

Copula-based multivariate models allow the researcher to specify the models for the marginal distributions separately from the dependence structure (copula) that links these distributions to form the joint distribution. This increases the flexibility of multivariate models that can be considered, and often reduces the computational complexity of estimating such models. This chapter reviews some of the empirical applications of copula-based methods in economics, and discusses in detail methods for estimation, inference, goodness-of-fit testing, and model selection that are useful when working with these models. Inference methods differ according to whether the marginal distributions are modelled parametrically or nonparametrically (leading respectively to a fully parametric or semiparametric multivariate model) and both cases are considered. A representative data set of two daily equity index returns is used to illustrate all of the main results.

In reviewing the literature to date, an outline of the "ideal" copula model emerges. An ideal copula model can accommodate dependence of either *sign* (positive or negative), it can capture both symmetric and *asymmetric* dependence, and it allows for the possibility of non-zero *tail dependence*. A truly ideal copula model might also possess a fourth attribute: *scalability*, to higher dimensions (more on this below). Most of the copulas in use empirically, see Table 3 for example, possess at least two of these attributes, and more recent research has lead to copula models that possess all three, and sometimes scalability, such as the skew t copula of Demarta and McNeil (2004) and the factor copula of Oh and Patton (2011b).

The literature on copula methods for economic and financial time series suggests two important directions for further research. The first is theoretical: methods for inference on semiparametric multivariate models with a time-varying conditional copula. These models have great empirical appeal: in many economic and financial applications there is sufficient data to reliably estimate a univariate distribution nonparametrically, and there is an abundance of evidence that the dependence between economic variables varies through time. Inference methods currently assume either the marginal distributions are parametric (Patton, 2006b), or the conditional copula is constant (Chen and Fan, 2006b; Rémillard, 2010). A block bootstrap method for inference for semiparametric multivariate models with a time-varying conditional copula was discussed in this chapter, but its use requires formal justification. An alternative approach based on a "multiplier central limit theorem", see Rémillard and Scaillet (2009) and Ruppert (2011) for details and discussion, may prove useful.

A second direction for further research is empirical: useful and feasible methods for modelling dependence in high dimensions. While bivariate and low dimension ($n < 10$) applications of copula-based models are still common, researchers have begun to consider higher dimension problems, up to around one hundred variables. For example, Daul *et al.* (2003) proposed a "grouped t" copula and show that this copula can be used in applications of up to 100 variables. Hofert and Scherer (2011) and Hering, *et al.* (2010) consider nested Archimedean copulas for modelling credit default swaps on 125 companies. Aas, *et al.* (2009) and Min and Czado (2010) consider multivariate "vine" copulas, which are constructed by sequentially applying bivariate copulas to build up a higher dimension copula, see Acar, *et al.* (2012) for an important critique of vine copulas. Oh and Patton (2011b) propose a new class of "factor copulas" for a collection of 100 equity returns. When taking models to high dimension applications one is inevitably forced to make some simplifying assumptions, and in different applications the set of plausible simplifying assumptions will vary. Increasing the variety of models available for such applications, and investigating their usefulness, will be an active area of research for some time.

References

[1] Aas, K., C. Czado, A. Frigessi, and H. Bakken, 2009, Pair-copula Constructions of Multiple Dependence, *Insurance: Mathematics and Economics*, 44, 182-198.

[2] Abegaz, F. and U.V. Naik-Nimbalkar, 2008, Modeling statistical dependence of Markov chains via copula models, *Journal of Statistical Planning and Inference*, 138, 1131-1146.

[3] Alexander, C., 2008, *Market Risk Analysis, Volume III*, Wiley & Sons, London.

[4] Amisano, G. and R. Giacomini, 2007, Comparing Density Forecasts via Weighted Likelihood Ratio Tests, *Journal of Business & Economic Statistics*, 25(2), 177-190.

[5] Andersen, T.G., T. Bollerslev, P. Christoffersen and F.X. Diebold., 2006, Volatility and Correlation Forecasting, in *Handbook of Economic Forecasting, Volume 1*, G. Elliott, C.W.J. Granger and A. Timmermann eds., Elsevier, Oxford.

[6] Andrews, D.W.K., 1993, Tests for Parameter Instability and Structural Change with an Unknown Change Point, *Econometrica*, 61, 821-856.

[7] Andrews, D.W.K., 2001, Testing When a Parameter Is on the Boundary of the Maintained Hypothesis, *Econometrica*, 69, 2001, 683-734.

[8] Andrews, D.W.K. and W. Ploberger, 2001, Optimal Tests When a Nuisance Parameter Is Present Only Under the Alternative, *Econometrica*, 62, 1994, 1383-1414.

[9] Bartram, S.M., S.J. Taylor and Y.-H. Wang, 2007, The euro and European financial market dependence, *Journal of Banking and Finance*, 51(5), 1461-1481.

[10] Beare, B.K., 2010, Copulas and temporal dependence, *Econometrica*, 78, 395-410.

[11] Bennett, M.N. and J.E. Kennedy, 2004, Quanto Pricing with Copulas, *Journal of Derivatives*, 12(1), 26-45.

[12] Berg, D., 2009, Copula goodness-of-fit testing: An overview and power comparison, *European Journal of Finance*, 15, 675-701.

[13] Bollerslev, T., 1986, Generalized Autoregressive Conditional Heteroskedasticity, *Journal of Econometrics*, 31, 307-327.

[14] Bollerslev, T., 1987, A Conditionally Heteroskedastic Time Series Model for Speculative Prices and Rates of Return, *Review of Economics and Statistics*, 69(3), 542-547.

[15] Bollerslev, T., 2010, Glossary to ARCH (GARCH), *in Volatility and Time Series Econometrics: Essays in Honor of Robert F. Engle*, T. Bollerslev, J.R. Russell and M.W. Watson eds, Oxford University Press, Oxford.

[16] Bonhomme, S., and J.-M. Robin, 2009, Assessing the Equalizing Force of Mobility Using Short Panels: France, 1990-2000, *Review of Economic Studies*, 76(1), 63-92.

[17] Bontemps, C., B. Feunou and N. Meddahi, 2011, Testing Distributional Assumptions: The Multivariate Case, working paper, Toulouse School of Economics.

[18] Breymann, W., A. Dias, and P. Embrechts, 2003, Dependence structures for multivariate high-frequency data in finance, *Quantitative Finance,* 3, 1-16.

[19] Busetti, F. and A. Harvey, 2011, When is a Copula Constant? A Test for Changing Relationships, *Journal of Financial Econometrics*, 9(1), 106-131.

[20] Capéraà, P., A.-L. Fourgères, and C. Genest, 1997, A nonparametric estimation procedure for bivariate extreme value copulas, *Biometrika*, 84(3), 567-577.

[21] Carrasco, M. and X. Chen, 2002, Mixing and moment properties of various GARCH and stochastic volatility models, *Econometric Theory* 18, 17-39.

[22] Casella, G. and R.L. Berger, 1990, *Statistical Inference*, Duxbury Press, U.S.A.

[23] Chan, N.-H., J. Chen, X. Chen, Y. Fan and L. Peng, 2009, Statistical Inference for Multivariate Residual Copula of GARCH Models, *Statistica Sinica*, 19, 53-70.

[24] Chen, Y.-T., 2007, Moment-Based Copula Tests for Financial Retrusn, *Journal of Business & Economic Statistics*, 25(4), 377-397.

[25] Chen, Y.-T., 2011, Moment Tests for Density Forecast Evaluation in the Presence of Parameter Estimation Uncertainty, *Journal of Forecasting*, 30, 409–450.

[26] Chen, X., and Y. Fan, 2006a, Estimation of copula-based semiparametric time series models, *Journal of Econometrics*, 130, 307-335.

[27] Chen, X., and Y. Fan, 2006b, Estimation and model selection of semiparametric copula-based multivariate dynamic models under copula misspecification, *Journal of Econometrics*, 135, 125-154.

[28] Chen, X., Y. Fan and V. Tsyrennikov, 2006, Efficient estimation of semiparametric multivariate copula models, *Journal of the American Statistical Association*, 101 (475), 1228-1240.

[29] Chen, X., Y. Fan, D. Pouzo and Z. Yang, 2010, Estimation and model selection of semiparametric multivariate survival functions under general censorship, *Journal of Econometrics*, 157, 129-142.

[30] Chen, X., W.B. Wu, and Y. Yi, 2009. Efficient estimation of copula-based semiparametric Markov models, *Annals of Statistics*, 37, 4214-4253.

[31] Cherubini, U. and E. Luciano, 2001, Value at Risk trade-off and capital allocation with copulas, *Economic Notes*, 30, 235-256.

[32] Cherubini, U., E. Luciano and W. Vecchiato, 2004, *Copula Methods in Finance*, John Wiley & Sons, England.

[33] Cherubini, U., F. Gobbi, S. Mulinacci and S. Romagnoli, 2012, *Dynamic Copula Methods in Finance*, John Wiley & Sons, England.

[34] Chollete, L., A. Heinen, and A. Valdesogo, 2009, Modeling International Financial Returns with a Multivariate Regime-switching Copula, *Journal of Financial Econometrics*, 7, 437-480.

[35] Choros, B., R. Ibragimov and E. Permiakova, 2010, Copula estimation, in *Workshop on Copula Theory and its Applications, Lecture Notes in Statistics - Proceedings*, F. Durante, W. Härdle, P. Jaworski, and T. Rychlik, eds., Springer.

[36] Christoffersen, P. and H. Langlois, 2011, The Joint Dynamics of Equity Market Factors, working paper, Rotman School of Management, University of Toronto.

[37] Christoffersen, P., V. Errunza, K. Jacobs and H. Langlois, 2011, Is the Potential for International Diversification Disappearing?, working paper, Rotman School of Management, University of Toronto.

[38] Corradi, V. and N.R. Swanson, 2006, Predictive Density Evaluation, in *Handbook of Economic Forecasting*, G. Elliott, C.W.J. Granger and A. Timmermann, eds., North Holland, Amsterdam.

[39] Creal, D., S.J. Koopman and A. Lucas, 2011, Generalized Autoregressive Score Models with Applications, *Journal of Applied Econometrics*, forthcoming.

[40] Darsow, W. F., B. Nguyen, and E.T. Olsen, 1992, Copulas and Markov processes, *Illinois Journal of Mathematics*, 36, 600-642.

[41] Daul, S., E. De Giorgi, F. Lindskog and A. McNeil, 2003, The grouped t-copula with an application to credit risk, *RISK*, 16, 73-76.

[42] de Haan, L., C. Neves and L. Peng, 2008, Parametric tail copula estimation and model testing, *Journal of Multivariate Analysis*, 99, 1260-1275.

[43] Dearden, L., E. Fitzsimons, A. Goodman and G. Kaplan, 2008, Higher Education Funding Reforms in England: The Distributional Effects and the Shifting Balance of Costs, *Economic Journal*, 118(526), 100-125.

[44] Demarta, S. and A.J. McNeil, 2005, The t copula and related copulas, *International Statistical Review*, 73, 111-129.

[45] Dias, A. and P. Embrechts, 2010, Modeling exchange rate dependence dynamics at different time horizons, *Journal of International Money and Finance* 29, 1687-1705.

[46] Diebold, F.X., J. Hahn and A.S. Tay, 1999, Multivariate Density Forecast Evaluation and Calibration in Financial Risk Management: High Frequency Returns on Foreign Exchange, *Review of Economics and Statistics*, 81, 661-673.

[47] Diks, C., V. Panchenko, and D. van Dijk, 2010, Out-of-sample comparison of copula specifications in multivariate density forecasts, *Journal of Economic Dynamics and Control*, 34(9), 1596-1609.

[48] Diks, C., V. Panchenko, and D. van Dijk, 2011, Likelihood-based scoring rules for comparing density forecasts in tails, *Journal of Econometrics*, 163(2), 215-230.

[49] Dudley, E. and M. Nimalendran, 2011, Margins and Hedge Fund Contagion, *Journal of Financial and Quantitative Analysis*, 46, 1227-1257.

[50] Duffie, D., 2011, *Measuring Corporate Default Risk*, Clarendon Lectures in Finance, Oxford University Press, Oxford.

[51] Embrechts, P. and A. Höing, 2006, Extreme VaR scenarios in higher dimensions, *Extremes*, 9, 177-192.

[52] Embrechts, P., A. McNeil and D. Straumann, 2002, Correlation and Dependence Properties in Risk Management: Properties and Pitfalls, in M. Dempster, ed., *Risk Management: Value at Risk and Beyond*, Cambridge University Press.

[53] Embrechts, P., A. Höing and A. Juri, 2003, Using Copulae to bound the Value-at-Risk for functions of dependent risks, *Finance & Stochastics*, 7, 145-167.

[54] Engle, R.F., 1982, Autoregressive Conditional Heteroscedasticity with Estimates of the Variance of UK Inflation, *Econometrica,* 50, 987-1007.

[55] Engle, R.F. and K.F. Kroner, 1995, Multivariate Simultaneous Generalized ARCH, *Econometric Theory*, 11(1), 122-150.

[56] Fermanian, J.-D. and O. Scaillet, 2003, Nonparametric estimation of copulas for time series, *Journal of Risk*, 5(4), 25-54.

[57] Fermanian, J.-D. and M. Wegkamp, 2012, Time dependent copulas, *Journal of Multivariate Analysis*, forthcoming.

[58] Fermanian, J.-D., D. Radulović and M. Wegkamp, 2004, Weak convergence of empirical copula processes, *Bernoulli*, 10(5), 847-860.

[59] Frahm, G., M. Junker and R. Schmidt, 2005, Estimating the tail-dependence coefficient: Properties and pitfalls, *Insurance: Mathematics and Economics,* 37, 80-100.

[60] Frees, E.W. and E.A. Valdez, 1998, Understanding Relationships using Copulas, *North American Actuarial Journal*, 2(1), 1-25.

[61] Frey, R. and A.J. McNeil, 2001, Modelling dependent defaults, ETH E-Collection, http://e-collection.ethbib.ethz.ch/show?type=bericht&nr=273.

[62] Gaißer, S., M. Ruppert and F. Schmid, 2010, A multivariate version of Hoeffding's Phi-Square, *Journal of Multivariate Analysis*, 101, 2571-2586.

[63] Garcia, R. and G. Tsafack, 2011, Dependence Structure and Extreme Comovements in International Equity and Bond Markets, *Journal of Banking & Finance*, 35(8), 1954-1970.

[64] Genest, C., 1987, Frank's family of bivariate distributions, *Biometrika*, 74(3), 549-555.

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

[66] Genest, C. and A.-C. Favre, 2007, Everything you always wanted to know about copula modeling but were afraid to ask, *Journal of Hydrologic Engineering*, 12, 347-368.

[67] Genest, C., K. Ghoudi and L.-P. Rivest, 1995, A semiparametric estimation procedure of dependence parameters in multivariate families of distributions, *Biometrika*, 82(3), 543-552.

[68] Genest, C. and B. Rémillard, 2008, Validity of the Parametric Bootstrap for Goodness-of-Fit Testing in Semiparametric Models, *Annales de l'Institut Henri Poincaré*, 44(6), 1096-1127.

[69] Genest, C., B. Rémillard and D. Beaudoin, 2009, Omnibus Goodness-of-Fit Tests for Copulas: A Review and a Power Study, *Insurance: Mathematics and Economics*, 44, 199-213.

[70] Ghoudi, K. and B. Rémillard, 2004, Empirical processes based on pseudo-observations II: The multivariate case, in *Asymptotic Methods in Stochastics: Festschrift for Miklós Csörgö*, C.M. Cuadras, J. Fortiana, and J.A. Rodríguez-Lallena, eds., Kluwer Academic, Dordrecht.

[71] Giacomini, R. and H. White, 2006, Tests of Conditional Predictive Ability, *Econometrica*, 74(6), 1545–1578.

[72] Giacomini, E., W. Härdle, and V. Spokoiny, 2009, Inhomogeneous dependence modeling with time-varying copulae, *Journal of Business & Economic Statistics*, 27, 224-234.

[73] Giesecke, K., 2004, Correlated Default with Incomplete Information, *Journal of Banking and Finance*, 28, 1521-1545.

[74] Glosten, L.R., R. Jagannathan, D.E. Runkle, 1993, On the Relation between the Expected Value and the Volatility of the Nominal Excess Return on Stocks, *Journal of Finance*, 48(5), 1779-1801.

[75] Gneiting, T. and R. Ranjan, 2011, Comparing Density Forecasts using Threshold- and Quantile-Weighted Scoring Rules, *Journal of Business & Economic Statistics*, 29(3), 411-422.

[76] Gonçalves, S. and H. White, 2004, Maximum Likelihood and the Bootstrap for Nonlinear Dynamic Models, *Journal of Econometrics*, 119, 199-220.

[77] Granger, C.W.J., T. Teräsvirta and A.J. Patton, 2006, Common factors in conditional distributions for bivariate time series, *Journal of Econometrics*, 132, 43-57.

[78] Grégoire, V., C. Genest and M. Gendron, 2008, Using copulas to model price dependence in energy markets, *Energy Risk*, 5(5), 58-64.

[79] Guégan, D. and J. Zhang, 2009, Change analysis of dynamic copula for measuring dependence in multivariate financial data, *Quantitative Finance*, 10(4), 421-430.

[80] Hafner, C.M. and H. Manner, 2010, Dynamic Stochastic Copula Models: Estimation, Inference and Applications, *Journal of Applied Econometrics*, forthcoming.

[81] Hamilton, J.D., 1989, A New Approach to the Economic Analysis of Nonstationary Time Series and the Business Cycle, *Econometrica*, 57, 357-384.

[82] Hansen, B.E., 1994, Autoregressive conditional density estimation, *International Economic Review*, 35(3), 705-730.

[83] Hansen, P.R., A. Lunde and J.M. Nason, 2011, Model Confidence Sets for Forecasting Models, *Econometrica*, forthcoming.

[84] Harvey, A., 2010, Tracking a Changing Copula, *Journal of Empirical Finance*, 17, 485-500.

[85] Heinen, A. and E. Rengifo, 2007, Multivariate autoregressive modeling of time series count data using copulas, *Journal of Empirical Finance*, 14(4), 564-583.

[86] Heinen, A. and A. Valdesogo, 2009, Asymmetric CAPM dependence for large dimensions: the Canonical Vine Autoregressive Model, CORE Discussion Paper 2009-069.

[87] Hering, C., M. Hofert, J.-F. Mai and M. Scherer, 2010, Constructing hierarchical Archimedean copulas with Lévy subordinators, *Journal of Multivariate Analysis*, 101(6), 1428-1433.

[88] Hofert, M. and M. Scherer, 2011, CDO pricing with nested Archimedean copulas, *Quantitative Finance*, 11(5), 775–787.

[89] Hong, Y., J. Tu and G. Zhou, 2007, Asymmetries in stock returns: Statistical tests and economic evaluation, *Review of Financial Studies* 20, 1547–1581.

[90] Hu, L., 2006, Dependence patterns across financial markets: a mixed copula approach, *Applied Financial Economics*, 16(10), 717-729.

[91] Hull, J. and A. White, 1998, Value at Risk when daily changes in market variables are not normally distributed, *Journal of Derivatives*, 5, 9-19.

[92] Ibragimov, R., 2009, Copula-based characterizations for higher-order Markov processes, *Econometric Theory,* 25, 819-846.

[93] Joe, H., 1997, *Multivariate Models and Dependence Concepts*, Monographs in Statistics and Probability 73, Chapman and Hall, London.

[94] Joe, H., 2005, Asymptotic efficiency of the two-stage estimation method for copula-based models, *Journal of Multivariate Analysis*, 94, 401-419.

[95] Joe, H. and J.J. Xu, 1996, The Estimation Method of Inference Functions for Margins for Multivariate Models, working paper, Department of Statistics, University of British Columbia.

[96] Jondeau, E. and M. Rockinger, 2003, Conditional Volatility, Skewness, and Kurtosis: Existence, Persistence, and Comovements, *Journal of Economic Dynamics and Control*, 27, 1699-1737.

[97] Jondeau, E. and M. Rockinger, 2006, The copula-GARCH model of conditional dependencies: an international stock market application, *Journal of International Money and Finance*, 25(5), 827-853.

[98] Kang, B.U., F. In, G. Kim and Y.S. Kim, 2010, A Longer Look at the Asymmetric Dependence between Hedge Funds and the Equity Market, *Journal of Financial and Quantitative Analysis*, 45(3), 763-789.

[99] Komunjer, I., 2011, Quantile Prediction, in *Handbook of Economic Forecasting, Volume 2*, G. Elliott and A. Timmermann eds., Elsevier, Oxford.

[100] Lee, T.-H. and X. Long, 2009, Copula-based multivariate GARCH model with uncorrelated dependent standardized returns, *Journal of Econometrics*, 150(2), 207-218.

[101] Li, D.X., 2000, On default correlation: a copula function approach, *Journal of Fixed Income*, 9, 43-54.

[102] Manner, H. and O. Reznikova, 2011, A survey on time-varying copulas: Specification, simulations and estimation, *Econometric Reviews*, forthcoming.

[103] Manner, H. and J. Segers, 2011, Tails of correlation mixtures of elliptical copulas, *Insurance: Mathematics and Economics*, 48(1), 153-160.

[104] Markwat, T.D., E. Kole and D.J.C. van Dijk, 2009, Time Variation in Asset Return Dependence: Strength or Structure?, working paper, Econometric Institute, Erasmus University Rotterdam.

[105] McCracken, M.W., 2007, Asymptotics for Out of Sample Tests of Granger Causality, *Journal of Econometrics*, 140, 719-752.

[106] McNeil, A.J., R. Frey and P. Embrechts, 2005, *Quantitative Risk Management: Concepts, Techniques and Tools*, Princeton University Press, New Jersey.

[107] Meitz, M. and P. Saikkonen, 2008, Ergodicity, mixing, and the existence of moments of a class of Markov models with applications to GARCH and ACD models, *Econometric Theory*, 24, 1291-1320.

[108] Min, A. and C. Czado, 2010, Bayesian inference for multivariate copulas using pair-copula constructions, *Journal of Financial Econometrics*, 8(4), 450-480.

[109] Nelsen, R.B., 2006, *An Introduction to Copulas*, Second Edition, Springer, U.S.A.

[110] Oh, D.-H. and A.J. Patton, 2011a, Simulated Method of Moments Estimation for Copula-Based Multivariate Models, working paper, Duke University.

[111] Oh, D.-H. and A.J. Patton, 2011b, Modelling Dependence in High Dimensions with Factor Copulas, working paper, Duke University.

[112] Okimoto, T., 2008, New evidence of asymmetric dependence structure in international equity markets, *Journal of Financial and Quantitative Analysis*, 43, 787–815.

[113] Patton, A.J., 2002, Applications of Copula Theory in Financial Econometrics, Unpublished Ph.D. dissertation, University of California, San Diego.

[114] Patton, A.J., 2004, On the Out-of-Sample Importance of Skewness and Asymmetric Dependence for Asset Allocation, *Journal of Financial Econometrics*, 2(1), 130-168.

[115] Patton, A.J., 2006a, Modelling Asymmetric Exchange Rate Dependence, *International Economic Review*, 47(2), 527-556.

[116] Patton, A.J., 2006b, Estimation of Multivariate Models for Time Series of Possibly Different Lengths, *Journal of Applied Econometrics*, 21(2), 147-173.

[117] Patton, A.J., 2009a, Copula-based models for financial time series, in T.G. Andersen, R.A. Davis, J.-P. Kreiss and T. Mikosch (eds.) *Handbook of Financial Time Series*, Springer Verlag.

[118] Patton, A.J., 2009b, Are "Market Neutral" Hedge Funds Really Market Neutral?, *Review of Financial Studies*, 22(7), 2495-2530.

[119] Patton, A.J., 2012, A Review of Copula Models for Economic Time Series, *Journal of Multivariate Analysis*, forthcoming.

[120] Pickands, J., 1981, Multivariate Extreme Value Distributions, *Bulletin de l'Institut International de Statistique*, 859-878.

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

[122] Rémillard, B., 2010, Goodness-of-fit tests for copulas of multivariate time series, working paper, HEC Montreal.

[123] Rémillard, B. and O. Scaillet, 2009, Testing for Equality Between Two Copulas, *Journal of Multivariate Analysis*, 100(3), 377-386.

[124] Rémillard, B., N. Papageorgiou and F. Soustra, 2012, Copula-based semiparametric models for multivariate time series, *Journal of Multivariate Analysis*, forthcoming.

[125] Rivers, D. and Q. Vuong, 2002, Model Selection Tests for Nonlinear Dynamic Models, *The Econometrics Journal*, 5(1), 1-39.

[126] Rodriguez, J.C., 2007, Measuring financial contagion: a copula approach, *Journal of Empirical Finance*, 14(3), 401-423.

[127] Romano, J.P. and M. Wolf, 2005. Stepwise Multiple Testing as Formalized Data Snooping. *Econometrica* 73, 1237-1282.

[128] Rosenberg, J.V., 2003, Nonparametric pricing of multivariate contingent claims, *Journal of Derivatives*, 10, 9-26.

[129] Rosenberg, J.V. and T. Schuermann, 2006, A general approach to integrated risk management with skewed, fat-tailed risks, *Journal of Financial Economics*, 79, 569-614.

[130] Ruppert, M., 2011, Consistent testing for a constant copula under strong mixing based on the tapered block multiplier technique, working paper, Department of Economic and Social Statistics, University of Cologne.

[131] Salmon, M. and C. Schleicher, 2006, Pricing Multivariate Currency Options with Copulas, in J. Rank, ed., *Copulas: From Theory to Application in Finance*, Risk Books, London.

[132] Sancetta, A. and S. Satchell, 2004, The Bernstein copula and its applications to modeling and approximations of multivariate distributions, *Econometric Theory*, 20, 2004, 535–562

[133] Schönbucher, P. and D. Schubert, 2001, Copula Dependent Default Risk in Intensity Models, mimeo, Bonn University.

[134] Shephard, N., 2005, *Stochastic Volatility: Selected Readings,* Oxford University Press, Oxford.

[135] Sklar, A., 1959, Fonctions de répartition à n dimensions et leurs marges, *Publications de l' Institut Statistique de l'Universite ́ de Paris*, 8, 229-231.

[136] Smith, M., A. Min, C. Almeida, and C. Czado, 2010a, Modeling longitudinal data using a pair-copula decomposition of serial dependence, *Journal of the American Statistical Association*, 105 (492), 1467-1479.

[137] Smith M., Q. Gan, and R. Kohn, 2010b, Modeling Dependence Using Skew t Copulas: Bayesian Inference and Applications, *Journal of Applied Econometrics*, forthcoming.

[138] Smith, M.S., 2011, Bayesian Approaches to Copula Modelling, in *Hierarchical Models and MCMC: A Tribute to Adrian Smith*, forthcoming.

[139] Taylor, S.J., and Y. Wang, 2010, Option prices and risk-neutral densities for currency cross-rates, *Journal of Futures Markets* 30, 324-360.

[140] Teräsvirta, T., 2006, Forecasting Economic Variables with Nonlinear Models, in *Handbook of Economic Forecasting, Volume 1*, G. Elliott, C.W.J. Granger and A. Timmermann eds., Elsevier, Oxford.

[141] Tsukahara, H., 2005, Semiparametric estimation in copula models, *Canadian Journal of Statistics*, 33(3), 357-375.

[142] van den Goorbergh, R.W.J., C. Genest and B.J.M. Werker, 2005, Multivariate Option Pricing Using Dynamic Copula Models, *Insurance: Mathematics and Economics*, 37, 101-114.

[143] West, K. D., 1996, Asymptotic Inference About Predictive Ability, *Econometrica*, 64(5), 1067-1084.

[144] West, K. D., 2006, Forecast Evaluation, in *Handbook of Economic Forecasting, Volume 1*, G. Elliott, C.W.J. Granger and A. Timmermann eds., Elsevier, Oxford.

[145] White, H., 1994, *Estimation, Inference and Specification Analysis*, Econometric Society Monographs No. 22, Cambridge University Press, Cambridge, U.K.

[146] White, H., 2000, A Reality Check for Data Snooping, *Econometrica* 68, 1097-1126.

[147] White, H., 2006, Approximate Nonlinear Forecasting Methods, in *Handbook of Economic Forecasting, Volume 1*, G. Elliott, C.W.J. Granger and A. Timmermann eds., Elsevier, Oxford.

[148] Zimmer, D., 2011, The Role of Copulas in the Housing Crisis, *Review of Economics and Statistics*, forthcoming.

Table 1: Summary statistics and marginal distribution parameter estimates

	S&P 100	*S&P 600*
	Summary statistics	
Mean	0.020	0.033
Std dev	1.297	1.426
Skewness	-0.151	-0.302
Kurtosis	10.021	7.962
Correl (lin/rnk)	0.837 / 0.782	
	Conditional mean	
ϕ_0	0.023	0.033
ϕ_1	-0.078	–
ϕ_2	-0.067	–
	Conditional variance	
ω	0.017	0.029
α	0.001	0.017
δ	0.134	0.149
β	0.919	0.892
	Skew t density	
λ	-0.145	-0.140
ν	9.936	19.808
	GoF tests	
KS p-value	0.124	0.093
CvM p-value	0.479	0.222

Notes: This table presents summary statistics and other results for daily returns on the S&P 100 and S&P 600 indices over the period August 1995 to May 2011. The top panel presents summary statistics, including linear and rank correlations; the second panel presents parameter estimates from AR(2) and AR(0) models for the conditional mean; the third panel presents parameter estimates from GJR-GARCH(1,1) models for the conditional variance; the fourth panel presents parameter estimates from skew t models for the distribution of the standardized residuals; the bottom panel presents simulation-based p-values from two Kolmogorov-Smirnov and Cramer-von Mises goodness-of-fit tests for the models of the conditional marginal distributions, using 1000 simulations.

Table 2: Testing for time-varying dependence

	Break				AR(p)		
	0.15	0.50	0.85	Anywhere	1	5	10
p-val	0.667	0.373	0.045	0.269	0.417	0.054	0.020

Notes: This table presents p-values from tests for time varying rank correlation between the standardized residuals of the S&P 100 and S&P 600 indices, based on 1000 bootstrap simulations. The left panel considers tests that allow for a one-time break in rank correlation. The right panel considers tests for autocorrelation in $U_{it}U_{jt}$.

Table 3: Some common copula models

	Parameter(s)	Parameter space	Indep	Pos & Neg dep?	Rank correlation	Kendall's τ	Lower tail dep	Upper tail dep
Normal	ρ	$(-1,1)$	0	*Yes*	$\frac{6}{\pi}\arcsin\frac{\rho}{2}$	$\frac{2}{\pi}\arcsin\rho$	0	0
Clayton	γ	$(0,\infty)$	0	*No*†	*n.a.*	$\frac{\gamma}{\gamma+2}$	$2^{-1/\gamma}$	0
Rotated Clayton	γ	$(0,\infty)$	0	*No*†	*n.a.*	$\frac{\gamma}{\gamma+2}$	0	$2^{-1/\gamma}$
Plackett	γ	$(0,\infty)$	1	*Yes*	$\frac{\gamma^2-2\gamma\log\gamma-1}{(\gamma-1)^2}$	*n.a.*	0	0
Frank	γ	$(-\infty,\infty)$	0	*Yes*	$g_\rho(\gamma)$	$g_\tau(\gamma)$	0	0
Gumbel	γ	$(1,\infty)$	1	*No*	*n.a.*	$\frac{\gamma-1}{\gamma}$	0	$2-2^{1/\gamma}$
Rotated Gumbel	γ	$(1,\infty)$	1	*No*	*n.a.*	$\frac{\gamma-1}{\gamma}$	$2-2^{1/\gamma}$	0
Sym Joe-Clayton	$\tau^L,\ \tau^U$	$[0,1)\times[0,1)$	$(0,0)$	*No*	*n.a.*	*n.a.*	τ^L	τ^U
Student's t	$\rho,\ \nu$	$(-1,1)\times(2,\infty)$	$(0,\infty)$	*Yes*	*n.a.*	$\frac{2}{\pi}\arcsin(\rho)$	$g_T(\rho,\nu)$	$g_T(\rho,\nu)$

Notes: This table presents some common parametric copula models, along with their parameter spaces, and analytical forms for some common measures of dependence, if available. . For more details on these copulas see Joe (1997, Chapter 5) or Nelsen (2006, Chapters 4-5). Measures that are not available in closed form are denoted "*n.a.*". Parameter values that lead to the independence copula are given in the column titled "Indep". Frank copula rank correlation: $g_\rho(\gamma) = 1 - 12\left(D_1(\gamma) - D_2(\gamma)\right)/\gamma$ and Frank copula Kendall's tau: $g_\tau(\gamma) = 1 - 4\left(1 - D_1(\gamma)\right)/\gamma$, where $D_k(x) = kx^{-k}\int_0^x t^k\left(e^t - 1\right)^{-1} dt$ is the "Debye" function, see Nelsen (2006). Student's t copula lower and upper tail dependence: $g_T(\rho,\nu) = 2 \times F_{Studt}\left(-\sqrt{(\nu+1)\frac{\rho-1}{\rho+1}}, \nu+1\right)$, see Demarta and McNeil (2005). †The Clayton (and rotated Clayton) copula allows for negative dependence for $\gamma \in (-1,0)$, however the form of this dependence is different from the positive dependence case $(\gamma > 0)$, and is not generally used in empirical work.

Table 4: Estimates of tail dependence

	Nonparametric		Parametric	
	"log"	*"sec"*	*Gumbel*	*Student's t*
	Lower tail dependence: $\hat{\lambda}^L$			
Estimate	0.411	0.414	0.390	0.266
90% CI	$[0.112, 0.664]$	$[0.105, 0.658]$	$[0.321, 457]$	$[0.221, 0.349]$
	Upper tail dependence: $\hat{\lambda}^U$			
Estimate	0.230	0.233	0.270	0.149
90% CI	$[0.021, 0.537]$	$[0.021, 0.549]$	$[0.185, 0.354]$	$[0.081, 0.170]$
pval for $\lambda^L = \lambda^U$	0.850	0.842	0.411	0.245

Notes: This table presents four estimates of the lower and upper tail dependence coefficients for the standardized residuals of the S&P 100 and S&P 600 indices. 90% confidence intervals based on 1000 bootstrap replications are also presented. The bottom row presents bootstrap *p*-values from tests that the upper and lower tail dependence coefficients are equal.

Table 5: Constant copula model parameter estimates

	Parametric		Semiparametric	
	Est. Param.	$\log \mathcal{L}$	Est. Param.	$\log \mathcal{L}$
Normal	0.7959	**1991.8**	0.7943	**1978.3**
Clayton	2.0279	1720.5	2.0316	1723.1
Rotated Clayton	1.6914	1414.5	1.6698	1396.2
Plackett	18.8405	1976.2	18.7224	1964.8
Frank	7.8969	1904.1	7.8019	1882.0
Gumbel	2.2637	1826.5	2.2480	1803.4
Rotated Gumbel	2.3715	**2013.6**	2.3673	**2008.4**
Sym Joe-Clayton (τ^L, τ^U)	0.6639 , 0.5378	1980.8	0.6649 , 0.5318	1967.8
Student's t (ρ, ν^{-1})	0.8019 , 0.1455	**2057.4**	0.8005 , 0.1428	**2041.9**

Notes: This panel presents the estimated parameters of nine different models for the copula of the standardized residuals of the S&P 100 and S&P 600 indices. The value of the copula log-likelihood at the optimum is also presented, and the best three models are in bold. The left panel presents results when the marginal distributions are modelled using a skew t distribution; the right panel presents results when the marginal distributions are estimated using the empirical distribution function.

Table 6: Standard errors on estimated constant copula parameters

		Parametric				Semiparametric			
		Naïve	*MSML*	*Boot*	*Sim*	*Naïve*	*MSML*	*Boot*	*Sim*
Normal	$\hat{\rho}$		0.7959				0.7943		
	s.e.	0.0046	0.0108	0.0099	0.0062	0.0046	0.0061	0.0065	0.0055
	$\log \mathcal{L}$		1991.8				1978.3		
Clayton	$\hat{\kappa}$		2.0279				2.0316		
	s.e.	0.0451	0.0961	0.0862	0.0664	0.0449	0.0545	0.0580	0.0701
	$\log \mathcal{L}$		1720.5				1723.1		
Rotated	$\hat{\kappa}$		2.3715				2.3673		
Gumbel	s.e.	0.0310	0.0610	0.0595	0.0386	0.0309	0.0421	0.0344	0.0420
	$\log \mathcal{L}$		2013.6				2008.4		
Student's t	$\hat{\rho}$		0.8019				0.8005		
	s.e.	0.0053	0.0101	0.0096	0.0070	0.0053	0.0055	0.0054	0.0067
	$\hat{\nu}^{-1}$		0.1455				0.1428		
	s.e.	0.0172	0.0206	0.0222	0.0186	0.0172	0.0182	0.0169	0.0203
	$\log \mathcal{L}$		2057.4				2041.9		

Note: This table presents the estimated parameters of four different copula models for the standardized residuals for the S&P 100 and the S&P 600 indices, when the marginal distributions are estimated using a skewed t distribution (left panel) or the empirical distribution function (right panel). For the parametric model four different estimators of the standard error on the estimated parameter are presented, and for the semiparametric model three different standard errors are presented. For all models the log-likelihood at the estimated parameter is also presented.

Table 7: Standard errors on estimated time-varying copula parameters

		Parametric				Semiparametric	
		Naïve	*MSML*	*Boot*	*Sim*	*Naïve*	*Boot*
Rotated	$\hat{\omega}$		0.0013			0.0015	
Gumbel		0.0012	0.0051	0.0069	0.0013	0.0011	0.0072
GAS	$\hat{\alpha}$		0.0404			0.0420	
		0.0124	0.0298	0.0175	0.0076	0.0110	0.0176
	$\hat{\beta}$		0.9961			0.9955	
		0.0028	0.0096	0.0165	0.0026	0.0029	0.0172
	$\log \mathcal{L}$		2127.3			2117.3	
Student's t	$\hat{\omega}$		0.0199			0.0192	
GAS		0.0012	0.0142	0.0381	0.0090	0.0093	0.0382
	$\hat{\alpha}$		0.0653			0.0603	
		0.0091	0.0166	0.0189	0.0100	0.0296	0.0182
	$\hat{\beta}$		0.9912			0.9913	
		1.9×10^{-6}	0.0119	0.0164	0.0038	0.0284	0.0167
	$\hat{\nu}^{-1}$		0.0887			0.0891	
		0.0133	0.0415	0.0181	0.0174	0.0515	0.0185
	$\log \mathcal{L}$		2203.6			2184.6	

Note: This table presents the estimated parameters of two different time-varying copula models for the standardized residuals for the S&P 100 and the S&P 600 indices, when the marginal distributions are estimated using a skewed t distribution (left panel) or the empirical distribution function (right panel). For the parametric model four different estimators of the standard error on the estimated parameter are presented, and for the semiparametric model two different standard errors are presented. For all models the log-likelihood at the estimated parameter is also presented.

Table 8: Goodness of fit tests for copula models

	Naïve				*Simulation*			
	KS_C	CvM_C	KS_R	CvM_R	KS_C	CvM_C	KS_R	CvM_R
				Parametric				
Normal	0.30	0.26	**0.00**	**0.00**	0.10	0.09	**0.00**	**0.00**
Clayton	**0.00**	**0.00**	**0.00**	0.06	**0.00**	**0.00**	**0.00**	**0.01**
Rot. Gumbel	0.42	0.32	0.18	0.15	0.09	**0.02**	0.09	0.06
Student's t	0.47	0.39	0.09	0.13	0.35	0.13	**0.04**	0.07
Rot. Gumbel-GAS	–	–	0.11	0.18	–	–	0.99	1.00
Student's t-GAS	–	–	0.07	0.07	–	–	0.08	0.08
				Semiparametric				
Normal	0.43	0.48	**0.04**	**0.00**	**0.00**	**0.00**	0.00	**0.00**
Clayton	**0.00**	**0.00**	0.08	0.014	**0.00**	**0.00**	**0.00**	**0.01**
Rot. Gumbel	0.43	0.53	0.61	0.41	**0.00**	**0.00**	**0.02**	**0.00**
Student's t	0.65	0.74	0.40	0.13	**0.00**	**0.00**	**0.02**	**0.00**
Rot. Gumbel-GAS	–	–	0.78	0.27	–	–	1.00	1.00
Student's t-GAS	–	–	0.47	0.08	–	–	**0.03**	**0.00**

Note: This table presents the p-values from various tests of goodness-of-fit for four different copula models for the standardized residuals for the S&P 100 index and the S&P 600 index, when the marginal distributions are estimated parametrically (top panel) or nonparametrically (lower panel). KS and CvM refer to the Kolmogorov-Smirnov and Cramer-von Mises tests respectively. The subscripts C and R refer to whether the test was applied to the empirical copula of the standardized residuals, or to the empirical copula of the Rosenblatt transform of these residuals. The p-values are based on 100 simulations. The left panel presents p-values that (incorrectly) ignore parameter estimation error, the right panel present results that take this estimation error into account. p-values less than 0.05 are in bold.

Table 9: In-sample model comparisons for constant copula models

	Normal	*Clayton*	*Rot Gumbel*	*Student's t*
		Parametric		
Normal	–			
Clayton	*-7.24*	–		
Rot. Gumbel	0.93	**15.59**	–	
Student's t	**7.06**†	**10.00**	**2.58**	–
$\log \mathcal{L}$	1991.8	1720.5	2013.6	2057.4
Rank	3	4	2	1
		Semiparametric		
Normal	–			
Clayton	*-6.27*	–		
Rot. Gumbel	1.16	**16.32**	–	
Student's t	**7.85**†	**8.80**	1.67	–
$\log \mathcal{L}$	1978.3	1723.1	2008.4	2041.9
Rank	3	4	2	1

Note: This table presents t-statistics from Rivers and Vuong (2002) (upper panel) and Chen and Fan (2006b) model comparison tests for four constant copula models. A positive value indicates that the model to the left is better than the model above, and a negative value indicates the opposite. The average value of the log-likelihood for each model is also presented. †The Student's t copula nests the Normal copula, and so a standard t-test can be used to compare these models. t-statistics that are greater than 1.96 are in bold, and those less than -1.96 are in italics.

Table 10: Out-of-sample model comparisons

	Normal	*Clayton*	*RGum*	*Stud t*	*RGum-GAS*	*Stud t-GAS*
				Parametric		
Normal	–					
Clayton	*-10.05*	–				
RGum	0.96	**18.81**	–			
Stud t	**9.39**	**12.67**	**3.87**	–		
RGum GAS	**5.94**	**15.81**	**8.57**	**4.43**	–	
Stud-t GAS	**9.89**	**14.74**	**10.35**	**9.46**	**4.99**	–
$\log L_C^{OOS}$	914.8	770.91	923.39	952.69	1017.16	1069.15
Rank	5	6	4	3	2	1
				Semiparametric		
Normal	–					
Clayton	*-9.90*	–				
RGum	0.71	**18.36**	–			
Stud t	**9.34**	**12.39**	**3.71**	–		
RGum GAS	**5.47**	**15.79**	**8.29**	**3.99**	–	
Stud-t GAS	**9.85**	**14.99**	**10.55**	**9.43**	**5.15**	–
$\log L_C^{OOS}$	912.74	765.90	919.30	948.33	1007.64	1062.07
Rank	5	6	4	3	2	1
				Parametric vs Nonparametric margins		
t-stat	0.91	1.35	1.23	1.39	**1.99**	1.80

Note: This table presents t-statistics from out-of-sample pair-wise comparisons of the log-likelihood values for four constant copula models and two time-varying copula models, with fully parametric or semiparametric marginal distribution models. A positive value indicates that the model to the left is better than the model above, and a negative value indicates the opposite. The out-of-sample value of the log-likelihood for each model is also presented. The bottom row of the table presents t-statistics from pair-wise comparisons of bivariate density models with the same copula specification but with either nonparametric or skew t marginal distributions, and a positive value indicates that the model with skew t marginal distributions is preferred. t-statistics that are greater than 1.96 are in bold, and those less than -1.96 are in italics.

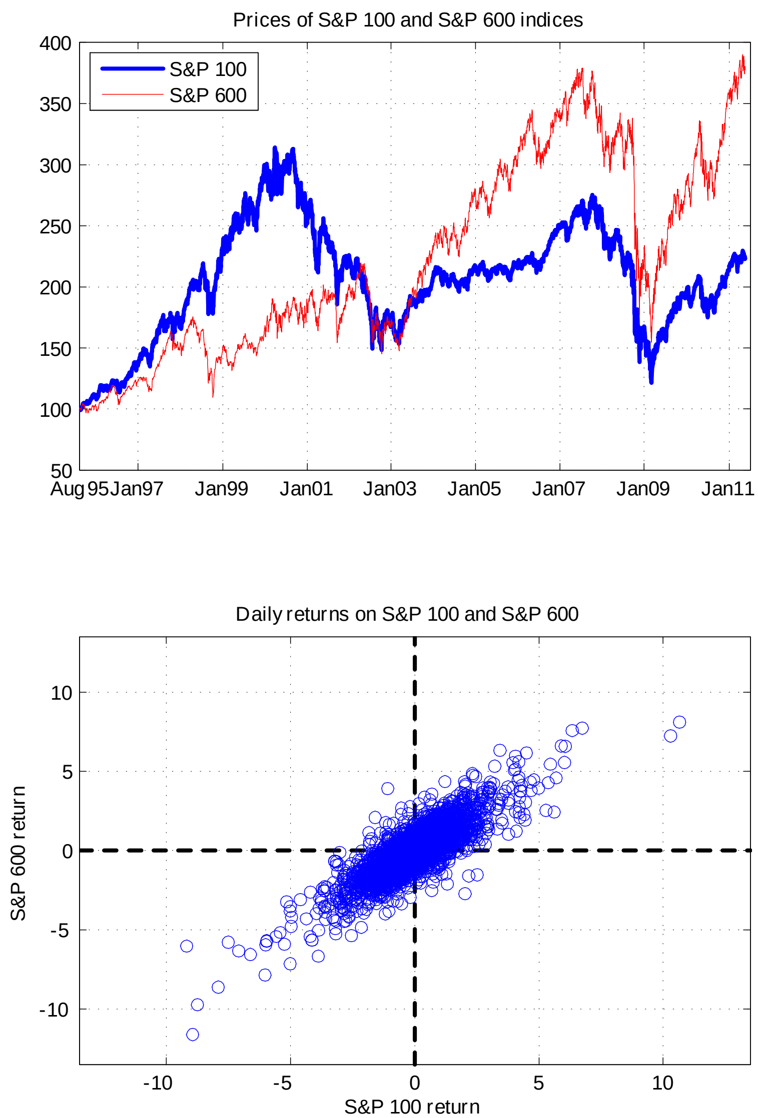


Figure 1: *The upper panel of this figure shows the level of the S&P100 and S&P 600 indices over the period August 1995 to May 2011, normalized to 100 at the start of the sample period. The lower panel shows a scatter plot of daily returns on these indices.*

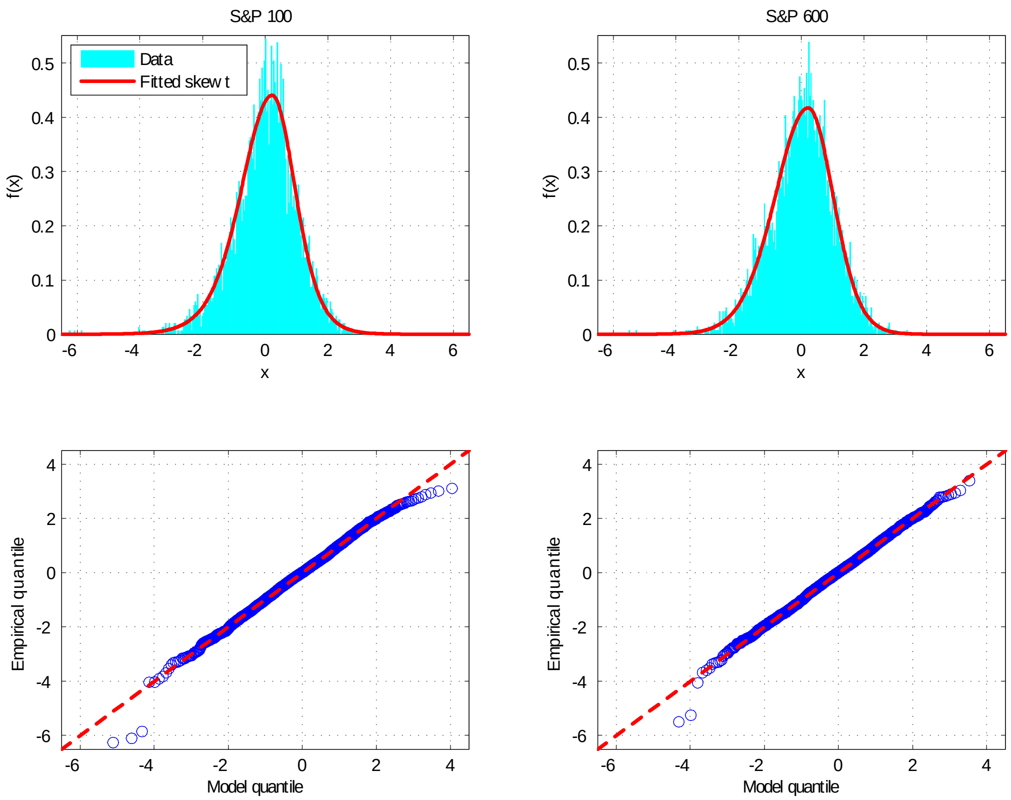


Figure 2: *The upper panels of this figure present the fitted skew t density for the S&P100 and S&P 600 standardized residuals, along with histograms of these residuals; the lower panels present QQ plots.*

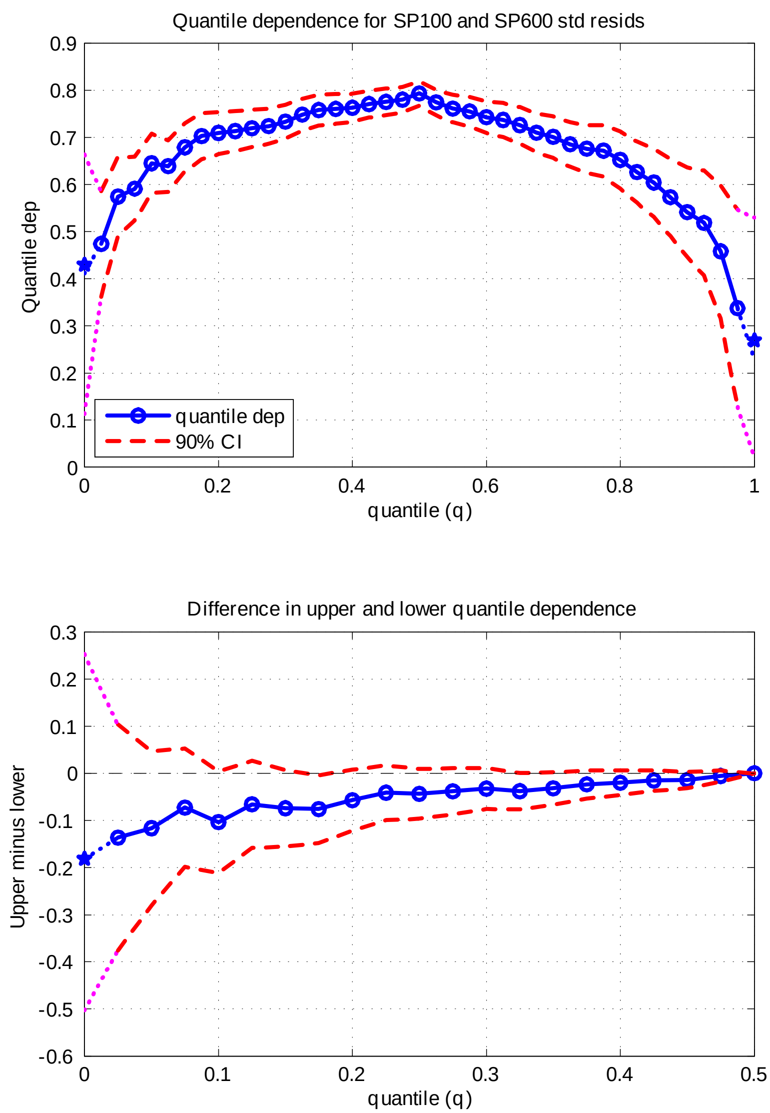


Figure 3: *The upper panel shows the estimated quantile dependence between the standardized residuals for the S&P 100 index and the S&P 600 index, and the upper and lower tail dependence coefficients estimated using a Gumbel tail copula, along with 90% bootstrap confidence intervals. The lower panel presents the difference between corresponding upper and lower quantile and tail dependence estimates, along with a 90% bootstrap confidence interval for this difference.*

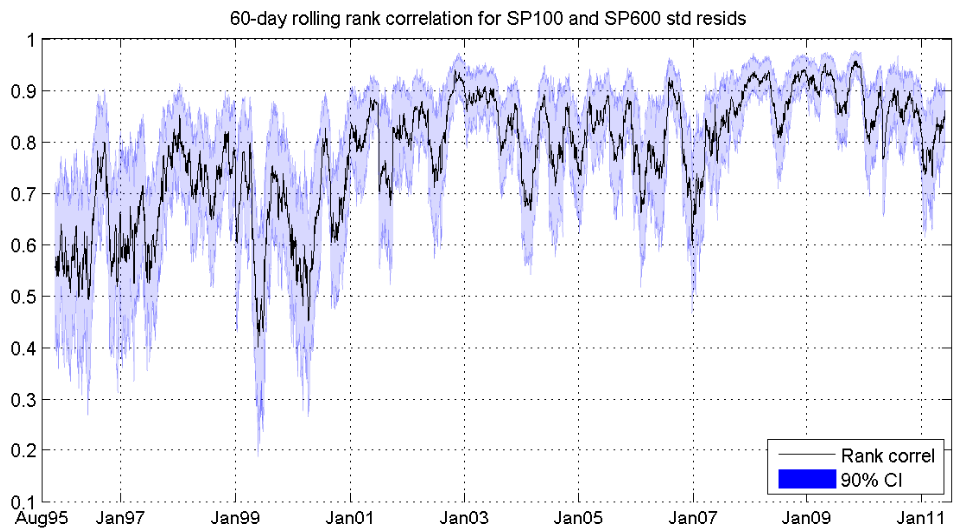


Figure 4: *This figure shows the rank correlation between the standardized residuals for the S&P 100 index and the S&P 600 index over a 60-day moving window, along with 90% bootstrap confidence intervals.*

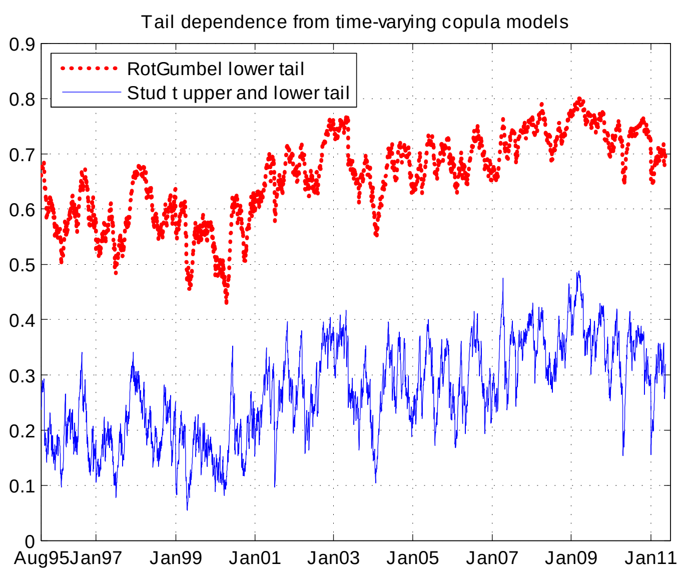


Figure 5: *Conditional tail dependence from the time-varying rotated Gumbel and Student's t copula models.*

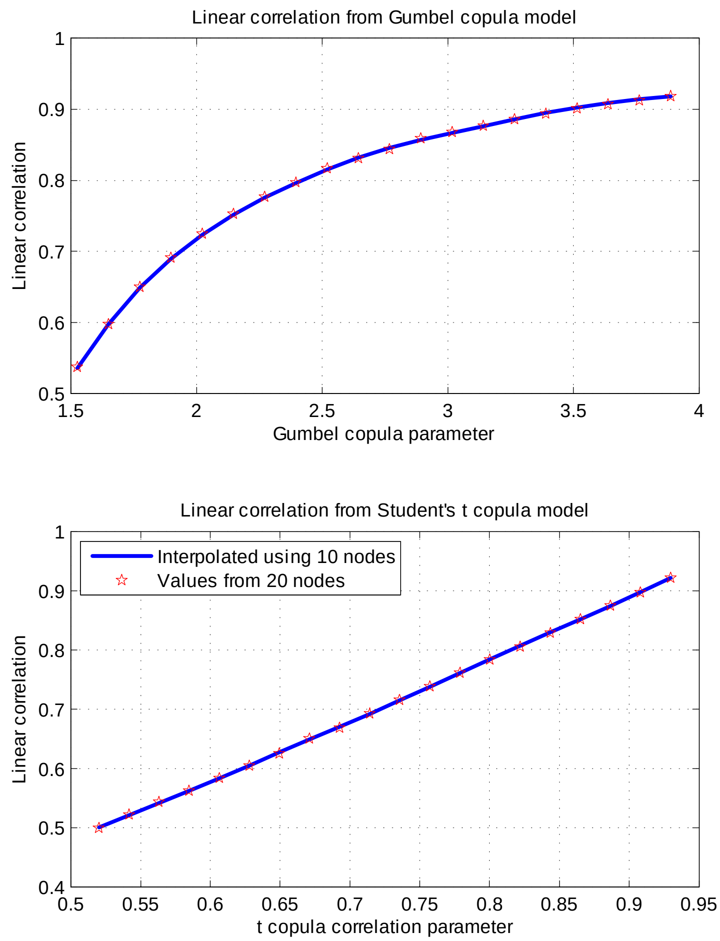


Figure 6: *Spline for linear correlation implied by Gumbel and Student's t copula models, when combined with the empirical distributions of the standardized residuals of the S&P 100 and S&P 600 indices, compared with actual values at 20 points.*

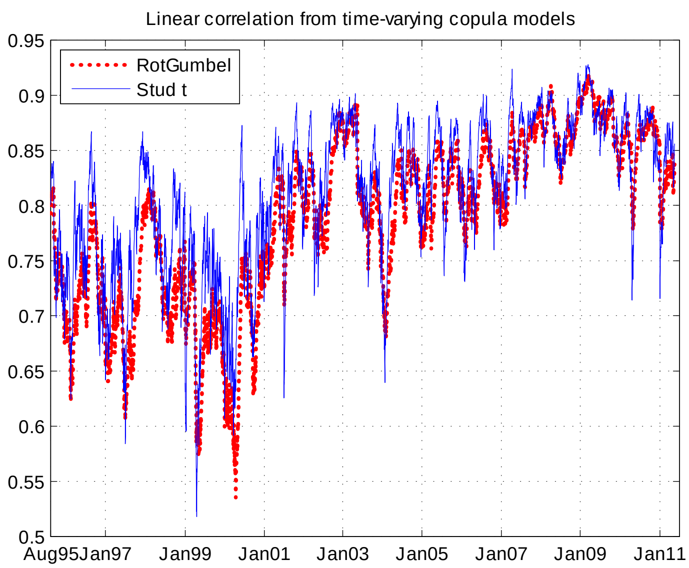
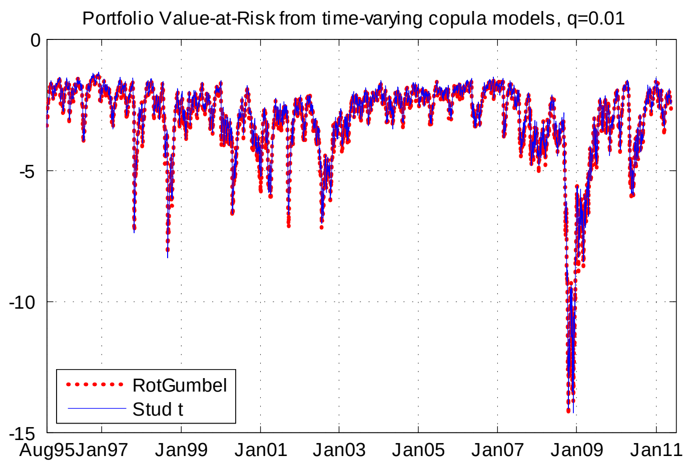


Figure 7: *Conditional correlation from the time-varying rotated Gumbel and Student's t copula models.*



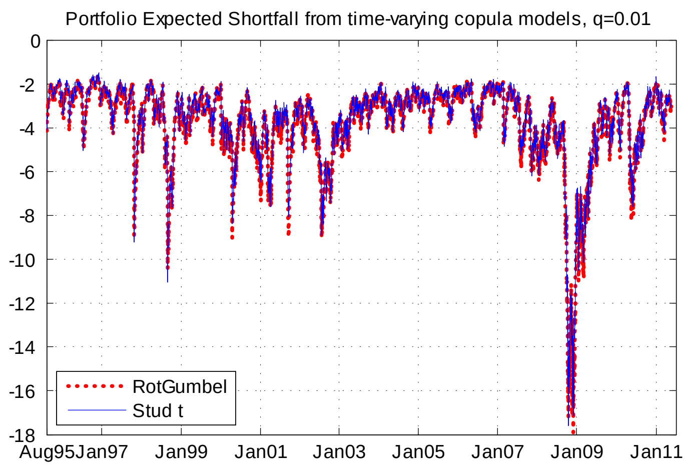


Figure 8: *Conditional 1% Value-at-Risk (upper panel) and Expected Shortfall (lower panel) for an equal-weighted portfolio, based on the time-varying rotated Gumbel and Student's t copula models.*

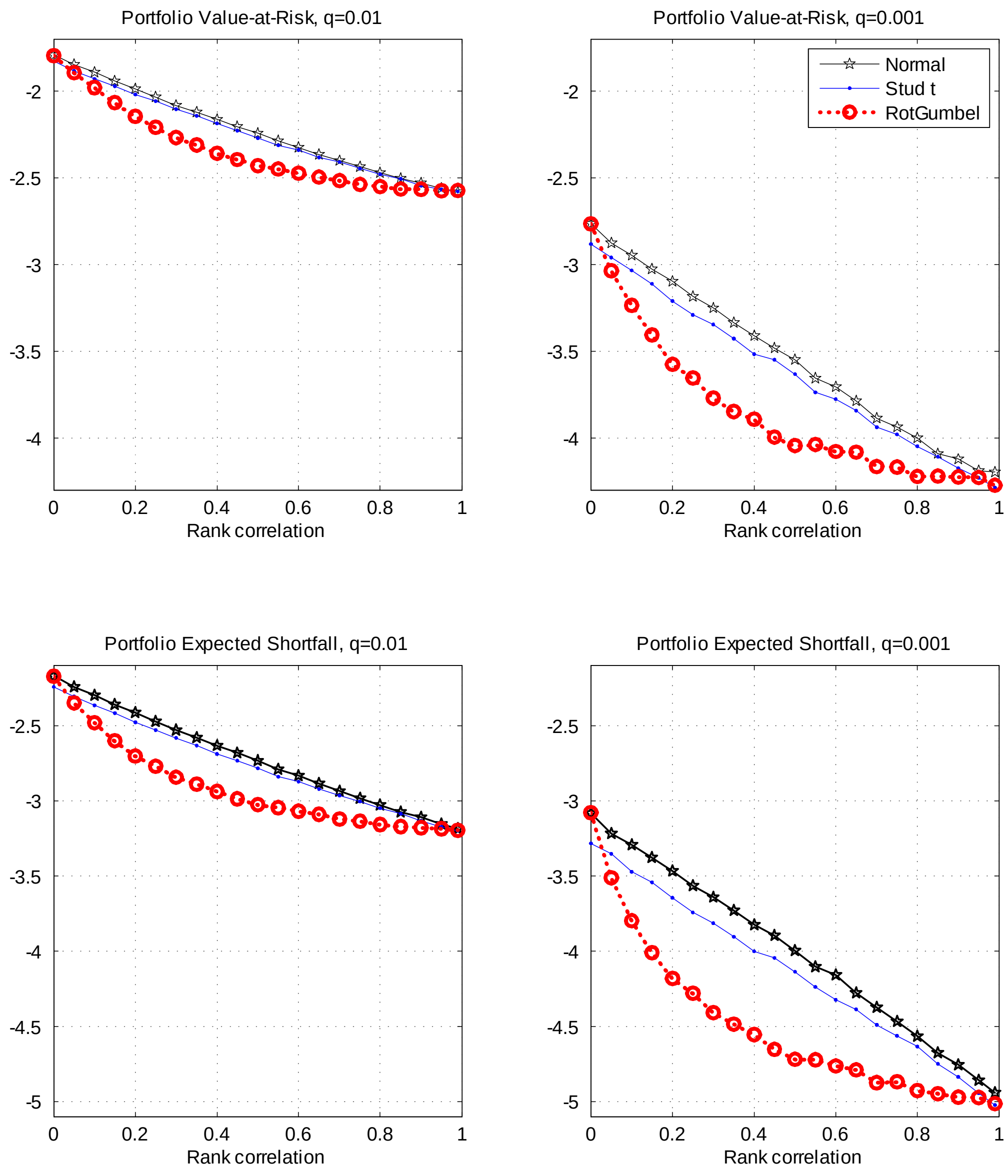


Figure 9: *Value-at-Risk (upper panels) and Expected Shortfall (lower panels), at the 1% (left panels) and 0.1% (right panels) confidence level, for an equal-weighted portfolio of two returns, with joint distribution formed from the empirical distributions of the standardized residuals of the S&P 100 and S&P 600 indices and combined with three different copulas. The rank correlation implied by these copulas is set to vary from zero to one.*