Modelling Dependence in High Dimensions with Factor Copulas*

Dong Hwan Oh and Andrew J. Patton
Duke University

31 May 2011

Preliminary. Comments welcome.

Abstract

This paper presents new models for the dependence structure, or copula, of economic variables, and asymptotic results for a new simulation-based estimator of these models. The proposed models are based on a factor structure for the copula and are particularly attractive for high dimensional applications, involving fifty or more variables. Estimation of this class of models is complicated by the lack of a closed-form likelihood, but estimation via a simulation-based method using rank statistics is simple, and we provide asymptotic results that show the consistency and asymptotic normality of such estimators. We analyze the finite-sample behavior of these estimators in an extensive simulation study. We apply the model to a group of 100 daily stock returns and find evidence of statistically significant tail dependence, and that the dependence between these assets is stronger in crashes than booms.

Keywords: correlation, dependence, copulas, tail risk.

J.E.L. codes: C31, C32, C51.

---

*We thank seminar participants at Duke University, Erasmus University Rotterdam and the Humboldt-Copenhagen Financial Econometrics workshop for helpful comments. Contact address: Andrew Patton, Department of Economics, Duke University, 213 Social Sciences Building, Box 90097, Durham NC 27708-0097. Email: andrew.patton@duke.edu.

1 Introduction

One of the many surprises from the financial crisis of late 2007 to 2008 was the extent to which assets that had previously behaved mostly independently suddenly moved together. This was particularly prominent in the financial sector, where poor models of the dependence between certain asset returns (such as those on housing, or those related to mortgage defaults) are thought to be one of the causes of the collapse of the market for CDOs and related securities, see Coval, *et al.* (2009) and Zimmer (2010) for example. Many models that were being used to capture the dependence between a large number of financial assets were revealed as being inadequate during the crisis. However, one of the difficulties in analyzing risks across many variables is the relative paucity of econometric models suitable for the task. Correlation-based models, while suitable when risk can be assumed to be summarized using the second moment, are often built on an assumption of multivariate Gaussianity, and face the risk of neglecting dependence between the variables in the tails, i.e., neglecting the possibility that arbitrarily large crashes may be correlated across assets.

This paper makes two primary contributions. First, we present new models for the dependence structure, or copula, of economic variables. The models are based on a simple factor structure for the copula and are particularly attractive for high dimensional applications, involving fifty or more variables[1]. These copula models may be combined with existing models for univariate distributions to construct flexible, tractable joint distributions for large collections of variables. The proposed copula models permit the researcher to determine the degree of flexibility/parsimony, based on the number of variables and the amount of data available. For example, by allowing for a fat-tailed common factor the model captures the possibility of correlated crashes, and by allowing the common factor to be asymmetrically distributed the model allows for the possibility that the dependence between the variables is stronger during downturns than during upturns. By allowing for multiple common factors, it is possible to capture heterogeneous pair-wise dependence within the overall multivariate copula. High dimension economic applications will often require some strong simplifying assumptions in order to keep the model tractable, and an important feature of the class of proposed models is that such assumptions can be made in an easily understandable manner.

[1] For related recent work on high dimensional conditional covariance matrix estimation, see Engle and Kelly (2007), Engle *et al.* (2008), and Hautsch *et al.* (2010).

The second contribution of this paper is a study of the asymptotic properties of a simulation-based estimator of the parameters of this model. The class of factor copulas that we present does not generally have a likelihood that is known in closed form, forcing us to consider other estimation methods. We propose a simulation-based estimation method based on rank dependence measures that shares features with the simulated method of moments (SMM), and we derive its asymptotic properties using recent work in empirical copula process theory, see Fermanian, *et al.* (2004), Chen and Fan (2006a,b) and Rémillard (2010), combined with existing results on simulation-based estimation, see Pakes and Pollard (1989) and Newey and McFadden (1994) for example.

We apply examples of our proposed factor copulas to a set of 100 daily stock returns, over the period 2008–2010. This is one of the highest dimension applications of copula theory in the econometrics literature. We find significant evidence in favor of a fat-tailed common factor for these stocks (indicative of non-zero tail dependence), implying that the Normal, or Gaussian, copula is not suitable for these assets. Moreover, we find significant evidence that the common factor is asymmetrically distributed, with crashes being more highly correlated than booms. Our empirical results suggest that risk management decisions made using the Normal copula may be based on too benign a view of these assets, and derivative securities based on baskets of these assets, or related securities such as CDOs, may be mis-priced if based on a Normal copula. The fact that large negative shocks may originate from a fat-tailed common factor, and thus affect all stocks at once, makes the diversification benefits of investing in these stocks lower than under Normality.

Certain types of factor copulas have already appeared in the literature. The models we consider are extensions of Hull and White (2004), in that we retain a simple linear, additive factor structure, but allow for the variables in the structure to have flexibly specified distributions. Other variations on factor copulas are presented in Andersen, *et al.* (2004) and van der Voort (2005), who consider certain non-linear factor structures. The papers to date, however, have not considered estimation of the unknown parameters of these copulas, instead examining simulation and pricing using these copulas. Our formal analysis of the estimation of copulas via a SMM-type procedure is new to the literature, as is our application of this class of models to a large collection of asset returns.

Some methods for modelling high dimension copulas have previously been proposed in the literature, though few consider dimensions greater than twenty[2]. The Normal copula, see Li (2000)

[2] For general reviews of copulas in economics and finance see Cherubini, *et al.* (2004), Patton (2009), and Manner and Reznikova (2010).

amongst many others, is simple to implement and to understand, but imposes the strong assumption of zero tail dependence, and symmetric dependence between booms and crashes. The (Student's) t copula, and variants of it, are discussed in Demarta and McNeil (2005). An attractive extension of the t copula, the "grouped t" copula, is proposed in Daul *et al.* (2003), who show that this copula can be used in applications of up to 100 variables. This copula allows for heterogeneous tail dependence between pairs of variables, but imposes that upper and lower tail dependence are equal (a finding we strongly reject for equity returns). Archimedean copulas such as the Clayton or Gumbel allow for tail dependence and particular forms of asymmetry, but usually have only a one or two parameters to characterize the dependence between all variables, and are thus quite restrictive when the number of variables is large. In a recent paper McNeil and Nešlehová (2010) propose "Liouville" copulas as a more flexible generalization of Archimedean copulas. Multivariate "vine" copulas are constructed by sequentially applying bivariate copulas to build up a higher dimension copula, see Aas, *et al.* (2009), Heinen and Valdesogo (2009) and Min and Czado (2010) for example. Smith, *et al.* (2010) extract the copula implied by a multivariate skew t distribution and use that to model groups of up to 15 variables.

The class of factor copulas we propose has two main advantages relative to existing models. First, many extensions of the simplest version of this model are possible, permitting the researcher great flexibility in whichever direction he/she believes is the most important for a given application. Second, the model is easily interpreted, particularly given economists' familiarity with factor models[3], and any restrictions that are required for tractability are easily understood and explained. The main drawback of this class of models is that it generally does not have a closed-form likelihood, however we show how to overcome this via simulation based estimation methods.

The remainder of the paper is structured as follows. Section 2 presents the class of factor copulas and Section 3 discusses their estimation via an SMM-type method. Section 4 presents a simulation study of the proposed new methods, and Section 5 presents an application using daily returns on individual constituents of the S&P 100 equity index over the period 2008-2010. Appendix A contains all proofs, and Appendix B contains a discussion of the dependence measures used in estimation.

[3] See, for example, the recent special issue of the *Journal of Econometrics* devoted to this topic, edited by Palm and Urbain (2011).

2 Factor copulas

For simplicity we will focus on unconditional distributions in the text below, and discuss the extension to conditional distributions in the next section. Consider a vector of N variables, $\mathbf{Y}$, with some joint distribution $\mathbf{F}$, marginal distributions F_i, and copula $\mathbf{C}$:

$$[Y_1, ..., Y_N]' \equiv \mathbf{Y} \sim \mathbf{F} = \mathbf{C}(F_1, ..., F_N) \tag{1}$$

The copula completely describes the dependence between the variables $Y_1, ..., Y_N$, and our task is to construct useful models for this dependence.

2.1 Description of a simple factor copula model

The class of copulas we consider are those that can be generated by the following simple factor structure, based on a set of $N+1$ latent variables:

$$\begin{aligned} X_i &= Z + \varepsilon_i, \;\; i = 1, 2, ..., N \\ Z &\sim F_z, \;\; \varepsilon_i \sim iid \; F_\varepsilon, \;\; Z \perp\!\!\!\perp \varepsilon_i \; \forall \; i \\ [X_1, ..., X_N]' &\equiv \mathbf{X} \sim \mathbf{F}_x = \mathbf{C}(G_1, ..., G_N) \end{aligned} \tag{2}$$

The copula of the latent variables $\mathbf{X}$ is used as the model for the copula of the observable variables $\mathbf{Y}$.[4] An important point about the above construction is that the marginal distributions of X_i may be different from those of the original variables Y_i, that is, $F_i \neq G_i$ in general. We use the structure for the vector $\mathbf{X}$ *only* for its copula, and completely discard the resulting marginal distributions. By doing so, we use equation (2) to construct a model for the copula of $\mathbf{Y}$, and leave the marginal distributions F_i to be specified and estimated in a separate step.

The copula implied by the above structure is not generally known in closed form. The leading case where it *is* known is when F_z and F_ε are both Gaussian distributions, in which case the variable $\mathbf{X}$ is multivariate Gaussian, implying a Gaussian copula, and with an equicorrelation dependence structure (with correlation between any pair of variables equal to $\sigma_z^2 / \left(\sigma_z^2 + \sigma_\varepsilon^2\right)$). For other choices of F_z and F_ε the joint distribution of $\mathbf{X}$, and more importantly the copula of $\mathbf{X}$, is not known in

[4] This method for constructing a copula model resembles the use of mixture models, e.g. the Normal-inverse Gaussian or generalized hyperbolic distributions, where the distribution of interest is obtained by considering a function of a collection of latent variables, see Barndorff-Nielsen (1978, 1997), Barndorff-Nielsen and Shephard (2009), McNeil, *et al.* (2005).

closed form. It is clear from the structure above that the copula will exhibit "equidependence", in that each pair of variables will have the same bivariate copula as any other pair. A similar assumption for correlations is made in Engle and Kelly (2007).

It is simple to simulate from F_z and F_ε for many classes of distributions, and from simulated data we can extract properties of the copula, such as rank correlation, other measures of concordance such as Kendall's tau, and upper and lower tail dependence and quantile dependence. These simulated moments can then be used in simulated method of moments (SMM) estimation of the unknown parameters. We describe how to implement this estimation procedure in Section 3. Note that by focussing on 'pure' dependence measures, see Nelsen (2006, Chapter 5) for discussion, we can extract copula information from the simulations of $\mathbf{X}$ without being affected at all by the marginal distributions (G_i) of $\mathbf{X}$.

To illustrate the flexibility of this simple class of copulas, Figure 1 presents 1000 random draws from bivariate distributions constructed using four different factor copulas. In all cases the marginal distributions, F_i, are set to $N(0,1)$, and the variance of the latent variables in the factor copula are set to $\sigma_z^2 = \sigma_\varepsilon^2 = 1$, so that the common factor (Z) accounts for one-half of the variance of each X_i. We set $F_\varepsilon = N(0,1)$, and generate the four factor copulas by considering four different choices for F_z. The first is $N(0,1)$, implying that the copula is Normal. The second sets $F_z = t(4)$, generating a symmetric copula with tail dependence. The third and fourth factor copulas use Hansen's (1994) skewed t distribution, with degrees of freedom and asymmetry parameters set to either $(\infty, -0.5)$, corresponding to a skewed Normal distribution, or to $(4, -0.5)$, corresponding to a skewed $t(4)$ distribution. When the degrees of freedom parameter is infinite (as in the Normal or skewed Normal case), Figure 1 shows that tail events tend to be uncorrelated across the two variables, while when the degrees of freedom is set to 4, we observe several draws in the joint upper and lower tails. When the skewness parameter is negative, as in the lower two panels of Figure 1, we observe stronger clustering of observations in the joint negative quadrant compared with the joint positive quadrant.

An alternative way to illustrate the differences in the dependence implied by these four models is to use a measure known as "quantile dependence". This measure captures the probability of observing a draw in the q-tail of one variable given that such an observation has been observed for the other variable. It is defined as:

$$\tau_q \equiv \begin{cases} \frac{1}{q} \Pr\left[U_1 \leq q, U_2 \leq q\right], & q \in (0, 0.5] \\ \frac{1}{1-q} \Pr\left[U_1 > q, U_2 > q\right], & q \in (0.5, 1) \end{cases} \tag{3}$$

where $U_i \equiv G_i(X_i) \sim Unif(0,1)$ are the probability integral transforms of the simulated X_i variables. As $q \to 0$ ($q \to 1$) this measure converges to lower (upper) tail dependence, and for values of q "near" zero or one we obtain an estimate of the dependence "near" the joint tails of the distribution.

Figure 2 presents the quantile dependence functions for these four copulas. For the symmetric copulas (Normal, and t-Normal factor copula) this function is symmetric about $q = 0.5$, while for the others it is not. The two copulas with a fat-tailed common factor exhibit quantile dependence that increases near the tails: in those cases an extreme observation is more likely to have come from the fat-tailed common factor (Z) than from the thin-tailed idiosyncratic variable (ε_i), and thus an extreme value for one variable makes an extreme value for the other variable more likely. The Skew t(4)-Normal factor copula illustrates the flexibility of this simple class of models, generating low upper quantile dependence but high lower quantile dependence, a feature that may be useful when modelling asset returns.

Figure 3 illustrates the differences between these copulas using a truly multivariate approach: Conditional on observing k out of 100 stocks crashing, we present the expected number of the remaining $(100 - k)$ stocks that will crash. For this illustration we define a "crash" as a realization in the lower 1/66 tail.[5] The upper panel shows that as we condition on more variables crashing, the expected number of other variables that will crash initially increases, and peaks at around $k = 30$. At that point, a Skew t(4)-Normal factor copula predicts that around another 35 variables will crash, while under the Normal copula we expect only around 12 more variables to crash. As we condition on even more variables crashing the plot converges to zero, which is a function of the fact that as we condition on having observed more crashes, there are fewer variables left to crash. The lower panel of Figure 3 shows that the expected *proportion* of remaining stocks that will crash generally increases all the way to $k = 99$.[6] For comparison, this figure also plots the

[5] This is motivated as a once-in-a-quarter event for daily asset returns. Results for once-in-a-month (1/22) and once-in-a-year (1/252) events are broadly similar.

[6] For the Normal copula this is not the case, however this is perhaps due to simulation error: even with the 10 million simulations used to obtain this figure, joint 1/66 tail crashes are so rare under a Normal copula that there is

results for a positively skewed Skew t(4)-Normal factor copula, where booms are more correlated than crashes. This figure illustrates some of the features of dependence that are unique to high dimension applications, and further motivates our proposal for a class of flexible, parsimonious models for such applications.

[INSERT FIGURES 1, 2 AND 3 ABOUT HERE]

2.2 Extensions of the model

The structure of this model immediately suggests three directions for extensions. The first is to allow for weights on the common factor that differ across variables. That is, let

$$X_i = \beta_i Z + \varepsilon_i, \;\; i = 1, 2, ..., N \tag{4}$$

with the rest of the model left unchanged. This might be called the "single factor, flexible weights" factor copula. In this case the implied copula is no longer equidependent: a given pair of variables may have weaker or stronger dependence than some other pair. This extension introduces $N - 1$ additional parameters to this model, increasing its flexibility to model heterogeneous pairs of variables, at the cost of a more difficult estimation problem. An intermediate model may be considered, in which sub-sets of variables are assumed to have the same weight on the common factor, similar in spirit to the "grouped t" copula of Daul, *et al.* (2003). This might be reasonable for financial applications with variables grouped ex ante using industry classifications, for example.

A second extension to consider is a multi-factor version of the model, where the dependence between the variables is assumed to come from a J-factor model, with possibly correlated factors:

$$\begin{aligned} X_i &= \sum_{j=1}^{J} \beta_{ij} Z_j + \varepsilon_i \\ \varepsilon_i &\sim iid\ F_\varepsilon, \;\; Z_j \perp\!\!\!\perp \varepsilon_i \ \forall\ i, j \\ [Z_1, ..., Z_J]' &\equiv \mathbf{Z} \sim \mathbf{F}_z = \mathbf{C}_z \left(G_{z_1}, ..., G_{z_J} \right) \end{aligned} \tag{5}$$

In the most general case we allow $\mathbf{Z}$ to have the copula $\mathbf{C}_Z$, to allow dependence between the common factors. An empirically useful simplification of this model is to impose that the common factors are independent, and thus remove the need to specify and estimate $\mathbf{C}_Z$. A further simplification of this factor model may be to assume that each common factor has a weight equal to one or

a fair degree of simulation error in this plot for $k \geq 80$.

zero, with the weights specified in advance by grouping variables, for example by grouping stocks by industry.

A third extension is to consider allowing for time variation in the factor copula. The factor copula models discussed above are similar in spirit to Bollerslev (1990), in that we (implicitly, so far) allow for time-varying marginal distributions, but impose that the copula is constant. The resulting multivariate density model will be time-varying but the dependence between the variables is assumed constant. In generalizing to allow for time-varying dependence, we are guided by recent work on high dimensional time-varying correlation matrices, see Engle and Kelly (2007) and Engle, *et al.* (2008), who propose methods that are feasible given the amount of data that is typically available in practice. A natural analog of the model by Engle and Kelly (2007) for our application is one where we impose a simple one-factor, unit-weight, structure, but allow the variance of the common factor, σ_z^2, to vary through time, for example:

$$\sigma_{z,t}^2 = \omega + \beta \sigma_{z,t-1}^2 + \alpha \cdot h\left(Y_{1,t-1}, ..., Y_{N,t-1}\right) \tag{6}$$

where $h\left(\cdot\right)$ is some function of the lagged values of the observable data, in the spirit of Patton (2006). In such a model, the all pair-wise dependence functions are identical at a given point in time, but these functions change through time, with higher values of σ_z^2 corresponding to times of higher dependence.

3 Simulation-based estimation of copula models

In this section we propose estimation of parametric copula models via a simulated method of moments (SMM) type approach. The method we consider is *not* strictly SMM, as the "moments" we consider are functions of rank statistics, however the results we obtain resemble those for SMM and so we use this moniker for simplicity. Our results combine well-known results for estimation involving non-smooth objective functions, see Pakes and Pollard (1989), McFadden (1989), and Newey and McFadden (1994) for example, and recent results from empirical process theory for copulas. To the best of our knowledge, simulation-based estimation of copula models has not previously been considered in the literature.

We consider the same class of data generating processes (DGPs) as Chen and Fan (2006b) and Rémillard (2010). This class allows each variable to have time-varying conditional mean and conditional variance, each governed by parametric models, with some unknown marginal distribution.

We estimate the marginal distributions using the empirical distribution function (EDF). The conditional copula of the data is assumed to belong to a parametric family, and is assumed constant; the extension to allow for time-varying conditional copulas requires a quite different asymptotic approach, and we do not pursue that here. The DGP we consider is:

$$\begin{aligned} \mathbf{Y}_t &= \boldsymbol{\mu}_t(\boldsymbol{\phi}_0) + \boldsymbol{\sigma}_t(\boldsymbol{\phi}_0)\boldsymbol{\eta}_t \\ \text{where } \mathbf{Y}_t &\equiv [Y_{1t}, \ldots, Y_{Nt}]' \\ \boldsymbol{\mu}_t(\boldsymbol{\phi}) &\equiv [\mu_{1t}(\boldsymbol{\phi}), \ldots, \mu_{Nt}(\boldsymbol{\phi})]' \\ \boldsymbol{\sigma}_t(\boldsymbol{\phi}) &\equiv diag\{\sigma_{1t}(\boldsymbol{\phi}), \ldots, \sigma_{Nt}(\boldsymbol{\phi})\} \\ \boldsymbol{\eta}_t &\equiv [\eta_{1t}, \ldots, \eta_{Nt}]' \sim iid \;\; \mathbf{F}_{\boldsymbol{\eta}} = \mathbf{C}(F_1, \ldots, F_N; \boldsymbol{\theta}_0) \end{aligned} \tag{7}$$

where $\boldsymbol{\mu}_t$ and $\boldsymbol{\sigma}_t$ are $\mathcal{F}_{t-1}$-measurable and independent of $\boldsymbol{\eta}_t$. $\mathcal{F}_{t-1}$ is the sigma-field containing information generated by $\{\mathbf{Y}_{t-1}, \mathbf{Y}_{t-2}, \ldots\}$. The $r \times 1$ vector of parameters governing the dynamics of the variables, $\boldsymbol{\phi}_0$, is assumed to be $\sqrt{T}$-consistently estimable. If $\boldsymbol{\phi}_0$ is known, or if $\boldsymbol{\mu}_t$ and $\boldsymbol{\sigma}_t$ are known constant, then the model becomes one for *iid* data. Our task is to estimate the $p \times 1$ vector of copula parameters, $\boldsymbol{\theta}_0 \in \boldsymbol{\Theta}$, based on the standardized residual $\left\{\hat{\boldsymbol{\eta}}_t \equiv \boldsymbol{\sigma}_t^{-1}\left(\hat{\phi}\right)\left[\mathbf{Y}_t - \boldsymbol{\mu}_t\left(\hat{\phi}\right)\right]\right\}_{t=1}^{T}$ and simulations from the copula model (for example, the factor copula model in equation 2).

3.1 Definition of the SMM estimator

We will consider simulation from some parametric joint distribution, $\mathbf{F}_x(\boldsymbol{\theta})$, with marginal distributions $G_i(\boldsymbol{\theta})$, and copula $\mathbf{C}(\boldsymbol{\theta})$. The case where it is possible to simulate directly from $\mathbf{C}(\boldsymbol{\theta})$ is nested in this scenario by allowing G_i to be the $Unif(0,1)$ cdf.[7] We use only "pure" dependence measures as moments since those are affected not by changes in the marginal distributions of simulated data ($\mathbf{X}$). For example, moments like means and variances, are pure functions of the marginal distributions (G_i) and thus contain no information on the copula. Dependence measures like rank correlation, Kendall's tau, and quantile dependence are pure functions of the copula and are unaffected by the marginal distributions, see Nelsen (2006) for example. Spearman's rank correlation

[7] In that case there would be no need to consider the estimation of G_i using the EDF, and some of the steps below would simplify. The EDF is still required for the standardized residuals and thus this case still requires the consideration of rank statistics. We focus on the more general case that it is only possible to simulate from $\mathbf{F}_x$ as that is the case we face when estimating factor copula models of the form proposed in Section 2.

and quantile dependence for the pair $\left(\eta_{it},\eta_{jt}\right)$ are defined as:

$$\rho^{ij} \equiv 12E\left[F_i\left(\eta_i\right)F_j\left(\eta_j\right)\right]-3=12\int\int uvdC_{ij}\left(u,v\right)-3 \tag{8}$$

$$\tau_q^{ij} \equiv \begin{cases} \frac{P\left(F_i(\eta_i)\le q, F_j\left(\eta_j\right)\le q\right)}{q}=\frac{C_{ij}(q,q)}{q}, & q\in(0,0.5] \\ \frac{P\left(F_i(\eta_i)> q, F_j\left(\eta_j\right)> q\right)}{1-q}=\frac{1-2q+C_{ij}(q,q)}{1-q}, & q\in(0.5,1)\end{cases} \tag{9}$$

where C_{ij} is the copula of $\left(\eta_i,\eta_j\right)$. The sample counterparts based on the estimated standardized residuals are defined as:

$$\hat{\rho}^{ij} \equiv \frac{12}{T}\sum_{t=1}^{T}\hat{F}_i\left(\hat{\eta}_{it}\right)\hat{F}_j\left(\hat{\eta}_{jt}\right)-3 \tag{10}$$

$$\hat{\tau}_q^{ij} \equiv \begin{cases} \frac{1}{Tq}\sum_{t=1}^{T}\mathbf{1}\left\{\hat{F}_i\left(\hat{\eta}_{it}\right)\le q, \hat{F}_j\left(\hat{\eta}_{jt}\right)\le q\right\}, & q\in(0,0.5] \\ \frac{1}{T(1-q)}\sum_{t=1}^{T}\mathbf{1}\left\{\hat{F}_i\left(\hat{\eta}_{it}\right)> q, \hat{F}_j\left(\hat{\eta}_{jt}\right)> q\right\}, & q\in(0.5,1)\end{cases} \tag{11}$$

$$\text{where } \hat{F}_i\left(y\right) \equiv \frac{1}{T+1}\sum_{t=1}^{T}1\left\{\hat{\eta}_{it}\le y\right\} \tag{12}$$

We will denote the counterparts based on simulated data as $\tilde{\rho}^{ij}\left(\boldsymbol{\theta}\right)$ and $\tilde{\tau}_q^{ij}\left(\boldsymbol{\theta}\right)$.

Let $\tilde{\mathbf{m}}_S\left(\boldsymbol{\theta}\right)$ be a $(m\times 1)$ vector of dependence measures computed using S simulations from $\mathbf{F}_x\left(\boldsymbol{\theta}\right)$, $\left\{\mathbf{X}_s\right\}_{s=1}^{S}$, and let $\hat{\mathbf{m}}_T$ be the corresponding vector of dependence measures computed using the the standardized residuals $\{\hat{\boldsymbol{\eta}}_t\}_{t=1}^{T}$. These vectors can also contain linear combinations of dependence measures, a feature that is useful when considering estimation of high-dimension models.[8] Define the difference between these as

$$\mathbf{g}_{T,S}\left(\boldsymbol{\theta}\right)\equiv\hat{\mathbf{m}}_T-\tilde{\mathbf{m}}_S\left(\boldsymbol{\theta}\right) \tag{13}$$

Our SMM estimator is based on searching across $\boldsymbol{\theta}\in\boldsymbol{\Theta}$ to make this difference as small as possible. The estimator is defined as:

$$\hat{\boldsymbol{\theta}}_{T,S} \equiv \underset{\boldsymbol{\theta}\in\boldsymbol{\Theta}}{\arg\min}\; Q_{T,S}\left(\boldsymbol{\theta}\right) \tag{14}$$

$$\text{where } Q_{T,S}\left(\boldsymbol{\theta}\right) \equiv \mathbf{g}_{T,S}'\left(\boldsymbol{\theta}\right)\hat{W}_T\mathbf{g}_{T,S}\left(\boldsymbol{\theta}\right)$$

and $\hat{W}_T$ is some positive definite weight matrix, which may depend on the data. As usual, for identification we require at least as many moment conditions as there are free parameters (i.e.,

[8] For example, in our empirical application with an equi-dependence copula model, we use the average of all pair-wise rank correlation coefficients, and the average of all pair-wise quantile dependence coefficients with $q\in\{0.05,0.10,0.90,0.95\}$.

$m \geq p$). In the subsections below we establish the consistency and asymptotic normality of this estimator, provide a consistent estimator of its asymptotic covariance matrix, and obtain a test based on over-identifying restrictions.

3.2 Consistency of the SMM estimator

The estimation problem here differs in two important ways from standard GMM or M-estimation: Firstly, the objective function, $Q_{T,S}(\boldsymbol{\theta})$ is not continuous in $\boldsymbol{\theta}$ as we use simulations to obtain $\tilde{\mathbf{m}}_S(\boldsymbol{\theta})$. This problem would vanish if, for the copula model being considered, we knew the mapping from $\boldsymbol{\theta}$ to the dependence measure(s) in closed form. The second difference is that a law of large numbers is not available to show the pointwise convergence of $\mathbf{g}_{T,S}(\boldsymbol{\theta})$, as the functions $\hat{\mathbf{m}}_T$ and $\tilde{\mathbf{m}}_S(\boldsymbol{\theta})$ both involve empirical distribution functions. We use recent developments in empirical process theory to overcome this difficulty.

We now list some assumptions that are required for our results to hold.

Assumption 1

(i) The distributions $\mathbf{F}_\eta$ and $\mathbf{F}_x$ are continuous.

(ii) Every bivariate marginal copula C_{ij} of $\mathbf{C}$ has continuous partial derivatives with respect to u_i and u_j.

If the data $\mathbf{Y}_t$ are *iid*, e.g. if $\boldsymbol{\mu}_t$ and $\boldsymbol{\sigma}_t$ are known constant in (7) or if $\boldsymbol{\phi}_0$ is known, then Assumption 1 is sufficient to prove Proposition 1 below, but if standardized residuals are used in the estimation of the copula then more assumptions are necessary in order to control the estimation error coming from the models for the conditional means and conditional variances. We combine assumptions A1–A6 in Rémillard (2010) in the following assumption. Firstly, define $\boldsymbol{\gamma}_{0t} = \boldsymbol{\sigma}_t^{-1}\left(\hat{\boldsymbol{\phi}}\right)\dot{\boldsymbol{\mu}}_t\left(\hat{\boldsymbol{\phi}}\right)$ and $\boldsymbol{\gamma}_{1kt} = \boldsymbol{\sigma}_t^{-1}\left(\hat{\boldsymbol{\phi}}\right)\dot{\boldsymbol{\sigma}}_{kt}\left(\hat{\boldsymbol{\phi}}\right)$ where $\dot{\boldsymbol{\mu}}_t(\boldsymbol{\phi}) = \frac{\partial \boldsymbol{\mu}_t(\boldsymbol{\phi})}{\partial \boldsymbol{\phi}'}, \dot{\boldsymbol{\sigma}}_{kt}(\boldsymbol{\phi}) = \frac{\partial [\boldsymbol{\sigma}_t(\boldsymbol{\phi})]_{\text{k-th column}}}{\partial \boldsymbol{\phi}'}$, $k = 1, \ldots, N$. Define $\mathbf{d}_t$ as

$$\mathbf{d}_t = \boldsymbol{\eta}_t - \hat{\boldsymbol{\eta}}_t - \left(\gamma_{0t} + \sum_{k=1}^{N} \eta_{kt}\gamma_{1kt}\right)\left(\hat{\boldsymbol{\phi}} - \boldsymbol{\phi}_0\right)$$

where η_{kt} is k-th row of $\boldsymbol{\eta}_t$ and both $\boldsymbol{\gamma}_{0t}$ and $\boldsymbol{\gamma}_{1kt}$ are $\mathcal{F}_{t-1}$-measurable.

Assumption 2

(i) $\frac{1}{T}\sum_{t=1}^{T}\boldsymbol{\gamma}_{0t}\overset{p}{\rightarrow}\Gamma_0$ *and* $\frac{1}{T}\sum_{t=1}^{T}\boldsymbol{\gamma}_{1kt}\overset{p}{\rightarrow}\Gamma_{1k}$ *where* Γ_0 *and* Γ_{1k} *are deterministic for* $k=1,\ldots,N$.

(ii) $\frac{1}{T}\sum_{t=1}^{T}E\left(\|\boldsymbol{\gamma}_{0t}\|\right), \frac{1}{T}\sum_{t=1}^{T}E\left(\|\boldsymbol{\gamma}_{0t}\|^2\right), \frac{1}{T}\sum_{t=1}^{T}E\left(\|\boldsymbol{\gamma}_{1kt}\|\right),$ *and* $\frac{1}{T}\sum_{t=1}^{T}E\left(\|\boldsymbol{\gamma}_{1kt}\|^2\right)$ *are bounded for* $k=1,\ldots,N$.

(iii) There exists a sequence of positive terms $r_t>0$ *so that* $\sum_{t\geq 1}r_t<\infty$ *and such that the sequence* $\max_{1\leq t\leq T}\|\mathbf{d}_t\|/r_t$ *is tight.*

(iv) $\max_{1\leq t\leq T}\|\boldsymbol{\gamma}_{0t}\|/\sqrt{T}=o_p(1)$ *and* $\max_{1\leq t\leq T}\eta_{kt}\|\boldsymbol{\gamma}_{1kt}\|/\sqrt{T}=o_p(1)$ *for* $k=1,\ldots,N$.

(v) $\left(\alpha_T,\sqrt{T}\left(\hat{\boldsymbol{\phi}}-\boldsymbol{\phi}_0\right)\right)$ *weakly converges to a continuous Gaussian process in* $[0,1]^N\times\mathbb{R}^r$, *where* α_T *is the empirical copula process of uniform random variables:*

$$\alpha_T=\frac{1}{\sqrt{T}}\sum_{t=1}^{T}\left\{\prod_{k=1}^{N}1\left(U_{kt}\leq u_k\right)-C\left(\mathbf{u}\right)\right\}$$

(vi) $\frac{\partial\mathbf{F}_{\boldsymbol{\eta}}}{\partial\eta_k}$ *and* $\eta_k\frac{\partial\mathbf{F}_{\boldsymbol{\eta}}}{\partial\eta_k}$ *are bounded and continuous on* $\bar{\mathbb{R}}^N=[-\infty,+\infty]^N$ *for* $k=1,\ldots,N$.

With these two assumptions we can show that sample rank dependence and quantile dependence converge in probability to their population counterparts, see Lemma 1 in Appendix A. When applied to simulated data this convergence holds pointwise for any $\boldsymbol{\theta}$. Thus $\mathbf{g}_{T,S}(\boldsymbol{\theta})$ converges in probability to the population moment functions defined as follows:

$$\mathbf{g}_{T,S}(\boldsymbol{\theta})\equiv\hat{\mathbf{m}}_T-\tilde{\mathbf{m}}_S(\boldsymbol{\theta})\overset{p}{\rightarrow}\mathbf{g}_0(\boldsymbol{\theta})\equiv\mathbf{m}(\boldsymbol{\theta}_0)-\mathbf{m}(\boldsymbol{\theta}),\ \text{for }\forall\boldsymbol{\theta}\in\boldsymbol{\Theta}\text{ as }T,S\rightarrow\infty \tag{15}$$

We define the population objective function as

$$Q_0(\boldsymbol{\theta})=\mathbf{g}_0(\boldsymbol{\theta})'W_0\mathbf{g}_0(\boldsymbol{\theta}) \tag{16}$$

where W_0 is the probability limit of $\hat{W}_T$. The convergence of $\mathbf{g}_{T,S}(\boldsymbol{\theta})$ and $\hat{W}_T$ implies that

$$Q_{T,S}(\boldsymbol{\theta})\overset{p}{\rightarrow}Q_0(\boldsymbol{\theta})\quad\text{for }\ \forall\boldsymbol{\theta}\in\boldsymbol{\Theta}\text{ as }T,S\rightarrow\infty$$

For consistency of our estimator we need, as usual, uniform convergence of $Q_{T,S}(\boldsymbol{\theta})$, but as this function is not continuous in $\boldsymbol{\theta}$ and a law of large numbers is not available, the standard approach based on an uniform law of large numbers is not available. We instead use results on the stochastic equicontinuity of $\mathbf{g}_{T,S}(\boldsymbol{\theta})$, based on Andrews (1994) and Newey and McFadden (1994).

Assumption 3

(i) $g_0(\boldsymbol{\theta}) \neq 0$ *for* $\boldsymbol{\theta} \neq \boldsymbol{\theta}_0$

(ii) $\boldsymbol{\Theta}$ *is compact.*

(iii) Every bivariate marginal copula C_{ij} *of* $\mathbf{C}$ *is continuous in* $\boldsymbol{\theta}$*.*

(iv) $\exists\ \alpha > 0$ *and* $\hat{B}_S = O_p(1)$ *such that for all* $\boldsymbol{\theta}_1, \boldsymbol{\theta}_2 \in \boldsymbol{\Theta}$,

$$\|\mathbf{g}_{T,S}(\boldsymbol{\theta}_1) - \mathbf{g}_{T,S}(\boldsymbol{\theta}_2)\| = \|\tilde{\mathbf{m}}_S(\boldsymbol{\theta}_1) - \tilde{\mathbf{m}}_S(\boldsymbol{\theta}_2)\| \leq \hat{B}_S \|\boldsymbol{\theta}_1 - \boldsymbol{\theta}_2\|^{\alpha}.$$

Assumption 3(iv) is called a global "in probability" Lipschitz condition in Newey (1991), and is a sufficient condition for stochastic equicontinuity of $\mathbf{g}_{T,S}(\boldsymbol{\theta})$. This condition is difficult to check in our framework, as $\tilde{\mathbf{m}}_S(\boldsymbol{\theta})$ depends only implicitly on $\boldsymbol{\theta}$, through the simulated data on which the sample moments are computed, and so we leave this as a high level assumption.

Proposition 1 *Suppose that Assumptions 1, 2 and 3 hold. Then* $\hat{\boldsymbol{\theta}}_{T,S} \xrightarrow{p} \boldsymbol{\theta}_0$ *as* $T, S \to \infty$

All proofs are presented in Appendix A. A key difference between this result and the corresponding result for a standard SMM application is that we require $S \to \infty$ as $T \to \infty$ for consistency. In standard SMM applications, see Gouriéroux and Monfort (1996a) for example, consistency is often obtained with finite S. The presence of EDFs inside our "moment" conditions compels us to require $S \to \infty$ for consistency.

3.3 Asymptotic Normality of the SMM estimator

As $Q_{T,S}(\boldsymbol{\theta})$ is non-differentiable the standard approach based on a Taylor expansion is not available, however the asymptotic normality of our estimator can still be established with some further assumptions:

Assumption 4

(i) $\boldsymbol{\theta}_0$ *is an interior point of* $\boldsymbol{\Theta}$

(ii) $\hat{W}_T$ *is* $O_p(1)$ *and converges in probability to* W_0*, a positive definite matrix.*

(iii) $\mathbf{g}_0(\boldsymbol{\theta})$ *is differentiable at* $\boldsymbol{\theta}_0$ *with derivative* G_0 *such that* $G_0' W_0 G_0$ *is nonsingular.*

(iv) $\mathbf{g}_{T,S}\left(\hat{\boldsymbol{\theta}}_{T,S}\right)'\hat{W}_T\mathbf{g}_{T,S}\left(\hat{\boldsymbol{\theta}}_{T,S}\right) \leq \inf_{\boldsymbol{\theta}\in\boldsymbol{\Theta}} \mathbf{g}_{T,S}\left(\boldsymbol{\theta}\right)'\hat{W}_T\mathbf{g}_{T,S}\left(\boldsymbol{\theta}\right) + o_p\left(T^{-1}\right)$

The first two assumptions above are standard, and the fourth assumption is standard in simulation-based estimation problems, see Newey and McFadden (1994) for example. The third assumption requires the population objective function, $\mathbf{g}_0$, to be differentiable even though its finite-sample counterpart is not, which is common in simulation-based estimation. The nonsingularity of $G_0'W_0G_0$ is sufficient for local identification of the parameters of this model at $\boldsymbol{\theta}_0$, see Hall (2005) and Rothenberg (1971). With these assumptions in hand we obtain the following result.

Proposition 2 *Suppose that Assumptions 1, 2, 3 and 4 hold. Then if* $T/S \to 0$ *as* $T, S \to \infty$,

$$\sqrt{T}\left(\hat{\boldsymbol{\theta}}_{T,S} - \boldsymbol{\theta}_0\right) \xrightarrow{d} N\left(0, \Omega_0\right) \text{ as } T, S \to \infty \tag{17}$$

where $\Omega_0 = \left(G_0'W_0G_0\right)^{-1} G_0'W_0\Sigma_0W_0G_0\left(G_0'W_0G_0\right)^{-1}$, *and* $\Sigma_0 \equiv avar\left[\hat{\mathbf{m}}_T\right]$.

The asymptotic variance of SMM estimator $\hat{\boldsymbol{\theta}}_{T,S}$ has the same "sandwich" form as that of the usual GMM estimator, and as usual it simplifies to $\Omega_0 = \left(G_0'\Sigma_0^{-1}G_0\right)^{-1}$ if W_0 is the efficient weight matrix, Σ_0^{-1}. Chen and Fan (2006b) and Rémillard (2010) show that estimation error from $\hat{\boldsymbol{\phi}}$ does not enter the asymptotic distribution of the copula parameter estimator for maximum likelihood or (analytical) moment-based estimators, and the above proposition shows that this also holds for SMM-type estimators proposed here.

The proof of the above proposition uses recent results for empirical copula processes presented in Fermanian, *et al.* (2004) and Rémillard (2010) to establish the asymptotic normality of the sample dependence measures, $\hat{\mathbf{m}}_T$, and requires us to establish the stochastic equicontinuity of the moment functions, $\mathbf{v}_{T,S}\left(\boldsymbol{\theta}\right) = \sqrt{T}\left[\mathbf{g}_{T,S}\left(\boldsymbol{\theta}\right) - \mathbf{g}_0\left(\boldsymbol{\theta}\right)\right]$. These are shown in Lemmas 5 and 6 in Appendix A. Establishing the stochastic equicontinuity of $\mathbf{v}_{T,S}\left(\boldsymbol{\theta}\right)$ imposes the constraint that the number of simulations grows faster than the sample size, i.e., $T/S \to 0$. This is in contrast with standard SMM applications, see Gouriéroux and Monfort (1996a) for example, where asymptotic normality is obtained with finite S (at a cost of inflated asymptotic variance), and efficiency is obtained if $\sqrt{T}/S \to 0$ as $T, S \to \infty$.

3.4 Consistent estimation of the asymptotic variance

The asymptotic variance of our estimator has the same form as in standard GMM applications, however the components Σ_0 and G_0 require more care in their estimation than in standard applications. We suggest using an *iid* bootstrap to estimate Σ_0 :

1. Estimate parameters $\boldsymbol{\phi}$ for conditional mean and conditional variance using the given sample $\{\mathbf{Y}_t\}_{t=1}^T$, and obtain the standardized residuals $\{\hat{\boldsymbol{\eta}}_t\}_{t=1}^T$.

2. Using $\{\hat{\boldsymbol{\eta}}_t\}_{t=1}^T$, compute the sample moments and denote as $\hat{\mathbf{m}}_T$.

3. Sample with replacement from the standardized residuals $\{\hat{\boldsymbol{\eta}}_t\}_{t=1}^T$ to obtain a bootstrap sample, $\left\{\hat{\boldsymbol{\eta}}_t^{(b)}\right\}_{t=1}^T$. Repeat this step B times.

4. Using $\left\{\hat{\boldsymbol{\eta}}_t^{(b)}\right\}_{t=1}^T$, $b = 1, ..., B$, compute the sample moments and denote as $\hat{\mathbf{m}}_T^{(b)}$, $b = 1, ..., B$.

5. Calculate the sample covariance matrix of $\hat{\mathbf{m}}_T^{(b)}$ across the bootstrap replications, and scale it by the sample size:

$$\hat{\Sigma}_{T,B} = \frac{T}{B}\sum_{b=1}^{B}\left(\hat{\mathbf{m}}_T^{(b)} - \hat{\mathbf{m}}_T\right)\left(\hat{\mathbf{m}}_T^{(b)} - \hat{\mathbf{m}}_T\right)' \tag{18}$$

For the estimation G_0, we suggest a numerical derivative of $\mathbf{g}_{T,S}(\boldsymbol{\theta})$ at $\hat{\boldsymbol{\theta}}_{T,S}$, however the fact that $\mathbf{g}_{T,S}$ is non-differentiable means that care is needed in choosing the step size for the numerical derivative. In particular, as in Newey and McFadden (1994), Proposition 3 below shows that we need to let the step size go to zero, as usual, but *slower* than $1/\sqrt{T}$. Let $\mathbf{e}_k$ denote the k-th unit vector whose dimension is the same as that of $\boldsymbol{\theta}$, and let ε_T denote the step size. A two-sided numerical derivative estimator $\hat{G}_{T,S}$ of G has k-th column

$$\hat{G}_{T,S,k} = \frac{\mathbf{g}_{T,S}\left(\hat{\boldsymbol{\theta}}_{T,S}+\mathbf{e}_k\varepsilon_T\right) - \mathbf{g}_{T,S}\left(\hat{\boldsymbol{\theta}}_{T,S}-\mathbf{e}_k\varepsilon_T\right)}{2\varepsilon_T},\ \ k = 1, 2, ..., p \tag{19}$$

Combine these two estimators with $\hat{W}_T$ to form the estimator:

$$\hat{\Omega}_{T,S,B} = \left(\hat{G}'_{T,S}\hat{W}_T\hat{G}_{T,S}\right)^{-1}\hat{G}'_{T,S}\hat{W}_T\hat{\Sigma}_{T,B}\hat{W}_T\hat{G}_{T,S}\left(\hat{G}'_{T,S}\hat{W}_T\hat{G}_{T,S}\right)^{-1} \tag{20}$$

Proposition 3 *Suppose that all assumptions of Proposition 2 are satisfied, and that* $\varepsilon_T \to 0$, $\varepsilon_T\sqrt{T} \to \infty$, $B \to \infty$ *as* $T \to \infty$. *Then* $\hat{\Sigma}_{T,B} \xrightarrow{p} \Sigma_0$, $\hat{G}_{T,S} \xrightarrow{p} G_0$ *and* $\hat{\Omega}_{T,S,B} \xrightarrow{p} \Omega_0$.

3.5 A test of overidentifying restrictions

If the number of moments used in estimation is greater than the number copula parameters, then, as in standard GMM and SMM applications, it is possible to conduct a simple test of the over-identifying restrictions. When the efficient weight matrix is used in estimation, the asymptotic distribution of this test statistic is the usual chi-squared, however the method of proof is different as we again need to deal with the lack of differentiability of the objective function. We also consider the distribution of this test statistic for general weight matrices.

Proposition 4 *Suppose that all assumptions of Proposition 2 are satisfied and that the number of moments* (m) *is greater than the number of copula parameters* (p). *Then*

$$\begin{aligned} J_{T,S} &\equiv T\mathbf{g}_{T,S}\left(\hat{\boldsymbol{\theta}}_{T,S}\right)'\hat{W}_T\mathbf{g}_{T,S}\left(\hat{\boldsymbol{\theta}}_{T,S}\right)\xrightarrow{d}\mathbf{u}'A_0'A_0\mathbf{u} \\ \textit{where } \mathbf{u} &\sim N(0,I) \end{aligned}$$

and $A_0 \equiv W_0^{1/2}\Sigma_0^{1/2}R_0$, $R_0 \equiv I - \Sigma_0^{-1/2}G_0\left(G_0'W_0G_0\right)^{-1}G_0'W_0\Sigma_0^{1/2}$. *If* $\hat{W}_T = \hat{\Sigma}_{T,B}^{-1}$, *then* $J_{T,S} \xrightarrow{d} \chi^2_{m-p}$ *as usual.*

As in standard applications, the above test statistic simplifies to $TQ_{T,S}\left(\hat{\boldsymbol{\theta}}_{T,S}\right)$ if the efficient weight matrix $(\hat{\Sigma}_{T,B}^{-1})$ is used. When any other weight matrix is used, the $J_{T,S}$ statistic has a sample-specific limiting distribution, and critical values in such cases can be obtained via a simple simulation:

1. Compute $\hat{R}$ using $\hat{G}_{T,S}$, $\hat{W}_T$, and $\hat{\Sigma}_{T,B}$.
2. Simulate $\mathbf{u}^{(k)} \sim iid\ N(0,I)$, for $k = 1,2,...,K$, where K is large.
3. For each simulation, compute $J_T^{(k)} = \mathbf{u}^{(k)\prime}\hat{R}'\Sigma_{T,B}^{1/2\prime}\hat{W}_T\Sigma_{T,B}^{1/2}\hat{R}\mathbf{u}^{(k)}$
4. The sample $(1-\alpha)$ quantile of $\left\{J_T^{(k)}\right\}_{k=1}^{K}$ is the critical value for this test statistic.

The need for simulations to obtain critical values from the limiting distribution is non-standard but is not uncommon; this arises in many other testing problems, see Wolak (1989), White (2000), Andrews (2001) for examples. Given that $\mathbf{u}^{(k)}$ is a simple standard Normal, and that no optimization is required in this simulation, and that the matrix $\hat{R}$ need only be computed once, obtaining critical values for this test is simple and fast.

4 Simulation study

In this section we present a study of the finite sample properties of the simulated method of moments (SMM) estimator of the parameters of various factor copulas. In the one case where a likelihood for the copula model is available in closed form we contrast the properties of the SMM estimator with those of the maximum likelihood estimator. We initially consider three different factor copulas, all of them of the form:

$$\begin{aligned} X_i &= Z + \varepsilon_i, \quad i = 1, 2, ..., N \\ Z &\sim \textit{Skew } t\left(0, \sigma_z^2, \nu_z, \lambda_z\right) \\ \varepsilon_i &\sim iid\ N(0,1), \text{ and } \varepsilon_i \perp\!\!\!\perp Z \ \forall\ i \\ [X_1, ..., X_N]' &\sim \mathbf{F}_x = \mathbf{C}(G_x, ..., G_x) \end{aligned} \tag{21}$$

For the simulation study we set $\sigma_z^2 = 1$, implying that the common factor (Z) accounts for one-half of the variance of each X_i. The "skewed t" distribution we use is that of Hansen (1994). In the first case we set $\nu_z \to \infty$ and $\lambda_z = 0$, which implies that the resulting factor copula is simply the Gaussian copula, with equicorrelation parameter $\rho = 0.5$. In this case we can estimate the model by SMM and also by GMM and MLE, and we use this case to study the loss of efficiency in moving from MLE to GMM to SMM. In the second case we set $\lambda_z = 0$ and $\nu_z = 4$, yielding a symmetric factor copula that generates tail dependence. In the third case we set $\lambda_z = -0.5$ and $\nu_z = 4$ yielding a factor copula that generates tail dependence as well as "asymmetric dependence", in that the lower tails of the copula are more dependent than the upper tails. We estimate the inverse degrees of freedom parameter, ν_z^{-1}, so that its parameter space is $[0, 0.5)$ rather than $(2, \infty]$.

As an extension of this initial specification, we also consider a model where we allow for deviations from equidependence by letting each X_i have a different coefficient on Z:

$$X_i = \beta_i Z + \varepsilon_i \tag{22}$$

and for identification in this case we set $\sigma_z^2 = 1$. For $N = 3$ we set $[\beta_1, \beta_2, \beta_3] = [0.5, 1, 1.5]$. For $N = 10$ we set $[\beta_1, \beta_2, ..., \beta_{10}] = [0.25, 0.50, ..., 2.5]$, which corresponds to pair-wise rank correlations ranging from approximately 0.1 to 0.8. Motivated by our empirical application below, for the $N = 100$ case we consider a "block equidependence" model, where we assume that the 100 variables can be grouped *ex ante* into 10 groups, and that all variables within each group have the same β_i. We use the same set of values for β_i as in the $N = 10$ case.

We consider two different scenarios for the marginal distributions of the variables of interest. In the first case we assume that the data are *iid* with standard Normal marginal distributions, meaning that the only parameters that need to be estimated are those of the factor copula. This simplified case is contrasted with a second scenario where the marginal distributions of the variables are assumed to follow an AR(1)-GARCH(1,1) process:

$$\begin{aligned} Y_{it} &= \phi_0 + \phi_1 Y_{i,t-1} + \sigma_{it}\eta_{it}, \quad t = 1, 2, ..., T \\ \sigma^2_{it} &= \omega + \beta\sigma^2_{i,t-1} + \alpha\sigma^2_{i,t-1}\eta^2_{i,t-1} \\ \boldsymbol{\eta}_t &\equiv [\eta_{1t}, ..., \eta_{Nt}] \sim iid \;\; \mathbf{F}_\eta = \mathbf{C}(\Phi, \Phi, ..., \Phi) \end{aligned} \tag{23}$$

where Φ is the standard Normal distribution function and $\mathbf{C}$ is the factor copula implied by equation (21). We set the parameters of the marginal distributions as $[\phi_0, \phi_1, \omega, \beta, \alpha] = [0.01, 0.05, 0.05, 0.85, 0.10]$, which broadly matches the values of these parameters when estimated using daily equity return data. In this scenario the parameters of the marginal distribution are estimated in a separate first stage, following which the estimated standardized residuals are obtained:

$$\hat{\eta}_{it} = \frac{Y_{it} - \hat{\phi}_0 - \hat{\phi}_1 Y_{i,t-1}}{\hat{\sigma}_{it}}. \tag{24}$$

These residuals are used in a second stage to estimate the factor copula parameters. In all cases we consider a time series of length $T = 1000$, corresponding to approximately 4 years of daily return data, and we use $S = 25 \times T$ simulations in the computation of the dependence measures to be matched in the SMM optimization. We repeat each scenario 100 times. In all results below we use the identity weight matrix for estimation; corresponding results based on the efficient weight matrix are available in an online appendix to this paper[9]. In Appendix B we describe the dependence measures we use for the estimation of these models.

Table 1 reveals that for all three dimensions ($N = 3$, 10 and 100) and for all three copula models the estimated parameters are centered on the true values, with the average estimated bias being small relative to the standard deviation, and with the median of the simulated distribution centered on the true values. The measures of estimator accuracy (the standard deviation and the 90-10 percentile difference) reveal that adding more parameters to the model, *ceteris paribus*, leads to greater estimation error; the σ^2_z parameter, for example, is more accurately estimated when it is

[9] The results based on the efficient weight matrix are generally comparable to those based on the identity weight matrix, however the coverage rates are worse than those based on the identity weight matrix.

the only unknown parameter compared with when it is one of three unknown parameters. Looking across the dimension size, we see that the copula model parameters are almost always more precisely estimated as the dimension grows. This is intuitive, given the "equidependence" nature of all three models: increasing the dimension of the model does not increase the number of parameters to be estimated but it does increase the amount of information available on the unknown parameters.

Comparing the SMM estimator with the ML estimator, which is only feasible for the Normal copula (as the other two factor copulas do not have a copula likelihood in closed form) we see that the SMM estimator performs quite well. As predicted by theory, the ML estimator is always more efficient than the SMM estimator, however the loss in efficiency is moderate, ranging from around 25% for $N = 3$ to around 10% for $N = 100$. This provides some confidence that our move to SMM, prompted by the lack of a closed-form likelihood, does not come at a cost of a large loss in efficiency. Comparing the SMM estimator to the GMM estimator provides us with a measure of the loss in accuracy of from having to estimate the population moment function via simulation. We find that this loss is at most 3% and in some cases ($N = 100$) is slightly negative. Thus little is lost from using SMM rather than GMM.

The simulation results in Table 2, where the copula parameters are estimated after the estimation of AR(1)-GARCH(1,1) models for the marginal distributions in a separate first stage, are very similar to the case when no marginal distribution parameters are required to be estimated, consistent with Proposition 2. Thus this somewhat surprising asymptotic result is also relevant in finite samples.

Table 3 shows results for the block equidependence model for the $N = 100$ case with AR-GARCH marginal distributions,[10] which can be compared to the results in the lower panel of Table 2. This table shows that the parameters of these models are well estimated using the proposed dependence measures described in Appendix B. The accuracy of the "shape" parameters, ν_z^{-1} and λ_z, is slightly lower in the more general model, consistent with the estimation error from having to estimate ten factor loadings (β_i) being greater than from having to estimate just a single other parameter $\left(\sigma_z^2\right)$, however this loss is not great.

[INSERT TABLES 1, 2 AND 3 ABOUT HERE]

In Tables 4 and 5 we present the finite-sample coverage probabilities of 95% confidence inter-

[10] The results for *iid* data and $N = 100$, and the results for $N = 3$ and 10, are available in the web appendix.

vals based on the asymptotic Normality result from Proposition 2 and the asymptotic covariance matrix estimator presented in Proposition 3. As shown in that proposition, a critical input to the asymptotic covariance matrix estimator is the step size used in computing the numerical derivative matrix $\hat{G}_{T,S}$. This step size, ε_T, must go to zero, but at a slower rate than $\sqrt{T}$. Ignoring constants, our simulation sample size of $T = 1000$ suggests setting $\varepsilon_T > 0.001$, which is much larger than standard step sizes used in computing numerical derivatives.[11] We study the impact of the choice of step size by considering a range of values from 0.0001 to 0.1. Table 4 shows that when the step size is set to 0.01, 0.03 or 0.1 the finite-sample coverage rates are close to their nominal levels. However if the step size is chosen too small (0.003 or smaller) then the coverage rates are much lower than nominal levels. For example, setting $\varepsilon_T = 0.0001$ (which is still 16 times larger than the default setting in Matlab) we find coverage rates as low as 38% for a nominal 95% confidence interval. Thus this table shows that the asymptotic theory provides a reliable means for obtaining confidence intervals, so long as care is taken not to set the step size too small.

[INSERT TABLES 4 AND 5 ABOUT HERE]

Finally in Table 6 we present the results of a study of the rejection rates for the test of over-identifying restrictions presented in Proposition 4. Given that we consider $W = I$ in this table, the test statistic has a non-standard distribution, and we use $K = 10,000$ simulations to obtain critical values. In this case, the limiting distribution also depends on $\hat{G}$, and based on the results in Tables 4 and 5 we compute $\hat{G}$ using a step size of $\varepsilon_T = 0.1$. Table 6 reveals that the rejection rates are close to their nominal levels, for both the equidependence models and the "different loading" models (which is a block equidependence model for the $N = 100$ case).

[INSERT TABLE 6 ABOUT HERE]

5 High-dimension copula models for equity returns

In this section we apply our proposed factor copulas to modeling the dependence between a large collection of U.S. equity returns. We study all 100 stocks that were constituents of the S&P 100 index as at December 2010. The sample period is April 2008 to December 2010, a total of $T = 696$

[11] For example, the default in many functions *Matlab* is a step size of $\varepsilon^{1/3} \approx 6 \times 10^{-6} \approx 1/(165,000)$, where $\varepsilon = 2.22 \times 10^{-16}$ is machine epsilon. This choice is optimal in certain applications, see Judd (1998) for example.

trade days. The starting point for our sample period was determined by the date of the latest addition to the S&P 100 index (Philip Morris Inc.), which has had no additions or deletions since April 2008. The stocks in our study are listed in Table 7, along with their 3-digit SIC codes, which we will use in part of our analysis below.

Table 8 presents some summary statistics of the data used in this analysis. The top panel presents sample moments of the daily returns for each stock, showing that the average daily return across all stocks for this sample period is 0.04%, corresponding to 10.08% annualized. The average volatility is 2.87% daily, which corresponds to 45.56% annually. The average individual skewness coefficient is positive, although the 25^{th} percentile of the cross-sectional distribution of this statistic is negative, indicating cross-sectional heterogeneity in the distribution of daily returns. This is also true in terms of kurtosis, which ranges from 7.60 to 11.45 from the 25^{th} to the 75^{th} percentile of the cross-sectional distribution. In the second panel of Table 8 we present information on the parameters of the AR(1)-EGARCH(1,1) models that are used to filter each of the individual return series[12]. Estimates of the parameters of these models are consistent with those reported in numerous other studies, with a small negative AR(1) coefficient found for most but not all stocks, and with EGARCH parameters strongly indicative of persistence in volatility. The asymmetry parameter, δ, in this model is negative for all but three of the 100 stocks in our sample, supporting the wide-spread finding of a "leverage effect" in the conditional volatility of equity returns.

In the lower panel of Table 8 we present summary statistics for four measures of dependence between pairs of AR-EGARCH filtered stock returns: linear correlation, rank correlation, average upper and lower 1% tail dependence (equal to $(\tau_{0.99} + \tau_{0.01})/2$), and the difference in upper and lower 10% tail dependence (equal to $\tau_{0.90} - \tau_{0.10}$). The two correlation statistics measure the sign and strength of dependence across the entire support, the third statistic measures the strength of dependence in the tails, and the fourth statistic is a measure of the asymmetry of dependence in the tails. The two correlation measures are similar, and are 0.42 and 0.44 on average. Across all possible pairs of assets (there are 4950 distinct pairs from the 100 stocks in our sample) the rank correlation varies from 0.30 to 0.50 from the 25^{th} and 75^{th} percentiles of the cross-sectional distribution, indicating the presence of mild heterogeneity in the correlation coefficients. The 1%

[12] We considered GARCH (Bollerslev, 1986), EGARCH (Nelson, 1991), and GJR-GARCH (Glosten, *et al.*, 1993) models for the conditional variance of these returns, and for almost all stocks the EGARCH model was preferred according to the BIC.

tail dependence measure is 0.07 on average, and varies from 0.00 to 0.07 across the inter-quartile range. The difference in the 10% tail dependence measures is negative on average, and indeed is negative for over 75% of the pairs of stocks, indicating asymmetric dependence between these stocks.

[INSERT TABLES 7 AND 8 ABOUT HERE]

We next present the first empirical results of this paper: the estimated parameters of seven different models for the copula of the three sets of stock returns. The models considered are the Normal copula, the Student's t copula, both with equicorrelation imposed for comparability, the Clayton copula, and four factor copulas, described by the distributions assumed for the common factor and the idiosyncractic shock: t-Normal, t-t, Skew t-Normal, Skew t-t. All models are estimated using the SMM-type method proposed in Section 3 using five dependence measures: rank correlation, and the $0.05, 0.10, 0.90, 0.95$ quantile dependence measures, averaged across pairs of assets, see Appendix B for details. The identity weight matrix is used in all cases. The value of the SMM objective function at the estimated parameters is presented for each model, along with the p-value from a test of the over-identifying restrictions based on Proposition 4. We use Proposition 3 to compute the standard errors, with $B = 1000$ bootstraps used to estimate $\Sigma_{T,S}$, and $\varepsilon_T = 0.1$ used as the step size to compute $\hat{G}$.

Table 9 reveals that the variance of the common factor, σ_z^2, is estimated by all models to be around 0.9, implying an average correlation coefficient of around 0.47. The estimated degrees of freedom parameter is around 15, and the standard error on this estimate indicates that this parameter is significantly greater than zero at the 10% level for all models, and at the 5% level for some models, indicating that allowing the common factor to have fat tails significantly improves the fit of the model[13]. Other papers have considered equicorrelation models for the dependence between large collections of stocks, see Engle and Kelly (2009) for example, but empirically showing the importance of allowing the implied common factor to have fat tails, and thus the assets to exhibit non-zero tail dependence, is novel. The most general models we consider, the Skew t-Normal and

[13] Note that the case of zero tail dependence corresponds to $\nu_z^{-1} = 0$, which is on the boundary of the parameter space for this parameter, implying that a standard t test is strictly not applicable. In such cases the squared t statistic no longer has an asymptotic χ_1^2 distribution under the null, rather it is distributed as an equal-weighted mixture of a χ_1^2 and χ_0^2, see Gourieroux and Monfort (1996b, Ch 21). The 90% and 95% critical values for this distribution are 1.64 and 2.71 (compared with 2.71 and 3.84 for the χ_1^2 distribution), which correspond to t-statistics of 1.28 and 1.65.

Skew t-t factor copulas, reveal another important feature about the implied common factor: it is asymmetrically distributed, with large crashes being more likely than large booms. The asymmetry parameter, λ_z, is estimated at -0.19, with a standard error of 0.06, which is statistically significantly different from zero at all usual levels, indicating strong evidence that these assets are more strongly correlated during market downturns than during market upturns.

[INSERT TABLE 9 ABOUT HERE]

Figure 4 presents the quantile dependence function from the estimated Normal copula and the estimated Skew t-t factor copula, along with the quantile dependence averaged across all pairs of stocks. In the lower tail we see that both copula models slightly over-estimate the probability of extreme joint crashes, while the Skew t-t factor copula provides a reasonable fit for less extreme crashes. In the upper tail we see that the Skew t-t factor copula provides a very good fit, while the Normal copula over-estimates the dependence in this tail.

Figure 5 exploits the high dimensional nature of our analysis, and plots the expected proportion of "crashes" in the remaining $(100 - k)$ stocks, conditional on observing a crash in k stocks. We show this for a "crash" defined as a once-in-a-month (1/22, around 4.6%) event and as a once-in-a-quarter (1/66, around 1.5%) event. For once-in-a-month crashes, the observed proportions track the Skew t-t factor copula well for k up to around 25 crashes, and then tracks the Normal copula for k up to around 60. When we condition on *many,* stocks having crashed, however, the data indicates stronger tail dependence than either of these models. It should be noted, however, that the empirical plot is estimated with increasing error as we condition on more and more crashes (i.e., when we move further into the joint tail). For once-in-a-quarter crashes, displayed in the lower panel of Figure 5, the empirical plot tracks that for the Normal copula well for k up to around 30, but for $k = 35$ the empirical plot jumps and follows the Skew t-t factor copula. Thus it appears that the Normal copula may be adequate for modeling moderate tail events, but a copula with greater tail dependence (such as the Skew t-t factor copula) is needed for more extreme tail events.

[INSERT FIGURES 4 AND 5 ABOUT HERE]

The last two columns of Table 9 report the value of the objective function (Q_{SMM}) and the p-value from a test of the over-identifying restrictions. The Q_{SMM} values reveal that the Skew t-Normal and Skew t-t factor copula models out-perform all the other models, and reinforce the above conclusion that allowing for a skewed common factor is important for this collection of assets. The p-values, however, are less than 0.05 for all models, indicating that none of them pass this test of goodness-of-fit. One possible source of these rejections is the assumption of equidependence, which was shown in the summary statistics in Table 8 to be questionable for this large set of stock returns.

In response to this, we next consider a model that allows for heterogeneous dependence between pairs of stocks using the "block equidependence" model discussed above. We use the first digit of Standard Industrial Classification (SIC) for each stock to group them into seven groups, see Table 10. All stocks in the same SIC group are assumed to have the same factor loading, but stocks in different groups may have different factor loadings. This greatly increases the flexibility of the model, but without generating too many additional parameters to estimate.

The results of this model are presented in Table 11. Similar to the equidependence models, the estimated degrees of freedom parameter, ν_z of factor copulas is estimated around 15 (ranging from 8 to around 20) and is significant at the 10% level in all cases. The asymmetry parameter, λ_z, is estimated at -0.17 with a standard error of 0.06, again strongly suggesting that the common factor has a left-skewed distribution. The parameters of seven different factor loadings range from 0.87 to 1.03, capturing the heterogeneity in the pair-wise dependence between these stocks. The p-value from a test that all of these loadings are equal (and thus that an equidependence assumption is adequate) is presented in the third last row, and is strongly rejected for all models. Thus allowing for block equidependence is a useful extension of this model. The SMM objective function again suggests that the Skew t-t factor copula model provides the best fit, however the test of the over-identifying restrictions yields a p-value of just over 0.05, indicating that this model is borderline rejected by the data. This perhaps suggests that an even greater degree of heterogeneity is needed to describe the dependence structure of these assets.

[INSERT TABLES 10 AND 11 ABOUT HERE]

6 Conclusion

This paper presents new models for the dependence structure, or copula, of economic variables based on a simple factor structure for the copula. These models are particularly attractive for high dimensional applications, involving fifty or more variables, as they allow the researcher to increase or decrease the flexibility of the model according the amount of data available and the dimension of the problem, and, importantly, to do so in a manner that is easily interpreted and explained. For example, allowing for a fat-tailed common factor the model captures the possibility of correlated crashes, and allowing the common factor to be asymmetrically distributed the model captures the possibility that the dependence between the variables is stronger during downturns than during upturns.

The class of factor copulas presented in this paper does not generally have a closed-form likelihood, and we present simulation based methods for the estimation of such models. These methods are simple to implement, and we derive the asymptotic properties of the resulting estimators using recent work in empirical process theory, see Fermanian, *et al.* (2004) and Rémillard (2010), combined with existing results on simulation-based estimation, see Newey and McFadden (1994) for example. Using a set of realistic simulation designs we show that the proposed estimator performs well in finite samples. We apply our various factor copulas to a set of 100 daily stock returns, over the period 2008–2010, and find significant evidence in favor of a fat-tailed common factor for these stocks, indicative of positive tail dependence, and very strong evidence of a left-skewed common factor, which is indicative of these stocks having stronger dependence during crashes than during booms.

Appendix A: Proofs

In order to prove Proposition 1, we use the following four lemmas. First, we recall the definition of stochastic equicontinuity.

Definition 1 *(Andrews (1994)) The sequence of functions* $\{\mathbf{h}_T(\cdot) : T \geq 1\}$ *is stochastically equicontinuous if* $\forall\ \varepsilon > 0$ *and* $\eta > 0, \exists\ \delta > 0$ *such that*

$$\overline{\lim}_{T\to\infty} P\left[\sup_{\|\boldsymbol{\theta}_1-\boldsymbol{\theta}_2\|<\delta} \|\mathbf{h}_T(\boldsymbol{\theta}_1) - \mathbf{h}_T(\boldsymbol{\theta}_2)\| > \eta\right] < \varepsilon$$

Lemma 1 *Under Assumptions 1 and 2,*

(i) $\frac{1}{T}\sum_{t=1}^{T} \hat{F}_i(\hat{\eta}_{it}) \hat{F}_j(\hat{\eta}_{jt}) \xrightarrow{p} \int\int uv dC_{\eta_i,\eta_j}(u, v; \boldsymbol{\theta}_0)$ *as* $T \to \infty$

(ii) $\frac{1}{T}\sum_{t=1}^{T} 1\left\{\hat{F}_i(\hat{\eta}_{it}) \leq q, \hat{F}_j(\hat{\eta}_{jt}) \leq q\right\} \xrightarrow{p} C_{\eta_i,\eta_j}(q, q; \boldsymbol{\theta}_0)$ *as* $T \to \infty$

(iii) $\frac{1}{S}\sum_{s=1}^{S}\hat{G}_i\left(x_{is}\left(\boldsymbol{\theta}\right)\right)\hat{G}_j\left(x_{js}\left(\boldsymbol{\theta}\right)\right)\xrightarrow{p}\int\int uvdC_{\eta_i,\eta_j}\left(u,v;\boldsymbol{\theta}\right)$ *for* $\forall\ \boldsymbol{\theta}\in\boldsymbol{\Theta}$ *as* $S\to\infty$

(iv) $\frac{1}{S}\sum_{s=1}^{S}1\left\{\hat{G}_i\left(x_{is}\left(\boldsymbol{\theta}\right)\right)\leq q,\hat{G}_j\left(x_{js}\left(\boldsymbol{\theta}\right)\right)\leq q\right\}\xrightarrow{p}C_{\eta_i,\eta_j}\left(q,q;\boldsymbol{\theta}\right)$ *for* $\forall\ \boldsymbol{\theta}\in\boldsymbol{\Theta}$ *as* $S\to\infty$

Proof of Lemma 1. Under Assumption 1, parts (iii) and (iv) of Lemma 1 can be proven by Theorem 3 and Theorem 6 of Fermanian, *et al.* (2004). Under Assumption 2, Corollary 1 of Rémillard (2010) proves that the empirical copula process constructed by the standardized residuals $\hat{\boldsymbol{\eta}}_t$ weakly converges to the limit of that constructed by the innovations $\boldsymbol{\eta}_t$, which combined with Theorem 3 and Theorem 6 of Fermanian, *et al.* (2004) yields parts (i) and (ii) above. ■

Lemma 2 *(Lemma 2.8 of Newey and McFadden (1994)) Suppose* $\boldsymbol{\Theta}$ *is compact and* $\mathbf{g}_0\left(\boldsymbol{\theta}\right)$ *is continuous. Then* $\sup_{\boldsymbol{\theta}\in\boldsymbol{\Theta}}\left\|\mathbf{g}_{T,S}\left(\boldsymbol{\theta}\right)-\mathbf{g}_0\left(\boldsymbol{\theta}\right)\right\|\xrightarrow{p}0$ *as* $T,S\to\infty$ *if and only if* $\mathbf{g}_{T,S}\left(\boldsymbol{\theta}\right)\xrightarrow{p}\mathbf{g}_0\left(\boldsymbol{\theta}\right)$ *for any* $\boldsymbol{\theta}\in\boldsymbol{\Theta}$ *as* $T,S\to\infty$ *and* $\mathbf{g}_{T,S}\left(\boldsymbol{\theta}\right)$ *is stochastically equicontinuous.*

Lemma 2 states that sufficient and necessary conditions of uniform convergence is pointwise convergence and stochastic equicontinuity. The following lemma shows that uniform converge of the moment functions $\mathbf{g}_{T,S}\left(\boldsymbol{\theta}\right)$ implies uniform convergence of the objective function $Q_{T,S}\left(\boldsymbol{\theta}\right)$.

Lemma 3 *If* $\sup_{\boldsymbol{\theta}\in\boldsymbol{\Theta}}\left\|\mathbf{g}_{T,S}\left(\boldsymbol{\theta}\right)-\mathbf{g}_0\left(\boldsymbol{\theta}\right)\right\|\xrightarrow{p}0$ *as* $T,S\to\infty$*, then* $\sup_{\boldsymbol{\theta}\in\boldsymbol{\Theta}}\left|Q_{T,S}\left(\boldsymbol{\theta}\right)-Q_0\left(\boldsymbol{\theta}\right)\right|\xrightarrow{p}0$ *as* $T,S\to\infty$.

Proof of Lemma 3. By the triangle inequality and Cauchy-Schwarz inequality

$$\begin{aligned}\left|Q_{T,S}\left(\boldsymbol{\theta}\right)-Q_0\left(\boldsymbol{\theta}\right)\right| &\leq \left|\left[\mathbf{g}_{T,S}\left(\boldsymbol{\theta}\right)-\mathbf{g}_0\left(\boldsymbol{\theta}\right)\right]'\hat{W}_T\left[\mathbf{g}_{T,S}\left(\boldsymbol{\theta}\right)-\mathbf{g}_0\left(\boldsymbol{\theta}\right)\right]\right|+\left|\mathbf{g}_0\left(\boldsymbol{\theta}\right)'\left(\hat{W}_T+\hat{W}_T'\right)\left[\mathbf{g}_{T,S}\left(\boldsymbol{\theta}\right)-\mathbf{g}_0\left(\boldsymbol{\theta}\right)\right]\right|\\ &\quad+\left|\mathbf{g}_0\left(\boldsymbol{\theta}\right)'\left(\hat{W}_T-W\right)\mathbf{g}_0\left(\boldsymbol{\theta}\right)\right|\\ &\leq \left\|\mathbf{g}_{T,S}\left(\boldsymbol{\theta}\right)-\mathbf{g}_0\left(\boldsymbol{\theta}\right)\right\|^2\left\|\hat{W}_T\right\|+2\left\|\mathbf{g}_0\left(\boldsymbol{\theta}\right)\right\|\left\|\mathbf{g}_{T,S}\left(\boldsymbol{\theta}\right)-\mathbf{g}_0\left(\boldsymbol{\theta}\right)\right\|\left\|\hat{W}_T\right\|\\ &\quad+\left\|\mathbf{g}_0\left(\boldsymbol{\theta}\right)\right\|^2\left\|\hat{W}_T-W\right\|\end{aligned}$$

Then note that $\mathbf{g}_0\left(\boldsymbol{\theta}\right)$ is bounded, $\hat{W}_T$ is $O_p\left(1\right)$ and converges to W by assumption 4(ii), and $\sup_{\boldsymbol{\theta}\in\boldsymbol{\Theta}}\left\|\mathbf{g}_{T,S}\left(\boldsymbol{\theta}\right)-\mathbf{g}_0\left(\boldsymbol{\theta}\right)\right\|=o_p\left(1\right)$ is given. So

$$\begin{aligned}\sup_{\boldsymbol{\theta}\in\boldsymbol{\Theta}}\left|Q_{T,S}\left(\boldsymbol{\theta}\right)-Q_0\left(\boldsymbol{\theta}\right)\right| &\leq \left(\sup_{\boldsymbol{\theta}\in\boldsymbol{\Theta}}\left\|\mathbf{g}_{T,S}\left(\boldsymbol{\theta}\right)-\mathbf{g}_0\left(\boldsymbol{\theta}\right)\right\|\right)^2 O_p\left(1\right)\\ &\quad+2O\left(1\right)\sup_{\boldsymbol{\theta}\in\boldsymbol{\Theta}}\left\|\mathbf{g}_{T,S}\left(\boldsymbol{\theta}\right)-\mathbf{g}_0\left(\boldsymbol{\theta}\right)\right\|O_p\left(1\right)+O\left(1\right)o_p\left(1\right)\\ &= o_p\left(1\right)O_p\left(1\right)+2O\left(1\right)o_p\left(1\right)O_p\left(1\right)+O\left(1\right)o_p\left(1\right)\\ &= o_p\left(1\right)\end{aligned}$$

■

Lemma 4 *(Theorem 2.1 of Newey and McFadden (1994)) Suppose that (i)* $Q_0(\boldsymbol{\theta})$ *is uniquely minimized at* $\boldsymbol{\theta}_0$; *(ii)* $\boldsymbol{\Theta}$ *is compact; (iii)* $Q_0(\boldsymbol{\theta})$ *is continuous (iv)* $\sup_{\boldsymbol{\theta}\in\boldsymbol{\Theta}}\left|\hat{Q}_T(\boldsymbol{\theta}) - Q_0(\boldsymbol{\theta})\right| \xrightarrow{p} 0$. *Then* $\hat{\boldsymbol{\theta}} \xrightarrow{p} \boldsymbol{\theta}_0$

Proof of Proposition 1. We prove this proposition by checking the conditions of Lemma 4.

(i) $Q_0(\boldsymbol{\theta})$ is uniquely minimized at $\boldsymbol{\theta}_0$ by Assumption 3(i) and positive definite W.

(ii) $\boldsymbol{\Theta}$ is compact by Assumption 3(ii).

(iii) $Q_0(\boldsymbol{\theta})$ consists of linear combinations of rank correlations and quantile dependence measures that are functions of pair-wise copula functions. Therefore, $Q_0(\boldsymbol{\theta})$ is continuous by Assumption 3(iii).

(iv) Pointwise convergence of $\mathbf{g}_{T,S}(\boldsymbol{\theta})$ to $\mathbf{g}_0(\boldsymbol{\theta})$ is shown by Lemma 1. Assumption 3(iv) is a sufficient condition for the stochastic equicontinuity of $\mathbf{g}_{T,S}$ by Lemma 2.9 of Newey and McFadden (1994). Thus $\mathbf{g}_{T,S}$ uniformly converges in probability to $\mathbf{g}_0$ by Lemma 2. This implies that $Q_{T,S}$ uniformly converges in probability to Q_0 by Lemma 3. ■

The proof of Proposition 2 uses the following three lemmas.

Lemma 5 *Let the dependence measures of interest include rank correlation and quantile dependence measures, and possibly linear combinations thereof. Then under Assumptions 1 and 2,*

$$\sqrt{T}\left(\hat{\mathbf{m}}_T - \mathbf{m}(\boldsymbol{\theta}_0)\right) \xrightarrow{d} N(0, \Sigma_0) \tag{25}$$

Proof of Lemma 5. Follows from Theorem 3 and Theorem 6 of Fermanian, *et al.* (2004) and Corollary 1, Proposition 2 and Proposition 4 of Rémillard (2010). ■

We use Theorem 7.2 of Newey & McFadden (1994) to establish the asymptotic normality of our estimator, and this relies on showing the stochastic equicontinuity of $\mathbf{v}_{T,S}(\boldsymbol{\theta})$ defined below.

Lemma 6 *Suppose that Assumptions 1, 2, and 3 hold, and further assume that* $T/S \to 0$ *as* $T, S \to \infty$. *Then* $\mathbf{v}_{T,S}(\boldsymbol{\theta}) = \sqrt{T}\left[\mathbf{g}_{T,S}(\boldsymbol{\theta}) - \mathbf{g}_0(\boldsymbol{\theta})\right]$ *is stochastically equicontinuous.*

Proof of Lemma 6. What we need to show is $\exists\ \delta$ such that as $T, S \to \infty$

$$\sup_{\|\boldsymbol{\theta}_1 - \boldsymbol{\theta}_2\| < \delta} \left| \mathbf{v}_{T,S}(\boldsymbol{\theta}_1) - \mathbf{v}_{T,S}(\boldsymbol{\theta}_2) \right| = o_p(1) \ \text{ for } \forall \boldsymbol{\theta}_1, \boldsymbol{\theta}_2 \in \boldsymbol{\Theta}$$

Without loss of generality, we assume that $\mathbf{v}_{T,S}(\boldsymbol{\theta})$ depends on just one quantile dependence.

First note that

$$\begin{aligned} \mathbf{v}_{T,S}(\boldsymbol{\theta}_1) - \mathbf{v}_{T,S}(\boldsymbol{\theta}_2) &\equiv \sqrt{T} \left(\mathbf{g}_{T,S}(\boldsymbol{\theta}_1) - \mathbf{g}_0(\boldsymbol{\theta}_1) - \mathbf{g}_{T,S}(\boldsymbol{\theta}_2) + \mathbf{g}_0(\boldsymbol{\theta}_2) \right) \\ &\equiv \sqrt{T} \left(\hat{\mathbf{m}}_T - \tilde{\mathbf{m}}_S(\boldsymbol{\theta}_1) - (\mathbf{m}(\boldsymbol{\theta}_0) - \mathbf{m}(\boldsymbol{\theta}_1)) - (\hat{\mathbf{m}}_T - \tilde{\mathbf{m}}_S(\boldsymbol{\theta}_2)) + (\mathbf{m}(\boldsymbol{\theta}_0) - \mathbf{m}(\boldsymbol{\theta}_2)) \right) \\ &= \sqrt{T} \left(\mathbf{m}(\boldsymbol{\theta}_1) - \tilde{\mathbf{m}}_S(\boldsymbol{\theta}_1) - (\mathbf{m}(\boldsymbol{\theta}_2) - \tilde{\mathbf{m}}_S(\boldsymbol{\theta}_2)) \right) \end{aligned}$$

so

$$\begin{aligned} &\sup_{\|\boldsymbol{\theta}_1 - \boldsymbol{\theta}_2\| < \delta} \left| \mathbf{v}_{T,S}(\boldsymbol{\theta}_1) - \mathbf{v}_{T,S}(\boldsymbol{\theta}_2) \right| \\ &= \sup_{\|\boldsymbol{\theta}_1 - \boldsymbol{\theta}_2\| < \delta} \sqrt{T} \left| \begin{array}{c} P(F_1(\eta_1) < q, F_2(\eta_2) < q; \boldsymbol{\theta}_1) - \frac{1}{S} \sum_{s=1}^{S} 1 \left\{ \hat{G}_1(x_{1s}(\boldsymbol{\theta}_1)) < q, \hat{G}_2(x_{2s}(\boldsymbol{\theta}_1)) < q \right\} \\ -P(F_1(\eta_1) < q, F_2(\eta_2) < q; \boldsymbol{\theta}_2) + \frac{1}{S} \sum_{s=1}^{S} 1 \left\{ \hat{G}_1(x_{1s}(\boldsymbol{\theta}_2)) < q, \hat{G}_2(x_{2s}(\boldsymbol{\theta}_2)) < q \right\} \end{array} \right| \\ &\leq \sup_{\|\boldsymbol{\theta}_1 - \boldsymbol{\theta}_2\| < \delta} \left\{ \sqrt{T} \left| P(F_1(\eta_1) < q, F_2(\eta_2) < q; \boldsymbol{\theta}_1) - \frac{1}{S} \sum_{s=1}^{S} 1 \left\{ \hat{G}_1(x_{1s}(\boldsymbol{\theta}_1)) < q, \hat{G}_2(x_{2s}(\boldsymbol{\theta}_1)) < q \right\} \right| \right. \\ &\quad \left. + \sqrt{T} \left| P(F_1(\eta_1) < q, F_2(\eta_2) < q; \boldsymbol{\theta}_2) - \frac{1}{S} \sum_{s=1}^{S} 1 \left\{ \hat{G}_1(x_{1s}(\boldsymbol{\theta}_2)) < q, \hat{G}_2(x_{2s}(\boldsymbol{\theta}_2)) < q \right\} \right| \right\} \\ &\leq \sup_{\boldsymbol{\theta}_1 \in \boldsymbol{\Theta}} \sqrt{T} \left| P(F_1(\eta_1) < q, F_2(\eta_2) < q; \boldsymbol{\theta}_1) - \frac{1}{S} \sum_{s=1}^{S} 1 \left\{ \hat{G}_1(x_{1s}(\boldsymbol{\theta}_1)) < q, \hat{G}_2(x_{2s}(\boldsymbol{\theta}_1)) < q \right\} \right| \\ &\quad + \sup_{\boldsymbol{\theta}_2 \in \boldsymbol{\Theta}} \sqrt{T} \left| P(F_1(\eta_1) < q, F_2(\eta_2) < q; \boldsymbol{\theta}_2) - \frac{1}{S} \sum_{s=1}^{S} 1 \left\{ \hat{G}_1(x_{1s}(\boldsymbol{\theta}_2)) < q, \hat{G}_2(x_{2s}(\boldsymbol{\theta}_2)) < q \right\} \right| \end{aligned}$$

In the proof of Proposition 1, we showed the uniform convergence of $\mathbf{g}_{T,S}(\boldsymbol{\theta})$ to $\mathbf{g}_0(\boldsymbol{\theta})$, and so

$$\sup_{\boldsymbol{\theta} \in \boldsymbol{\Theta}} \left| P(F_1(\eta_1) < q, F_2(\eta_2) < q; \boldsymbol{\theta}) - \frac{1}{S} \sum_{s=1}^{S} 1 \left\{ \hat{G}_1(x_{1s}(\boldsymbol{\theta})) < q, \hat{G}_2(x_{2s}(\boldsymbol{\theta})) < q \right\} \right| \xrightarrow{p} 0$$

and it is $\sqrt{S}$-consistent by Lemma 5. Since $T/S \to 0$, we have

$$\sup_{\boldsymbol{\theta} \in \boldsymbol{\Theta}} \sqrt{T} \left| P(F_1(\eta_1) < q, F_2(\eta_2) < q; \boldsymbol{\theta}) - \frac{1}{S} \sum_{s=1}^{S} 1 \left\{ \hat{G}_1(x_{1s}(\boldsymbol{\theta})) < q, \hat{G}_2(x_{2s}(\boldsymbol{\theta})) < q \right\} \right| = o_p(1)$$

Therefore,

$$\sup_{\|\boldsymbol{\theta}_1 - \boldsymbol{\theta}_2\| < \delta} \left| \mathbf{v}_{T,S}(\boldsymbol{\theta}_1) - \mathbf{v}_{T,S}(\boldsymbol{\theta}_2) \right| = o_p(1) \ \text{ as } T, S \to \infty$$

■

Lemma 7 *(Theorem 7.2 of Newey & McFadden (1994)) Suppose that* $\mathbf{g}_{T,S}\left(\hat{\boldsymbol{\theta}}\right)' \hat{W} \mathbf{g}_{T,S}\left(\hat{\boldsymbol{\theta}}\right) \leq \inf_{\boldsymbol{\theta}\in\boldsymbol{\Theta}} \mathbf{g}_{T,S}(\boldsymbol{\theta})' \hat{W} \mathbf{g}_{T,S}(\boldsymbol{\theta}) + o_p\left(T^{-1}\right), \hat{\boldsymbol{\theta}} \xrightarrow{p} \boldsymbol{\theta}_0$ *and* $\hat{W} \xrightarrow{p} W_0, W_0$ *is positive semi-definite, where there is* $\mathbf{g}_0(\boldsymbol{\theta})$ *such that (i)* $\mathbf{g}_0(\boldsymbol{\theta}_0) = 0$,*(ii)* $\mathbf{g}_0(\boldsymbol{\theta})$ *is differentiable at* $\boldsymbol{\theta}_0$ *with derivative* G_0 *such that* $G_0' W_0 G_0$ *is nonsingular,(iii)* $\boldsymbol{\theta}_0$ *is an interior point of* $\boldsymbol{\Theta}$,*(iv)* $\sqrt{T}\mathbf{g}_{T,S}(\boldsymbol{\theta}_0) \xrightarrow{d} N(0, \Sigma_0)$,*(v)* $\exists\delta$ *such that* $\sup_{\|\boldsymbol{\theta}-\boldsymbol{\theta}_0\|\leq\delta} \sqrt{T}\left\|\mathbf{g}_{T,S}(\boldsymbol{\theta}) - \mathbf{g}_{T,S}(\boldsymbol{\theta}_0) - \mathbf{g}_0(\boldsymbol{\theta})\right\| / \left[1 + \sqrt{T}\|\boldsymbol{\theta} - \boldsymbol{\theta}_0\|\right] \xrightarrow{p} 0$.*Then* $\sqrt{T}\left(\hat{\boldsymbol{\theta}} - \boldsymbol{\theta}_0\right) \xrightarrow{d} N\left(0, (G_0' W_0 G_0)^{-1} G_0' W_0 \Sigma_0 W_0 G_0 (G_0' W_0 G_0)^{-1}\right)$.

Proof of Proposition 2. We prove this proposition by checking conditions of Lemma 7.

(i) $\mathbf{g}_0(\boldsymbol{\theta}_0) = 0$ by construction of $\mathbf{g}_0(\boldsymbol{\theta}) = \mathbf{m}(\boldsymbol{\theta}_0) - \mathbf{m}(\boldsymbol{\theta})$

(ii) $\mathbf{g}_0(\boldsymbol{\theta})$ is differentiable at $\boldsymbol{\theta}_0$ with derivative G_0 such that $G_0' W_0 G_0$ is nonsingular by Assumption 4(iii).

(iii) $\boldsymbol{\theta}_0$ is an interior point of $\boldsymbol{\Theta}$ by Assumption 4(i).

(iv) For simplicity, we use only one sample quantile dependence for this proof.

$$
\begin{aligned}
\sqrt{T}\mathbf{g}_{T,S}(\boldsymbol{\theta}_0) &= \frac{1}{\sqrt{T}} \sum_{i=1}^{T} \left[1\left\{ \hat{F}_1(\hat{\eta}_{1i}) < q, \hat{F}_2(\hat{\eta}_{2i}) < q \right\} - \frac{1}{S} \sum_{j=1}^{S} 1\left\{ \hat{G}_1(\tilde{x}_{1j}(\boldsymbol{\theta}_0)) < q, \hat{G}_2(\tilde{x}_{2j}(\boldsymbol{\theta}_0)) < q \right\} \right] \\
&= \frac{1}{\sqrt{T}} \sum_{i=1}^{T} \left[\begin{array}{l} 1\left\{ \hat{F}_1(\hat{\eta}_{1i}) < q, \hat{F}_2(\hat{\eta}_{2i}) < q \right\} - P(F_1(X_1) < q, F_2(X_2) < q; \boldsymbol{\theta}_0) + \\ P(F_1(X_1) < q, F_2(X_2) < q; \boldsymbol{\theta}_0) - \frac{1}{S}\sum_{j=1}^{S} 1\left\{ \hat{G}_1(\tilde{x}_{1j}(\boldsymbol{\theta}_0)) < q, \hat{G}_2(\tilde{x}_{2j}(\boldsymbol{\theta}_0)) < q \right\} \end{array} \right] \\
&= \frac{1}{\sqrt{T}} \sum_{i=1}^{T} \left[1\left\{ \hat{F}_1(\hat{\eta}_{1i}) < q, \hat{F}_2(\hat{\eta}_{2i}) < q \right\} - P(F_1(X_1) < q, F_2(X_2) < q; \boldsymbol{\theta}_0) \right] \\
&\quad - \sqrt{T}\frac{1}{S} \sum_{j=1}^{S} \left[1\left\{ \hat{G}_1(\tilde{x}_{1j}(\boldsymbol{\theta}_0)) < q, \hat{G}_2(\tilde{x}_{2j}(\boldsymbol{\theta}_0)) < q \right\} - P(F_1(X_1) < q, F_2(X_2) < q; \boldsymbol{\theta}_0) \right] \\
&= \underbrace{\frac{1}{\sqrt{T}} \sum_{i=1}^{T} \left[1\left\{ \hat{F}_1(\hat{\eta}_{1i}) < q, \hat{F}_2(\hat{\eta}_{2i}) < q \right\} - P(F_1(X_1) < q, F_2(X_2) < q; \boldsymbol{\theta}_0) \right]}_{\xrightarrow{d} N(0,\Sigma_1) \text{ by Lemma 5}} \\
&\quad - \underbrace{\frac{\sqrt{T}}{\sqrt{S}}}_{=o(1)} \underbrace{\frac{1}{\sqrt{S}} \sum_{j=1}^{S} \left[1\left\{ \hat{G}_1(\tilde{x}_{1j}(\boldsymbol{\theta}_0)) < q, \hat{G}_2(\tilde{x}_{2j}(\boldsymbol{\theta}_0)) < q \right\} - P(F_1(X_1) < q, F_2(X_2) < q; \boldsymbol{\theta}_0) \right]}_{=O_p(1) \text{ by Lemma 5}}
\end{aligned}
$$

The similar proofs for other quantile dependence measures, rank correlations and linear combinations of these measures give the desired results. Thus,

$$
\sqrt{T}\mathbf{g}_{T,S}(\boldsymbol{\theta}_0) \xrightarrow{d} N(0, \Sigma_0)
$$

(v) We already established the stochastic equicontinuity of $\mathbf{v}_{T,S}(\boldsymbol{\theta}) = \sqrt{T}\left[\boldsymbol{g}_{T,S}(\boldsymbol{\theta}) - \boldsymbol{g}_0(\boldsymbol{\theta})\right]$ by Lemma 6, i.e. for $\forall\varepsilon > 0$, $\eta > 0, \exists\delta$ such that

$$\begin{aligned} & \overline{\lim}_{T\to\infty} P\left[\sup_{\|\boldsymbol{\theta}-\boldsymbol{\theta}_0\|<\delta} \|\mathbf{v}_{T,S}(\boldsymbol{\theta}) - \mathbf{v}_{T,S}(\boldsymbol{\theta}_0)\| > \eta\right] \\ = & \overline{\lim}_{T\to\infty} P\left[\sup_{\|\boldsymbol{\theta}-\boldsymbol{\theta}_0\|<\delta} \sqrt{T}\left\|\boldsymbol{g}_{T,S}(\boldsymbol{\theta}) - \boldsymbol{g}_{T,S}(\boldsymbol{\theta}_0) - \boldsymbol{g}_0(\boldsymbol{\theta})\right\| > \eta\right] < \varepsilon \end{aligned}$$

and

$$\sqrt{T}\left\|\boldsymbol{g}_{T,S}(\boldsymbol{\theta}) - \boldsymbol{g}_{T,S}(\boldsymbol{\theta}_0) - \boldsymbol{g}_0(\boldsymbol{\theta})\right\| / \left[1 + \sqrt{T}\|\boldsymbol{\theta} - \boldsymbol{\theta}_0\|\right] \leq \sqrt{T}\left\|\boldsymbol{g}_{T,S}(\boldsymbol{\theta}) - \boldsymbol{g}_{T,S}(\boldsymbol{\theta}_0) - \boldsymbol{g}_0(\boldsymbol{\theta})\right\|$$

$$\begin{aligned} & \overline{\lim}_{T\to\infty} P\left[\sup_{\|\boldsymbol{\theta}-\boldsymbol{\theta}_0\|<\delta} \sqrt{T}\left\|\boldsymbol{g}_{T,S}(\boldsymbol{\theta}) - \boldsymbol{g}_{T,S}(\boldsymbol{\theta}_0) - \boldsymbol{g}_0(\boldsymbol{\theta})\right\| / \left[1 + \sqrt{T}\|\boldsymbol{\theta} - \boldsymbol{\theta}_0\|\right] > \eta\right] \\ \leq & \overline{\lim}_{T\to\infty} P\left[\sup_{\|\boldsymbol{\theta}-\boldsymbol{\theta}_0\|<\delta} \sqrt{T}\left\|\boldsymbol{g}_{T,S}(\boldsymbol{\theta}) - \boldsymbol{g}_{T,S}(\boldsymbol{\theta}_0) - \boldsymbol{g}_0(\boldsymbol{\theta})\right\| > \eta\right] < \varepsilon \end{aligned}$$

■

Proof of Proposition 3. First, we prove the consistency of the numerical derivatives $\hat{G}_{T,S}$. This part of the proof is similar to that of Theorem 7.4 in Newey and McFadden (1994). We will consider one-sided derivatives first, with the same arguments applying to two-sided derivatives. We know that $\left\|\hat{\boldsymbol{\theta}}_{T,S} - \boldsymbol{\theta}_0\right\| = O_p\left(T^{-1/2}\right)$ by the conclusion of Proposition 2. Also, by assumption we have $\varepsilon_T \to 0$ and $\varepsilon_T\sqrt{T} \to \infty$, so

$$\left\|\hat{\boldsymbol{\theta}}_{T,S} + \mathbf{e}_k\varepsilon_T - \boldsymbol{\theta}_0\right\| \leq \left\|\hat{\boldsymbol{\theta}}_{T,S} - \boldsymbol{\theta}_0\right\| + \|\mathbf{e}_k\varepsilon_T\| = O_p\left(T^{-1/2}\right) + O(\varepsilon_T) = O_p(\varepsilon_T)$$

(Recall that $\mathbf{e}_k$ is the k^{th} unit vector.) In the proof of Proposition 2, it is shown that $\exists\delta$ such that

$$\sup_{\|\boldsymbol{\theta}-\boldsymbol{\theta}_0\|\leq\delta} \sqrt{T}\|\mathbf{g}_{T,S}(\boldsymbol{\theta}) - \mathbf{g}_{T,S}(\boldsymbol{\theta}_0) - \mathbf{g}_0(\boldsymbol{\theta})\| / \left[1 + \sqrt{T}\|\boldsymbol{\theta} - \boldsymbol{\theta}_0\|\right] = o_p(1)$$

Substituting $\hat{\boldsymbol{\theta}}_{T,S} + \mathbf{e}_k\varepsilon_T$ for $\boldsymbol{\theta}$, then for T, S large, it follows that

$$\sqrt{T}\left\|\mathbf{g}_{T,S}\left(\hat{\boldsymbol{\theta}}_{T,S} + \mathbf{e}_k\varepsilon_T\right) - \mathbf{g}_{T,S}(\boldsymbol{\theta}_0) - \mathbf{g}_0\left(\hat{\boldsymbol{\theta}}_{T,S} + \mathbf{e}_k\varepsilon_T\right)\right\| / \left[1 + \sqrt{T}\left\|\hat{\boldsymbol{\theta}}_{T,S} + \mathbf{e}_k\varepsilon_T - \boldsymbol{\theta}_0\right\|\right] \leq o_p(1)$$

$$\begin{aligned} \text{so } \left\|\mathbf{g}_{T,S}\left(\hat{\boldsymbol{\theta}}_{T,S} + \mathbf{e}_k\varepsilon_T\right) - \mathbf{g}_{T,S}(\boldsymbol{\theta}_0) - \mathbf{g}_0\left(\hat{\boldsymbol{\theta}}_{T,S} + \mathbf{e}_k\varepsilon_T\right)\right\| & \leq \left[1 + \sqrt{T}\underbrace{\left\|\hat{\boldsymbol{\theta}}_{T,S} + \mathbf{e}_k\varepsilon_T - \boldsymbol{\theta}_0\right\|}_{=O_p(\varepsilon_T)}\right] o_p\left(\frac{1}{\sqrt{T}}\right) \\ & = \sqrt{T}O_p(\varepsilon_T)\, o_p\left(\frac{1}{\sqrt{T}}\right) = O_p(\varepsilon_T)\, o_p(1) \\ & = o_p(\varepsilon_T) \end{aligned} \tag{26}$$

On the other hand, since $\mathbf{g}_0(\boldsymbol{\theta})$ is differentiable at $\boldsymbol{\theta}_0$ with derivative G_0 by Assumption 4(iii), a Taylor expansion of $\mathbf{g}_0\left(\hat{\boldsymbol{\theta}}_{T,S}+\mathbf{e}_k\varepsilon_T\right)$ around $\boldsymbol{\theta}_0$ is

$$\mathbf{g}_0\left(\hat{\boldsymbol{\theta}}_{T,S}+\mathbf{e}_k\varepsilon_T\right) = \mathbf{g}_0(\boldsymbol{\theta}_0) + G_0\cdot\left(\hat{\boldsymbol{\theta}}_{T,S}+\mathbf{e}_k\varepsilon_T-\boldsymbol{\theta}_0\right) + o\left(\left\|\hat{\boldsymbol{\theta}}_{T,S}+\mathbf{e}_k\varepsilon_T-\boldsymbol{\theta}_0\right\|\right)$$

with $\mathbf{g}_0(\boldsymbol{\theta}_0) = 0$ by Assumption 3(i). Then divide by ε_T,

$$\begin{aligned}
\mathbf{g}_0\left(\hat{\boldsymbol{\theta}}_{T,S}+\mathbf{e}_k\varepsilon_T\right)/\varepsilon_T &= G_0\cdot\left(\hat{\boldsymbol{\theta}}_{T,S}+\mathbf{e}_k\varepsilon_T-\boldsymbol{\theta}_0\right)/\varepsilon_T + o\left(\varepsilon_T^{-1}\left\|\hat{\boldsymbol{\theta}}_{T,S}+\mathbf{e}_k\varepsilon_T-\boldsymbol{\theta}_0\right\|\right) \\
\text{so}\ \ \mathbf{g}_0\left(\hat{\boldsymbol{\theta}}_{T,S}+\mathbf{e}_k\varepsilon_T\right)/\varepsilon_T - G_0\mathbf{e}_k &= G_0\cdot\left(\hat{\boldsymbol{\theta}}_{T,S}-\boldsymbol{\theta}_0\right)/\varepsilon_T + o\left(\varepsilon_T^{-1}\left\|\hat{\boldsymbol{\theta}}_{T,S}+\mathbf{e}_k\varepsilon_T-\boldsymbol{\theta}_0\right\|\right)
\end{aligned}$$

The triangle inequality implies that

$$\begin{aligned}
\left\|\mathbf{g}_0\left(\hat{\boldsymbol{\theta}}_{T,S}+\mathbf{e}_k\varepsilon_T\right)/\varepsilon_T - G_0\mathbf{e}_k\right\| &\leq \left\|G_0\cdot\left(\hat{\boldsymbol{\theta}}_{T,S}-\boldsymbol{\theta}_0\right)/\varepsilon_T\right\| + o\left(\varepsilon_T^{-1}\left\|\hat{\boldsymbol{\theta}}_{T,S}+\mathbf{e}_k\varepsilon_T-\boldsymbol{\theta}_0\right\|\right) \\
&= \frac{1}{\sqrt{T}\varepsilon_T}\left\|G_0\cdot\sqrt{T}\left(\hat{\boldsymbol{\theta}}_{T,S}-\boldsymbol{\theta}_0\right)\right\| + \varepsilon_T^{-1}\left\|\hat{\boldsymbol{\theta}}_{T,S}+\mathbf{e}_k\varepsilon_T-\boldsymbol{\theta}_0\right\|o(1) \\
&= o(1)\,O_p(1) + \varepsilon_T^{-1}O_p(\varepsilon_T)\,o(1) \qquad (27) \\
&= o_p(1)
\end{aligned}$$

Combining the inequalities in equations (26) and (27) gives

$$\begin{aligned}
\left(\frac{\mathbf{g}_{T,S}\left(\hat{\boldsymbol{\theta}}_{T,S}+\mathbf{e}_k\varepsilon_T\right)-\mathbf{g}_{T,S}(\boldsymbol{\theta}_0)}{\varepsilon_T} - G_0\mathbf{e}_k\right) &= \left(\frac{\mathbf{g}_{T,S}\left(\hat{\boldsymbol{\theta}}_{T,S}+\mathbf{e}_k\varepsilon_T\right)-\mathbf{g}_{T,S}(\boldsymbol{\theta}_0)-\mathbf{g}_0\left(\hat{\boldsymbol{\theta}}_{T,S}+\mathbf{e}_k\varepsilon_T\right)}{\varepsilon_T}\right) \\
&\quad + \left(\mathbf{g}_0\left(\hat{\boldsymbol{\theta}}_{T,S}+\mathbf{e}_k\varepsilon_T\right)/\varepsilon_T - G_0\mathbf{e}_k\right) \\
\left\|\frac{\mathbf{g}_{T,S}\left(\hat{\boldsymbol{\theta}}_{T,S}+\mathbf{e}_k\varepsilon_T\right)-\mathbf{g}_{T,S}(\boldsymbol{\theta}_0)}{\varepsilon_T} - G_0\mathbf{e}_k\right\| &\leq \left\|\frac{\mathbf{g}_{T,S}\left(\hat{\boldsymbol{\theta}}_{T,S}+\mathbf{e}_k\varepsilon_T\right)-\mathbf{g}_{T,S}(\boldsymbol{\theta}_0)-\mathbf{g}_0\left(\hat{\boldsymbol{\theta}}_{T,S}+\mathbf{e}_k\varepsilon_T\right)}{\varepsilon_T}\right\| \\
&\quad + \left\|\mathbf{g}_0\left(\hat{\boldsymbol{\theta}}_{T,S}+\mathbf{e}_k\varepsilon_T\right)/\varepsilon_T - G_0\mathbf{e}_k\right\| \\
&\leq o_p(1)
\end{aligned}$$

Then,

$$\frac{\mathbf{g}_{T,S}\left(\hat{\boldsymbol{\theta}}_{T,S}+\mathbf{e}_k\varepsilon_T\right)-\mathbf{g}_{T,S}(\boldsymbol{\theta}_0)}{\varepsilon_T} \xrightarrow{p} G_0\mathbf{e}_k$$

and the same arguments can be applied to the two-sided derivative:

$$\frac{\mathbf{g}_{T,S}\left(\hat{\boldsymbol{\theta}}_{T,S}+\mathbf{e}_k\varepsilon_T\right)-\mathbf{g}_{T,S}\left(\hat{\boldsymbol{\theta}}_{T,S}-\mathbf{e}_k\varepsilon_T\right)}{2\varepsilon_T} \xrightarrow{p} G_0\mathbf{e}_k$$

This holds for each column $k = 1, 2..., p$. Thus $\hat{G}_{T,S} \xrightarrow{p} G_0$.

Next, we show the consistency of $\hat{\Sigma}_{T,B}$. If $\boldsymbol{\mu}_t$ and $\boldsymbol{\sigma}_t$ are known constant, or if $\boldsymbol{\phi}_0$ is known, then the result follows from Theorems 5 and 6 of Fermanian, *et al.* (2004). When $\boldsymbol{\phi}_0$ is estimated, the result is obtained by combining the results in Fermanian, *et al.* with those of Rémillard (2010): For simplicity, assume that only one dependence measure is used. Let $\hat{\tau}_{ij}$ and $\hat{\tau}_{ij}^{(b)}$ be the sample quantile dependence constructed from the standardized residuals $\left\{\hat{\eta}_t^i, \hat{\eta}_t^j\right\}_{t=1}^T$ and from the bootstrap counterpart $\left\{\hat{\eta}_t^{(b)i}, \hat{\eta}_t^{(b)j}\right\}_{t=1}^T$. Also, define the corresponding estimates, $\ddot{\tau}_{ij}$ and $\ddot{\tau}_{ij}^{(b)}$, using the true innovations $\left\{\eta_t^i, \eta_t^j\right\}_{t=1}^T$ and the bootstrapped true innovations $\left\{\eta_t^{(b)i}, \eta_t^{(b)j}\right\}_{t=1}^T$ (where the same bootstrap time indices are used for both $\left\{\hat{\eta}_t^{(b)i}, \hat{\eta}_t^{(b)j}\right\}_{t=1}^T$ and $\left\{\eta_t^{(b)i}, \eta_t^{(b)j}\right\}_{t=1}^T$). Define true τ as τ_0. Theorem 5 of Fermanian *et al.* (2004) shows that

$$\sqrt{T}\left(\ddot{\tau}_{ij} - \tau_0\right) = \sqrt{T}\left(\ddot{\tau}_{ij}^{(b)} - \ddot{\tau}_{ij}\right) + o_p(1)$$

Corollary 1 and Proposition 4 of Rémillard (2010) shows, under Assumption 2, that

$$\begin{aligned} \sqrt{T}\left(\hat{\tau}_{ij} - \ddot{\tau}_{ij}\right) &= o_p(1) \\ \text{and} \quad \sqrt{T}\left(\hat{\tau}_{ij}^{(b)} - \ddot{\tau}_{ij}^{(b)}\right) &= o_p(1) \end{aligned}$$

Combining those three equations, we obtain

$$\sqrt{T}\left(\hat{\tau}^{ij} - \tau_0\right) = \sqrt{T}\left(\hat{\tau}_{ij}^{(b)} - \hat{\tau}^{ij}\right) + o_p(1), \text{ as } T, B \to \infty$$

and so equation (18) is a consistent estimator of Σ_0. ■

Proof of Proposition 4. A Taylor expansion of $\mathbf{g}_0\left(\hat{\boldsymbol{\theta}}_{T,S}\right)$ around $\boldsymbol{\theta}_0$ yields

$$\sqrt{T}\mathbf{g}_0\left(\hat{\boldsymbol{\theta}}_{T,S}\right) = \sqrt{T}\mathbf{g}_0\left(\boldsymbol{\theta}_0\right) + G_0 \cdot \sqrt{T}\left(\hat{\boldsymbol{\theta}}_{T,S} - \boldsymbol{\theta}_0\right) + o\left(\sqrt{T}\left\|\hat{\boldsymbol{\theta}}_{T,S} - \boldsymbol{\theta}_0\right\|\right)$$

and since $\mathbf{g}_0\left(\boldsymbol{\theta}_0\right) = 0$ and $\sqrt{T}\left\|\hat{\boldsymbol{\theta}}_{T,S} - \boldsymbol{\theta}_0\right\| = O_p(1)$

$$\sqrt{T}\mathbf{g}_0\left(\hat{\boldsymbol{\theta}}_{T,S}\right) = G_0 \cdot \sqrt{T}\left(\hat{\boldsymbol{\theta}}_{T,S} - \boldsymbol{\theta}_0\right) + o_p(1) \tag{28}$$

Then consider the following expansion of $\mathbf{g}_{T,S}\left(\hat{\boldsymbol{\theta}}_{T,S}\right)$ around $\boldsymbol{\theta}_0$

$$\sqrt{T}\mathbf{g}_{T,S}\left(\hat{\boldsymbol{\theta}}_{T,S}\right) = \sqrt{T}\mathbf{g}_{T,S}\left(\boldsymbol{\theta}_0\right) + \hat{G}_{T,S} \cdot \sqrt{T}\left(\hat{\boldsymbol{\theta}}_{T,S} - \boldsymbol{\theta}_0\right) + R_{T,S}\left(\hat{\boldsymbol{\theta}}_{T,S}\right) \tag{29}$$

where the remaining term is captured by $R_{T,S}\left(\hat{\boldsymbol{\theta}}_{T,S}\right)$. Combining equations (28) and (29) we obtain

$$\sqrt{T}\left[\mathbf{g}_{T,S}\left(\hat{\boldsymbol{\theta}}_{T,S}\right) - \mathbf{g}_{T,S}\left(\boldsymbol{\theta}_0\right) - \mathbf{g}_0\left(\hat{\boldsymbol{\theta}}_{T,S}\right)\right] = \left(\hat{G}_{T,S} - G_0\right) \cdot \sqrt{T}\left(\hat{\boldsymbol{\theta}}_{T,S} - \boldsymbol{\theta}_0\right) + R_{T,S}\left(\hat{\boldsymbol{\theta}}_{T,S}\right) + o_p(1)$$

Lemma 6 shows the stochastic equicontinuity of $\mathbf{v}_{T,S}(\boldsymbol{\theta})$, which implies (see proof of Proposition 2) that

$$\sqrt{T}\left[\mathbf{g}_{T,S}\left(\hat{\boldsymbol{\theta}}_{T,S}\right)-\mathbf{g}_{T,S}(\boldsymbol{\theta}_0)-\mathbf{g}_0\left(\hat{\boldsymbol{\theta}}_{T,S}\right)\right]=o_p(1)$$

By Proposition 3, $\hat{G}_{T,S}-G_0=o_p(1)$, which implies $R_{T,S}\left(\hat{\boldsymbol{\theta}}_{T,S}\right)=o_p(1)$. Thus, we obtain the expansion of $\mathbf{g}_{T,S}\left(\hat{\boldsymbol{\theta}}_{T,S}\right)$ around $\boldsymbol{\theta}_0$:

$$\sqrt{T}\mathbf{g}_{T,S}\left(\hat{\boldsymbol{\theta}}_{T,S}\right)=\sqrt{T}\mathbf{g}_{T,S}(\boldsymbol{\theta}_0)+\hat{G}_{T,S}\cdot\sqrt{T}\left(\hat{\boldsymbol{\theta}}_{T,S}-\boldsymbol{\theta}_0\right)+o_p(1) \tag{30}$$

The remainder of the proof is the same as in standard GMM applications: From the proof of Proposition 2, we have $\sqrt{T}\mathbf{g}_{T,S}(\boldsymbol{\theta}_0)\xrightarrow{d}N(0,\Sigma_0)$ and rewrite this as $-\Sigma_0^{-1/2}\sqrt{T}\mathbf{g}_{T,S}(\boldsymbol{\theta}_0)\equiv\mathbf{u}_{T,S}\xrightarrow{d}\mathbf{u}\sim N(0,I)$, and from Proposition 2, we have $\sqrt{T}\left(\hat{\boldsymbol{\theta}}_{T,S}-\boldsymbol{\theta}_0\right)=(G_0'W_0G_0)^{-1}G_0'W_0\Sigma_0^{1/2}\mathbf{u}_{T,S}+o_p(1)$. By these two equation and Proposition 3, equation (30) becomes

$$\begin{aligned}\sqrt{T}\mathbf{g}_{T,S}\left(\hat{\boldsymbol{\theta}}_{T,S}\right) &= -\hat{\Sigma}_{T,B}^{1/2}\mathbf{u}_{T,S}+\hat{G}_{T,S}\left(\hat{G}_{T,S}'\hat{W}_T\hat{G}\right)^{-1}\hat{G}_{T,S}'\hat{W}_T\Sigma_{T,B}^{1/2}\mathbf{u}_{T,S}+o_p(1)\\ &= -\hat{\Sigma}_{T,B}^{1/2}\hat{R}\mathbf{u}_{T,S}+o_p(1)\end{aligned}$$

where $\hat{R}\equiv\left(I-\hat{\Sigma}_{T,B}^{-1/2}\hat{G}_{T,S}\left(\hat{G}_{T,S}'\hat{W}_T\hat{G}\right)^{-1}\hat{G}_{T,S}'\hat{W}_T\Sigma_{T,B}^{1/2}\right)$. The test statistic is

$$\begin{aligned}T\mathbf{g}_{T,S}\left(\hat{\boldsymbol{\theta}}_{T,S}\right)'\hat{W}_T\mathbf{g}_{T,S}\left(\hat{\boldsymbol{\theta}}_{T,S}\right) &= \mathbf{u}_{T,S}'\hat{R}'\hat{\Sigma}_{T,B}^{1/2\prime}\hat{W}_T\hat{\Sigma}_{T,B}^{1/2}\hat{R}\mathbf{u}_{T,S}+o_p(1)\\ &= \mathbf{u}'R_0\Sigma_0^{1/2\prime}W_0\Sigma_0^{1/2}R_0\mathbf{u}+o_p(1)\end{aligned}$$

where $R_0\equiv\left(I-\Sigma_0^{-1/2}G_0(G_0'W_0G_0)^{-1}G_0'W_0\Sigma_0^{1/2}\right)$. When $\hat{W}_T=\hat{\Sigma}_{T,B}^{-1}$, $\hat{R}$ is symmetric and idempotent with $rank\left(\hat{R}\right)=tr\left(\hat{R}\right)=m-p$, and the test statistic converges to a χ^2_{m-p} random variable, as usual. In general, the asymptotic distribution is a sample-dependent combination of m independent standard Normal variables, namely that of $\mathbf{u}'R_0\Sigma_0^{1/2\prime}W_0\Sigma_0^{1/2}R_0\mathbf{u}$ where $\mathbf{u}\sim N(0,I)$.

■

Appendix B: Choice of dependence measures for estimation

To implement the SMM estimator of these copula models we must first choose which dependence measures to use in the SMM estimation. We draw on "pure" measures of dependence, in the sense that they are solely affected by changes in the copula, and not by changes in the marginal distributions. For examples of such measures, see Joe (1997, Chapter 2) or Nelsen (2006, Chapter 5). Our preliminary studies of estimation accuracy and identification lead us to use pair-wise rank correlation and quantile dependence, see equations (10) and (11), with $q = [0.01, 0.10, 0.90, 0.99]$, giving us five dependence measures for each pair of variables[14]. Let δ_{ij} denote one of the dependence measures (i.e., rank correlation or quantile dependence at different levels of q) between variables i and j, and define the "pair-wise dependence matrix":

$$D = \begin{bmatrix} 1 & \delta_{12} & \cdots & \delta_{1N} \\ \delta_{12} & 1 & \cdots & \delta_{2N} \\ \vdots & \vdots & \ddots & \vdots \\ \delta_{1N} & \delta_{2N} & \cdots & 1 \end{bmatrix} \tag{31}$$

Where applicable, we exploit the (block) equidependence feature of the models in defining the "moments" to match. For the initial set of simulation results and for the first model in the empirical section, the model implies equidependence, and we use as "moments" the average of these five dependence measures across all pairs, reducing the number of moments to match from $5N(N-1)/2$ to just 5:

$$\bar{\delta} \equiv \frac{2}{N(N-1)} \sum_{i=1}^{N-1} \sum_{j=i+1}^{N} \hat{\delta}_{ij} \tag{32}$$

For a model with different loadings on the common factor (as in equation 22) equidependence does not hold. Yet the common factor aspect of the model implies that there are $\mathcal{O}(N)$, not $\mathcal{O}(N^2)$, parameters driving the pair-wise dependence matrices. In light of this, we use the $N \times 1$ vector $[\bar{\delta}_1, ..., \bar{\delta}_N]'$, where

$$\bar{\delta}_i \equiv \frac{1}{N} \sum_{j=1}^{n} \hat{\delta}_{ij}$$

and so $\bar{\delta}_i$ is the average of all pair-wise dependence measures that involve variable i. This yields a total of $5N$ moments for estimation.

For the block-equidependence version of this model (used for the $N = 100$ case in the simulation, and in the second set of models for the empirical section), we exploit the fact that (i) all variables in the same group exhibit equidependence, and (ii) any pair of variables (i, j) in groups (r, s) has the same dependence as any other pair (i', j') in the same two groups (r, s). This allows us to average all intra- and inter-group dependence measures. Consider the following general design, where we have N variables, m groups, and $k = N/m$ variables per group. Then decompose the $(N \times N)$

[14] The selection of an "optimal" set of dependence measures is an interesting problem, and is left for future research. For related work in GMM applications see Hansen (1985), Andrews (1999) and Hall (2005).

matrix D into sub-matrices according to the groups:

$$\underset{(N\times N)}{D} = \begin{bmatrix} D_{11} & D_{12} & \cdots & D_{1m} \\ D_{12} & D_{22} & \cdots & D_{2m} \\ \vdots & \vdots & \ddots & \vdots \\ D_{1m} & D_{2m} & \cdots & D_{mm} \end{bmatrix}, \text{ where each } D_{ij} \text{ is } (k \times k) \tag{33}$$

Then create a matrix of average values from each of these matrices, taking into account the fact that the diagonal blocks are symmetric:

$$\underset{(m\times m)}{D^*} = \begin{bmatrix} \delta^*_{11} & \delta^*_{12} & \cdots & \delta^*_{1m} \\ \delta^*_{12} & \delta^*_{22} & \cdots & \delta^*_{2m} \\ \vdots & \vdots & \ddots & \vdots \\ \delta^*_{1m} & \delta^*_{2m} & \cdots & \delta^*_{mm} \end{bmatrix} \tag{34}$$

$$\text{where } \delta^*_{ss} \equiv \frac{2}{k(k-1)} \sum\sum \hat{\delta}_{ij}, \text{ avg of \textit{all upper triangle} values in } D_{ss}$$

$$\delta^*_{rs} = \frac{1}{k^2} \sum\sum \hat{\delta}_{ij}, \text{ avg of \textit{all} elements in matrix } D_{rs}, \quad r \neq s$$

Finally, similar to the previous model, create the vector of average measures $\left[\bar{\delta}^*_1, ..., \bar{\delta}^*_m\right]$, where

$$\bar{\delta}^*_i \equiv \frac{1}{m} \sum_{j=1}^{m} \delta^*_{ij} \tag{35}$$

This gives as a total of m moments for each dependence measure, so $5m$ in total.

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] Andersen, L. J. Sidenius, 2004, Extensions to the Gaussian Copula: Random Recovery and Random Factor Loadings, *Journal of Credit Risk*, 1, 29-70.

[3] Andrews, D.W.K., 1994, Empirical Process Methods in Econometrics, in R. Engle and D.McFadden (eds.), *Handbook of Econometrics*, 4, 2247-2294.

[4] Andrews, D.W.K., 1999, Consistent moment selection procedures for Generalized Method of Moments estimation, *Econometrica*, 67, 543-564.

[5] Andrews, D.W.K., 2001, Testing When a Parameter is on the Boundary of the Maintained Hypothesis, *Econometrica*, 69(3), 683-734.

[6] Barndorff-Nielsen, O.E., 1978, Hyperbolic distributions and distributions on hyperbolae, *Scandinavian Journal of Statistics*, 5, 151-157.

[7] Barndorff-Nielsen, O.E., 1997, Normal inverse Gaussian distributions and stochastic volatility modelling, *Scandinavian Journal of Statistics*, 24, 1-13.

[8] Barndorff-Nielsen, O.E., and N. Shephard, 2009, *Financial Volatility in Continuous Time*, Cambridge University Press, Cambridge, forthcoming.

[9] Bollerslev, T., 1986, Generalized autoregressive conditional heteroskedasticity, *Journal of Econometrics*, 31, 307-327.

[10] Bollerslev, T., 1990, Modelling the coherence in short-run nominal exchange rates: a multivariate generalized ARCH approach, *Review of Economics and Statistics*, 72, 498-505.

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

[12] 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.

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

[14] Coval, J., J. Jurek and E. Stafford, 2009, The economics of structured finance, *Journal of Economic Perspectives*, 23, 3-25.

[15] 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.

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

[17] Engle, R.F., and B. Kelly, 2008, Dynamic Equicorrelation, working paper, Stern School of Business, New York University.

[18] Engle, R.F., N. Shephard and K. Sheppard, 2008, Fitting vast dimensional time-varying covariance models, working paper, Oxford-Man Institute, University of Oxford.

[19] Fermanian J., D. Radulovic, and M. Wegkamp, 2004, Weak Convergence of Empirical Copula Process, *Bernoulli*, 10, 847-860.

[20] 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.

[21] Gouriéroux, C., and A. Monfort, 1996a, *Simulation-Based Econometric Methods*, Oxford University Press, Oxford.

[22] Gouriéroux, C. and A. Monfort, 1996b, *Statistics and Econometric Models, Volume 2*, translated from the French by Q. Vuong, Cambridge University Press, Cambridge.

[23] Hall, A.R., 2005, *Generalized Method of Moments*, Oxford University Press, Oxford.

[24] Hansen, L.P., 1985, A method of calculating bounds on the asymptotic covariance matrices of generalized method of moments estimators, *Journal of Econometrics*, 30, 203-238.

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

[26] Hautsch, N. L.M. Kyj and R.C.A. Oomen, 2010, A blocking and regularization approach to high dimensional realized covariance estimation, *Journal of Applied Econometrics*, forthcoming.

[27] Heinen, A., and A. Valdesogo, 2009, Asymmetric CAPM dependence for large dimensions: the canonical vine autoregressive model, working paper, Univerisidad Carlos III.

[28] Hull, J. and A. White, 2004, Valuation of a CDO and an n^{th} to default CDS without Monte Carlo simulation, *Journal of Derivatives*, 12, 8-23.

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

[30] Judd, K.L, 1998, *Numerical Methods in Economics*, MIT Press, Cambridge, USA.

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

[32] Manner, H., amd O. Reznikova, 2010, A survey on time-varying copulas: Specification, simulations and estimation, *Econometric Reviews*, forthcoming.

[33] McFadden, D., 1989, A method of simulated moments for estimation of discrete response models without numerical integration, *Econometrica*, 47, 995-1026.

[34] McNeil, A.J., R. Frey and P. Embrechts, 2005, *Quantitative Risk Management*, Princeton University Press, New Jersey.

[35] McNeil, A.J., and J. Nešlehová, 2010, From Archimedean to Liouville copulas, *Journal of Multivariate Analysis*, 101, 1772-1790.

[36] Min, A., and C. Czado, 2010, Bayesian inference for multivariate copulas using pair-copula constructions, *Journal of Financial Econometrics*, 8, 511-546.

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

[38] Nelson, D. B., 1991, Conditional heteroskedasticity in asset returns: A new approach, *Econometrica*, 59, 347-370.

[39] Newey, W. K., 1991, Uniform Convergence in Probability and Stochastic Equicontinuity, *Econometrica*, 59, 1161-1167.

[40] Newey, W. K., and D. McFadden, 1994, Large Sample Estimation and Hypothesis Testing, in R.Engle and D. McFadden (eds.), *Handbook of Econometrics*, 4, 2111-2245.

[41] Pakes, A., and D. Pollard, 1989, Simulation and the asymptotics of optimization estimators, *Econometrica*, 47, 1027-1057.

[42] Palm, F.C. and J.-P. Urbain, 2011, Factor structures for panel and multivariate time series data, *Journal of Econometrics*, 163(1), 1-3.

[43] Patton, A.J., 2006, Modelling asymmetric exchange rate dependence, *International Economic Review*, 47(2), 527-556.

[44] Patton, A.J., 2009, 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.

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

[46] Rothenberg, T.J., 1971, Identification of Parametric Models, *Econometrica*, 39(3), 577-591.

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

[48] van der Voort, M., 2005, Factor copulas: totally external defaults, working paper, Erasmus University Rotterdam.

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

[50] Wolak, F.A., 1989, Testing Inequality Contraints in Linear Econometic Models, *Journal of Econometrics,* 41, 205-235.

[51] Zimmer, D.M., 2010, The role of copulas in the housing crisis, *Review of Economics and Statistics*, forthcoming.

Table 1: Simulation results for iid data

	Normal MLE	Normal GMM	Normal SMM	$t\,(4)$-Normal		*Skew* $t\,(4,-0.5)$-Normal		
	σ_z^2	σ_z^2	σ_z^2	σ_z^2	ν_z^{-1}	σ_z^2	ν_z^{-1}	λ_z
True value	1.00	1.00	1.00	1.00	0.25	1.00	0.25	-0.50
				$N=3$				
Bias	0.0170	-0.0120	-0.0128	0.0418	-0.0052	0.0790	-0.0044	-0.0191
Std	0.0808	0.0997	0.1028	0.2220	0.0694	0.3167	0.0679	0.1219
Median	1.0125	0.9863	0.9952	1.0097	0.2527	1.0055	0.2411	-0.5108
90%	1.1184	1.1137	1.1158	1.2600	0.3177	1.3520	0.3327	-0.3713
10%	0.9196	0.8615	0.8502	0.8345	0.1611	0.8437	0.1779	-0.6886
90-10 Diff	0.1988	0.2521	0.2657	0.4255	0.1566	0.5083	0.1548	0.3173
				$N=10$				
Bias	0.0151	-0.0044	-0.0055	0.0027	-0.0056	0.0558	0.0012	-0.0052
Std	0.0565	0.0668	0.0687	0.1068	0.0403	0.1969	0.0486	0.0659
Median	1.0154	0.9900	0.9968	0.9999	0.2391	1.0235	0.2469	-0.5030
90%	1.0871	1.0720	1.0724	1.1374	0.2973	1.2919	0.3112	-0.4191
10%	0.9452	0.9158	0.8967	0.8828	0.1947	0.8826	0.1943	-0.5935
90-10 Diff	0.1419	0.1563	0.1757	0.2546	0.1026	0.4093	0.1169	0.1744
				$N=100$				
Bias	0.0183	-0.0033	-0.0046	-0.0002	-0.0045	0.0419	-0.0014	-0.0033
Std	0.0508	0.0566	0.0560	0.0987	0.0346	0.1604	0.0407	0.0531
Median	1.0150	0.9962	0.9970	0.9893	0.2450	1.0225	0.2496	-0.5058
90%	1.0875	1.0693	1.0725	1.1187	0.2911	1.1983	0.2964	-0.4370
10%	0.9560	0.9268	0.9314	0.8936	0.1985	0.8872	0.2006	-0.5656
90-10 Diff	0.1315	0.1425	0.1411	0.2251	0.0927	0.3111	0.0958	0.1287

Notes: This table presents the results from 100 simulations of three different factor copulas, the Normal copula, the t(4)-Normal factor copula and the Skew t$(4,-0.5)$-Normal factor copula. The Normal copula is estimated by ML, GMM and SMM, and the other two copulas are estimated by SMM. The marginal distributions of the data are assumed to be *iid*. These copulas are considered for problems of dimension $N = 3$, 10 and 100, and in all cases the sample size is $T = 1000$ and the number of simulations used is $S = 25 \times T$. The first row of each panel presents the average difference between the estimated parameter and its true value. The second row presents the standard deviation in the estimated parameters. The third, fourth and fifth rows present the 50^{th}, 90^{th} and 10^{th} percentiles of the distribution of estimated parameters, and the final row presents the difference between the 90^{th} and 10^{th} percentiles.

Table 2: Simulation results for AR-GARCH data

	Normal MLE	Normal GMM	Normal SMM	$t\,(4)$-Normal		*Skew* $t\,(4,-0.5)$-Normal		
	σ_z^2	σ_z^2	σ_z^2	σ_z^2	ν_z^{-1}	σ_z^2	ν_z^{-1}	λ_z
True value	1.00	1.00	1.00	1.00	0.25	1.00	0.25	-0.50
				$N=3$				
Bias	0.0141	-0.0143	-0.0164	0.0504	-0.0048	0.0607	-0.0077	-0.0201
Std	0.0803	0.1014	0.1033	0.2640	0.0715	0.3066	0.0675	0.1206
Median	1.0095	0.9880	0.9949	1.0023	0.2500	1.0021	0.2392	-0.5123
90%	1.1180	1.1103	1.1062	1.2450	0.3244	1.3369	0.3383	-0.3934
10%	0.9172	0.8552	0.8434	0.8413	0.1609	0.8363	0.1631	-0.6616
90-10 Diff	0.2008	0.2551	0.2628	0.4037	0.1635	0.5005	0.1751	0.2682
				$N=10$				
Bias	0.0113	-0.0099	-0.0119	0.0007	-0.0102	0.0403	-0.0027	-0.0056
Std	0.0559	0.0651	0.0666	0.1162	0.0446	0.1804	0.0488	0.0658
Median	1.0125	0.9874	0.9898	0.9928	0.2399	1.0108	0.2469	-0.5054
90%	1.0789	1.0644	1.0706	1.1514	0.2923	1.2351	0.3135	-0.4212
10%	0.9406	0.9027	0.8946	0.8834	0.1892	0.8676	0.1908	-0.5874
90-10 Diff	0.1383	0.1617	0.1761	0.2680	0.1031	0.3676	0.1228	0.1662
				$N=100$				
Bias	0.0167	-0.0068	-0.0080	-0.0031	-0.0101	0.0185	-0.0053	-0.0001
Std	0.0500	0.0554	0.0546	0.1078	0.0390	0.1411	0.0376	0.0535
Median	1.0164	0.9912	0.9956	0.9945	0.2409	1.0093	0.2449	-0.4997
90%	1.0805	1.0625	1.0696	1.1073	0.2880	1.1595	0.2894	-0.4325
10%	0.9534	0.9235	0.9279	0.8921	0.1904	0.8848	0.1992	-0.5623
90-10 Diff	0.1270	0.1390	0.1418	0.2153	0.0975	0.2747	0.0902	0.1298

Notes: This table presents the results from 100 simulations of three different factor copulas, the Normal copula, the t(4)-Normal factor copula and the Skew $t(4,-0.5)$-Normal factor copula. The Normal copula is estimated by ML, GMM, and SMM, and the other two copulas are estimated by SMM. The marginal distributions of the data are assumed to follow AR(1)-GARCH(1,1) processes, as described in Section 4. Problems of dimension $N = 3$, 10 and 100 are considered, the sample size is $T = 1000$ and the number of simulations used is $S = 25 \times T$. The first row of each panel presents the average difference between the estimated parameter and its true value. The second row presents the standard deviation in the estimated parameters. The third, fourth and fifth rows present the 50^{th}, 90^{th} and 10^{th} percentiles of the distribution of estimated parameters, and the final row presents the difference between the 90^{th} and 10^{th} percentiles.

Table 3: Simulation results for different weights factor copula model with N=100

	ν_z^{-1}	λ_z	β_1	β_2	β_3	β_4	β_5	β_6	β_7	β_8	β_9	β_{10}
True value	0.25	-0.5	0.25	0.5	0.75	1	1.25	1.5	1.75	2	2.25	2.5
						Normal						
Bias	-	-	-0.0014	-0.0032	-0.0041	-0.0065	-0.0079	-0.0115	-0.0157	-0.0095	-0.0145	-0.0180
Std	-	-	0.0128	0.0191	0.0253	0.0327	0.0392	0.0503	0.0689	0.0846	0.1046	0.1276
Median	-	-	0.2489	0.4967	0.7429	0.9956	1.2420	1.4849	1.7274	1.9893	2.2225	2.4891
90%	-	-	0.2648	0.5193	0.7816	1.0317	1.2974	1.5569	1.8179	2.0957	2.3617	2.6325
10%	-	-	0.2303	0.4710	0.7165	0.9500	1.1958	1.4266	1.6482	1.8852	2.0996	2.3244
90-10 diff	-	-	0.0345	0.0483	0.0651	0.0818	0.1016	0.1303	0.1697	0.2106	0.2621	0.3081
						t(4)-Normal						
Bias	-0.0071	-	0.0006	0.0015	0.0043	0.0008	0.0021	0.0011	-0.0019	0.0025	-0.0053	0.0010
Std	0.0419	-	0.0234	0.0470	0.0688	0.0881	0.1203	0.1538	0.1830	0.2172	0.2627	0.3543
Median	0.2415	-	0.2496	0.5003	0.7455	0.9865	1.2376	1.4801	1.7200	1.9824	2.2214	2.4582
90%	0.2879	-	0.2689	0.5341	0.7974	1.0634	1.3231	1.5825	1.8605	2.1269	2.4095	2.6790
10%	0.1967	-	0.2271	0.4648	0.7008	0.9332	1.1646	1.3961	1.6302	1.8657	2.0608	2.2919
90-10 diff	0.0913	-	0.0419	0.0693	0.0966	0.1302	0.1585	0.1864	0.2302	0.2612	0.3487	0.3871
						Skew t(4,-0.5)-Normal						
Bias	-0.0123	0.0030	0.0003	-0.0021	-0.0032	-0.0104	-0.0084	-0.0166	-0.0224	-0.0263	-0.0319	-0.0406
Std	0.0420	0.0439	0.0209	0.0365	0.0497	0.0668	0.0919	0.1094	0.1254	0.1505	0.1782	0.1936
Median	0.2402	-0.4993	0.2483	0.4897	0.7384	0.9785	1.2262	1.4674	1.7105	1.9591	2.1938	2.4394
90%	0.2958	-0.4412	0.2739	0.5465	0.8045	1.0892	1.3631	1.6260	1.9128	2.1543	2.4671	2.6959
10%	0.1809	-0.5484	0.2237	0.4552	0.6915	0.9054	1.1309	1.3490	1.5822	1.7916	2.0024	2.2518
90-10 diff	0.1149	0.1072	0.0502	0.0912	0.1130	0.1838	0.2322	0.2770	0.3306	0.3628	0.4646	0.4440

Notes: This table presents the results from 100 simulations of three different factor copulas: the Normal copula, the t(4)-Normal factor copula and the Skew t$(4, -0.5)$-Normal factor copula. We equally divide the $N = 100$ variables into ten groups and assume that all variables in the same group have the same loading on the common factor. The marginal distributions of the data are assumed to follow AR(1)-GARCH(1,1) processes, as described in Section 4. The sample size is $T = 1000$ and the number of simulations used is $S = 25 \times T$. The first row of each panel presents the average difference between the estimated parameter and its true value. The second row presents the standard deviation in the estimated parameters. The third, fourth and fifth rows present the 50^{th}, 90^{th} and 10^{th} percentiles of the distribution of estimated parameters, and the final row presents the difference between the 90^{th} and 10^{th} percentiles.

Table 4: Simulation results on coverage rates

	Normal	t(4)-Normal		Skew t(4,–0.5)-Normal			Normal	t(4)-Normal		Skew t(4,–0.5)-Normal		
	σ_z^2	σ_z^2	ν_z^{-1}	σ_z^2	ν_z^{-1}	λ_z	σ_z^2	σ_z^2	ν_z^{-1}	σ_z^2	ν_z^{-1}	λ_z
	iid data, $N = 3$						AR-GARCH data, $N = 3$					
ε_T												
0.1	91	93	93	96	98	95	89	94	94	95	95	95
0.03	91	93	93	94	94	94	90	92	93	95	94	94
0.01	91	93	93	95	93	94	88	92	92	94	96	95
0.003	87	87	89	93	88	93	85	90	88	92	90	93
0.001	83	77	87	81	87	93	83	76	85	80	89	90
0.0003	64	51	75	64	81	67	58	58	72	59	73	80
0.0001	39	39	63	43	67	45	38	38	54	43	64	55
	iid data, $N = 10$						AR-GARCH data, $N = 10$					
ε_T												
0.1	89	93	97	96	96	98	87	91	96	96	98	97
0.03	88	93	97	96	95	98	87	91	95	96	95	97
0.01	87	93	96	96	94	97	87	91	95	96	94	97
0.003	87	93	95	95	93	93	87	91	95	95	92	95
0.001	87	91	94	91	88	96	87	93	93	93	90	95
0.0003	86	87	90	81	82	88	86	84	88	85	86	92
0.0001	74	72	82	67	74	86	71	70	82	75	78	88
	iid data, $N = 100$						AR-GARCH data, $N = 100$					
ε_T												
0.1	95	97	96	97	94	87	95	96	93	97	95	89
0.03	95	96	94	97	94	90	95	95	92	95	95	90
0.01	95	96	95	97	94	90	95	94	92	95	94	91
0.003	95	96	95	96	94	90	94	94	94	94	94	91
0.001	94	96	91	93	91	91	94	94	91	94	91	92
0.0003	90	94	90	93	88	93	92	95	91	93	89	93
0.0001	87	91	90	88	84	92	84	88	86	91	86	94

Notes: This table presents the results from 100 simulations of three different factor copulas, the Normal copula, the t(4)-Normal factor copula and the Skew t$(4, -0.5)$-Normal factor copula, all estimated by SMM. The marginal distributions of the data are assumed to either be *iid* (left panels) or to follow AR(1)-GARCH(1,1) processes, as described in Section 4 (right panels). Problems of dimension $N = 3$, 10 and 100 are considered, the sample size is $T = 1000$ and the number of simulations used is $S = 25 \times T$. The rows of each panel contain the step size, ε_T, used in computing the matrix of numerical derivatives, $\hat{G}_{T,S}$. The numbers in the table present the percentage of simulations for which the 95% confidence interval based on the estimated covariance matrix contained the true parameter.

Table 5: Coverage rate for different weights factor copula model with N=100 AR-GARCH data

	ν_z^{-1}	λ_z	β_1	β_2	β_3	β_4	β_5	β_6	β_7	β_8	β_9	β_{10}
	Normal											
ε_T												
0.1	-	-	96	91	92	91	95	95	92	93	95	91
0.03	-	-	96	91	92	92	95	95	92	93	95	91
0.01	-	-	97	91	92	92	95	94	92	92	94	91
0.003	-	-	97	91	91	92	95	95	91	92	94	91
0.001	-	-	96	90	91	92	95	96	91	93	92	93
0.0003	-	-	97	90	95	94	96	95	91	93	93	93
0.0001	-	-	95	93	96	93	90	93	89	93	87	87
	t(4)-Normal											
ε_T												
0.1	94	-	94	94	98	96	93	93	95	94	91	90
0.03	93	-	94	94	98	96	93	93	95	94	92	92
0.01	92	-	93	93	98	96	93	93	95	94	92	91
0.003	92	-	92	92	97	95	93	93	93	93	91	92
0.001	91	-	93	93	95	95	93	93	94	93	93	91
0.0003	89	-	96	91	96	94	93	92	93	94	89	91
0.0001	84	-	96	92	88	89	88	91	89	89	93	82
	Skew t(4,-0.5)-Normal											
ε_T												
0.1	96	94	95	95	98	94	92	88	93	88	91	92
0.03	94	93	94	95	97	93	91	87	91	88	90	92
0.01	93	93	94	95	97	92	91	87	93	88	91	92
0.003	93	94	94	93	96	93	91	87	90	89	92	92
0.001	94	93	94	95	97	92	90	87	88	89	90	89
0.0003	88	89	96	95	98	90	87	84	84	82	82	84
0.0001	75	86	93	91	94	84	78	73	78	70	72	75

Notes: This table presents the results from 100 simulations of three different factor copulas: the Normal copula, the t(4)-Normal factor copula and the Skew $t(4, -0.5)$-Normal factor copula. We equally divide the $N = 100$ variables into ten groups and assume that all variables in the same group have the same loading on the common factor. The marginal distributions of the data are assumed to follow AR(1)-GARCH(1,1) processes, as described in Section 4. The sample size is $T = 1000$ and the number of simulations used is $S = 25 \times T$. The rows of each panel contain the step size, ε_T, used in computing the matrix of numerical derivatives, $\hat{G}_{T,S}$. The numbers in the table present the percentage of simulations for which the 95% confidence interval based on the estimated covariance matrix contained the true parameter.

Table 6: Rejection frequencies for the test of overidentifying restrictions

	Normal	t(4)-Normal	Skew t(4,-0.5)-Normal	Normal	t(4)-Normal	Skew t(4,–0.5)-Normal
	Equidependence, N=3			Different loadings, N=3		
90%	95	90	93	92	92	96
95%	97	96	98	95	96	99
99%	100	100	100	99	100	99
	Equidependence, N=10			Different loadings, N=10		
90%	93	89	90	96	93	96
95%	98	94	95	98	96	98
99%	99	98	99	99	99	99
	Equidependence, N=100			Different loadings, N=100		
90%	92	90	96	92	85	94
95%	96	92	97	96	92	95
99%	100	98	100	100	98	98

Notes: This table presents the results from 100 simulations of three different factor copulas, the Normal copula, the t(4)-Normal factor copula and the Skew $t(4, -0.5)$-Normal factor copula, all estimated by SMM. The marginal distributions of the data are assumed to follow AR(1)-GARCH(1,1) processes, as described in Section 4. Problems of dimension $N = 3$, 10 and 100 are considered, the sample size is $T = 1000$ and the number of simulations used is $S = 25 \times T$. The step size used to compute the numerical derivative matrix, $\hat{G}$, needed for the critical value, is 0.1. The rows of each panel contain the confidence level for the test of over-identifying restrictions (0.9, 0.95 or 0.99), and the numbers in the table present the percentage of simulations for which the test statistic was smaller than its computed critical value.

Table 7: **Stocks used in the empirical analysis**

Ticker	Name	SIC	Ticker	Name	SIC	Ticker	Name	SIC
AA	Alcoa	333	EXC	Exelon	493	NKE	Nike	302
AAPL	Apple	357	F	Ford	371	NOV	National Oilwell	353
ABT	Abbott Lab.	283	FCX	Freeport	104	NSC	Norfolk Sth	671
AEP	American Elec	491	FDX	Fedex	451	NWSA	News Corp	271
ALL	Allstate Corp	633	GD	GeneralDynam	373	NYX	NYSE Euronxt	623
AMGN	Amgen Inc.	283	GE	General Elec	351	ORCL	Oracle	737
AMZN	Amazon.com	737	GILD	GileadScience	283	OXY	OccidentalPetrol	131
AVP	Avon	284	GOOG	Google Inc	737	PEP	Pepsi	208
AXP	American Ex	671	GS	GoldmanSachs	621	PFE	Pfizer	283
BA	Boeing	372	HAL	Halliburton	138	PG	Procter&Gamble	284
BAC	Bank of Am	602	HD	Home Depot	525	QCOM	Qualcomm Inc	366
BAX	Baxter	384	HNZ	Heinz	203	RF	Regions Fin	602
BHI	Baker Hughes	138	HON	Honeywell	372	RTN	Raytheon	381
BK	Bank of NY	602	HPQ	HP	357	S	Sprint	481
BMY	Bristol-Myers	283	IBM	IBM	357	SLB	Schlumberger	138
BRK	Berkshire Hath	633	INTC	Intel	367	SLE	Sara Lee Corp.	203
C	Citi Group	602	JNJ	Johnson&J.	283	SO	Southern Co.	491
CAT	Caterpillar	353	JPM	JP Morgan	672	T	AT&T	481
CL	Colgate	284	KFT	Kraft	209	TGT	Target	533
CMCSA	Comcast	484	KO	Coca Cola	208	TWX	Time Warner	737
COF	Capital One	614	LMT	Lock'dMartn	376	TXN	Texas Inst	367
COP	Conocophillips	291	LOW	Lowe's	521	UNH	UnitedHealth	632
COST	Costco	533	MA	Master card	615	UPS	United Parcel	451
CPB	Campbell	203	MCD	MaDonald	581	USB	US Bancorp	602
CSCO	Cisco	367	MDT	Medtronic	384	UTX	United Tech	372
CVS	CVS	591	MET	Metlife Inc.	671	VZ	Verizon	481
CVX	Chevron	291	MMM	3M	384	WAG	Walgreen	591
DD	DuPont	289	MO	Altria Group	211	WFC	Wells Fargo	602
DELL	Dell	357	PM	Philip Morris	211	WMB	Williams	492
DIS	Walt Disney	799	MON	Monsanto	287	WMT	WalMart	533
DOW	Dow Chem	282	MRK	Merck	283	WY	Weyerhauser	241
DVN	Devon Energy	131	MS	MorganStanley	671	XOM	Exxon	291
EMC	EMC	357	MSFT	Microsoft	737	XRX	Xerox	357
ETR	ENTERGY	491						

Notes: This table presents the ticker symbols, names and 3-digit SIC codes of the 100 stocks used in the empirical analysis of this paper.

Table 8: Summary statistics

	Cross-sectional distribution					
	Mean	5%	25%	Median	75%	95%
Mean	0.0004	-0.0003	0.0001	0.0003	0.0006	0.0013
Std dev	0.0287	0.0153	0.0203	0.0250	0.0341	0.0532
Skewness	0.3458	-0.4496	-0.0206	0.3382	0.6841	1.2389
Kurtosis	11.3839	5.9073	7.5957	9.1653	11.4489	19.5939
ϕ_0	0.0004	-0.0007	0.0000	0.0003	0.0007	0.0016
ϕ_1	-0.0439	-0.1206	-0.0717	-0.0468	-0.0135	0.0310
ω	-0.1421	-0.2949	-0.1608	-0.1061	-0.0669	-0.0296
β	0.9812	0.9622	0.9779	0.9855	0.9905	0.9957
α	0.1642	0.0298	0.1203	0.1580	0.2067	0.3148
δ	-0.0828	-0.1536	-0.1051	-0.0833	-0.0588	-0.0291
ρ	0.4238	0.2760	0.3528	0.4160	0.4805	0.6036
ρ_s	0.4418	0.2967	0.3742	0.4330	0.5000	0.6171
$(\tau_{0.99} + \tau_{0.01})/2$	0.0678	0.0000	0.0000	0.0718	0.0718	0.2155
$(\tau_{0.90} - \tau_{0.10})$	-0.0839	-0.1868	-0.1293	-0.0862	-0.0431	0.0287

Notes: This table presents some summary statistics of the daily equity returns data used in the empirical analysis. The top panel presents simple unconditional moments of the daily return series. The second panel presents summaries of the estimated AR(1)-EGARCH(1,1) models estimated on these returns. The lower panel presents linear correlation, rank correlation, average 1% upper and lower tail dependence, and the difference between the 10% tail dependence measures, computed using the standardized residuals from the estimated AR-EGARCH model. The columns present the mean and quantiles from the cross-sectional distribution of the measures listed in the rows. The top two panels present summaries across the $N = 100$ marginal distributions, while the lower panel presents a summary across the $N(N-1)/2 = 4950$ distinct pairs of stocks.

Table 9: Estimation results for daily returns on S&P 100 stocks

	σ_z^2		ν_z^{-1}		λ_z		Q_{SMM}	p-val
	Est	Std Err	Est	Std Err	Est	Std Err		
	$N = 100$							
Normal	0.9237	0.0582	-	-	-	-	0.0081	0.0000
Student's t	0.8592	0.0550	0.0610	0.0288	-	-	0.0051	0.0025
Clayton†	0.6333	0.0302	-	-	-	-	0.0474	0.0000
t(ν)-Normal	0.8875	0.0528	0.0469	0.0302	-	-	0.0069	0.0005
Skew t(ν)-Normal	0.9158	0.0555	0.0420	0.0343	-0.1978	0.0631	0.0008	0.0004
t(ν)-t(ν)	0.8666	0.0557	0.0772	0.0471	-	-	0.0072	0.0004
Skew t(ν)-t(ν)	0.8904	0.0572	0.0827	0.0522	-0.1813	0.0577	0.0007	0.0008

Notes: This table presents estimation results for various copula models applied to 100 daily stock returns over the period April 2008 to December 2010. Estimates and asymptotic standard errors for the copula model parameters are presented, as well as the value of the SMM objective function at the estimated parameters and the p-value of the overidentifying restriction test. †Note that the parameter of the Clayton copula is not σ_z^2 but we report it in this column for simplicity.

Table 10: Seven groups according to the first digit of SIC code

First digit of SIC code	Industry	Number of Stocks
1	Mining, Construction, etc	6
2	Manufacturing: Food, Tobacco, Apprel, Furniture, etc	26
3	Manufacturing: Electronic, Machinery, Metal, etc	25
4	Transportation, Communications, Gas, etc	11
5	Wholesale and Retail Trade	8
6	Finance, Insurance, Real Estate, etc	18
7	Services	6
Total		100

Notes: The Standard Industrial Classification (SIC) is a United States government system for classifying industries. We use first digit of the SIC to classify our 100 stocks to seven industries.

Table 11: Estimation results for daily returns on S&P 100 stocks using different weights factor copulas

	Normal		Student's t		t(ν)-Normal		Skew t(ν)-Normal		t(ν)-t(ν)		Skew t(ν)-t(ν)	
	Est	Std Err	Est	Std Err	Est	Std Err	Est	Std Err	Est	Std Err	Est	Std Err
	N=100											
ν_z^{-1}	-	-	0.0731	0.0331	0.0366	0.0358	0.0533	0.0352	0.0712	0.0518	0.1029	0.0581
λ_z	-	-	-	-	-	-	-0.2078	0.0619	-	-	-0.1906	0.0567
β_1	1.0195	0.0543	0.9866	0.0523	1.0232	0.0527	1.0050	0.0521	1.0229	0.0541	1.0006	0.0539
β_2	0.8950	0.0372	0.8213	0.0369	0.8736	0.0354	0.8792	0.0366	0.8602	0.0358	0.8587	0.0368
β_3	0.9708	0.0339	0.9831	0.0334	0.9868	0.0323	1.0125	0.0333	0.9983	0.0349	0.9984	0.0352
β_4	0.9244	0.0388	0.8371	0.0358	0.9125	0.0373	0.9108	0.0378	0.9031	0.0387	0.9040	0.0389
β_5	0.8842	0.0458	0.8939	0.0449	0.8840	0.0439	0.8971	0.0454	0.8842	0.0432	0.8806	0.0446
β_6	0.9958	0.0368	0.9187	0.0387	0.9984	0.0369	1.0025	0.0380	0.9743	0.0395	0.9769	0.0406
β_7	1.0580	0.0530	1.0009	0.0481	1.0468	0.0490	1.0559	0.0501	1.0551	0.0497	1.0327	0.0499
β p-val	0.0000	-	0.0000	-	0.0000	-	0.0000	-	0.0000	-	0.0000	-
Q_{SMM}	0.0763	-	0.0629	-	0.0581	-	0.0057	-	0.0569	-	0.0047	-
J p-val	0.0001	-	0.0000	-	0.0001	-	0.0170	-	0.0001	-	0.0504	-

Notes: This table presents estimation results for various copula models with seven different weights applied to filtered daily returns on collections of 100 stocks over the period April 2008 to December 2010. Estimates and asymptotic standard errors for the copula model parameters are presented. The third to last row presents the p-value from a test that $\beta_1 = ... = \beta_7$. The bottom two rows contain the value of the objective function at the optimum and the p-value from a test of over-identifying restrictions.

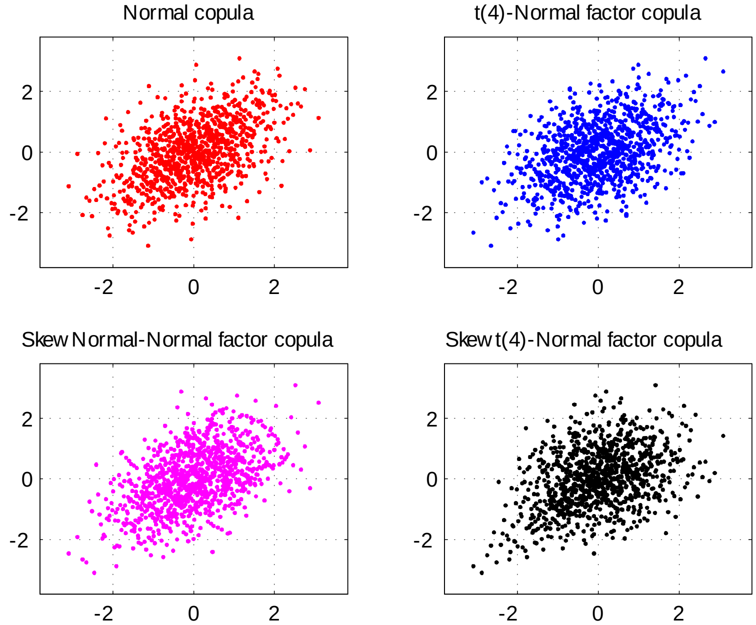


Figure 1: *Scatter plots from four bivariate distributions, all with N(0,1) margins and linear correlation of 0.5, constructed using four different factor copulas.*

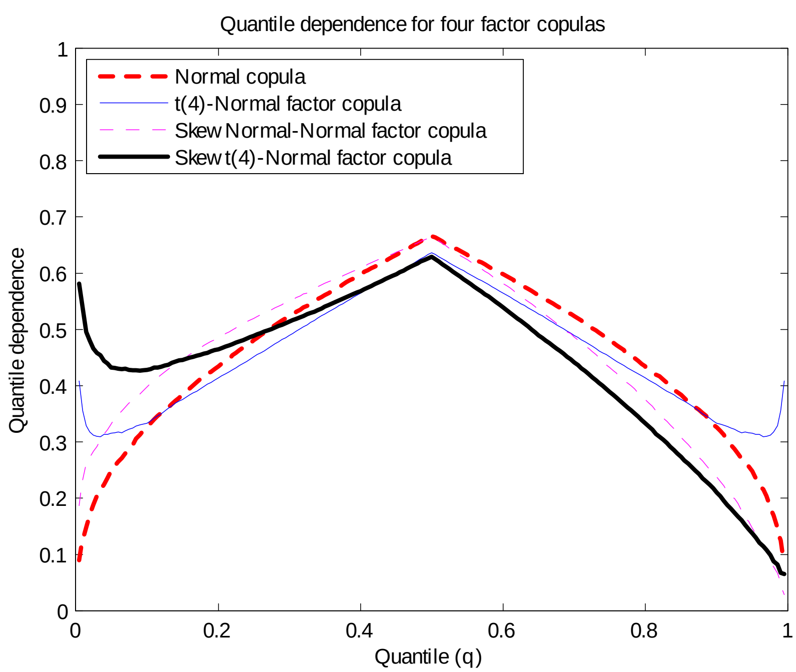
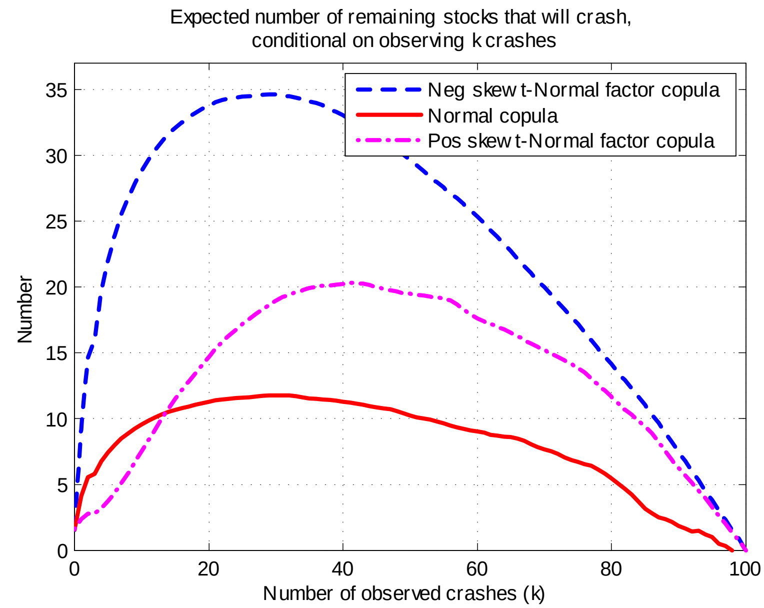


Figure 2: *Quantile dependence implied by four factor copulas, all with linear correlation of 0.5.*



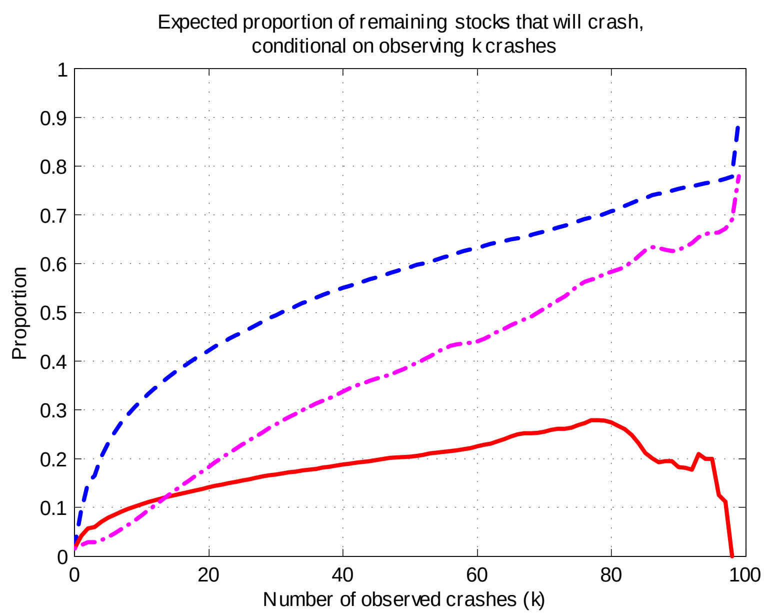


Figure 3: *Conditional on observing k out of 100 stocks crashing, this figure presents the expected number (upper panel) and proportion (lower panel) of the remaining (100-k) stocks that will crash. "Crash" events are defined as returns in the lower 1/66 tail.*

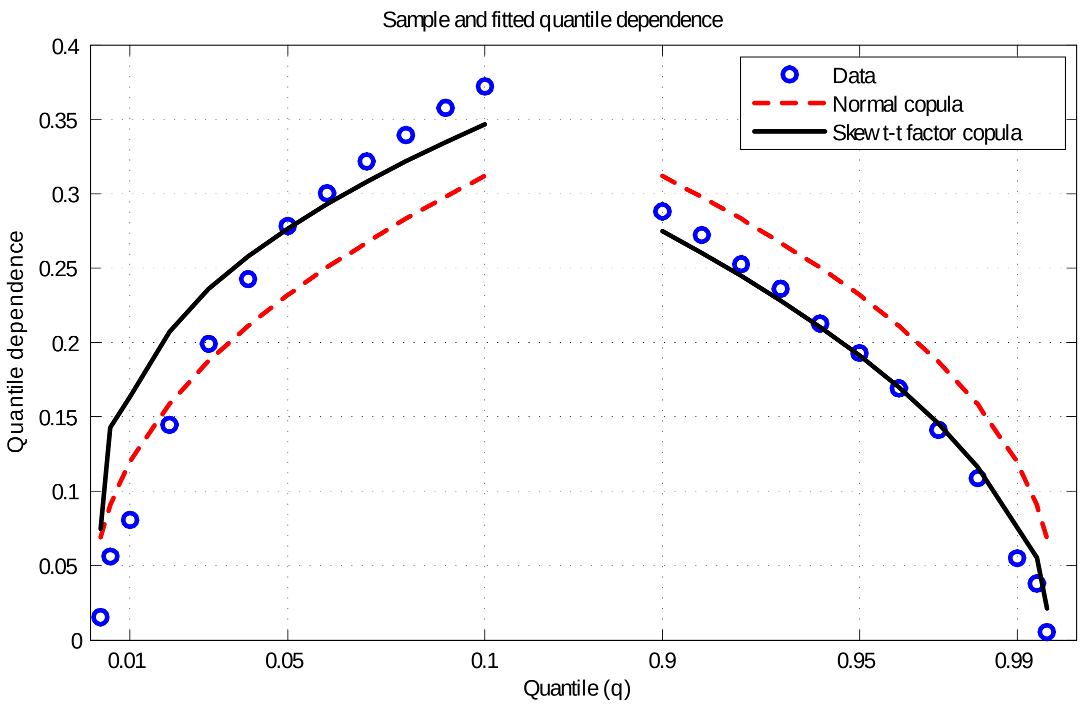


Figure 4: *Sample quantile dependence for 100 daily stock returns, along with the fitted quantile dependence from a Normal copula and from a Skew t-t factor copula, for the lower and upper tails.*

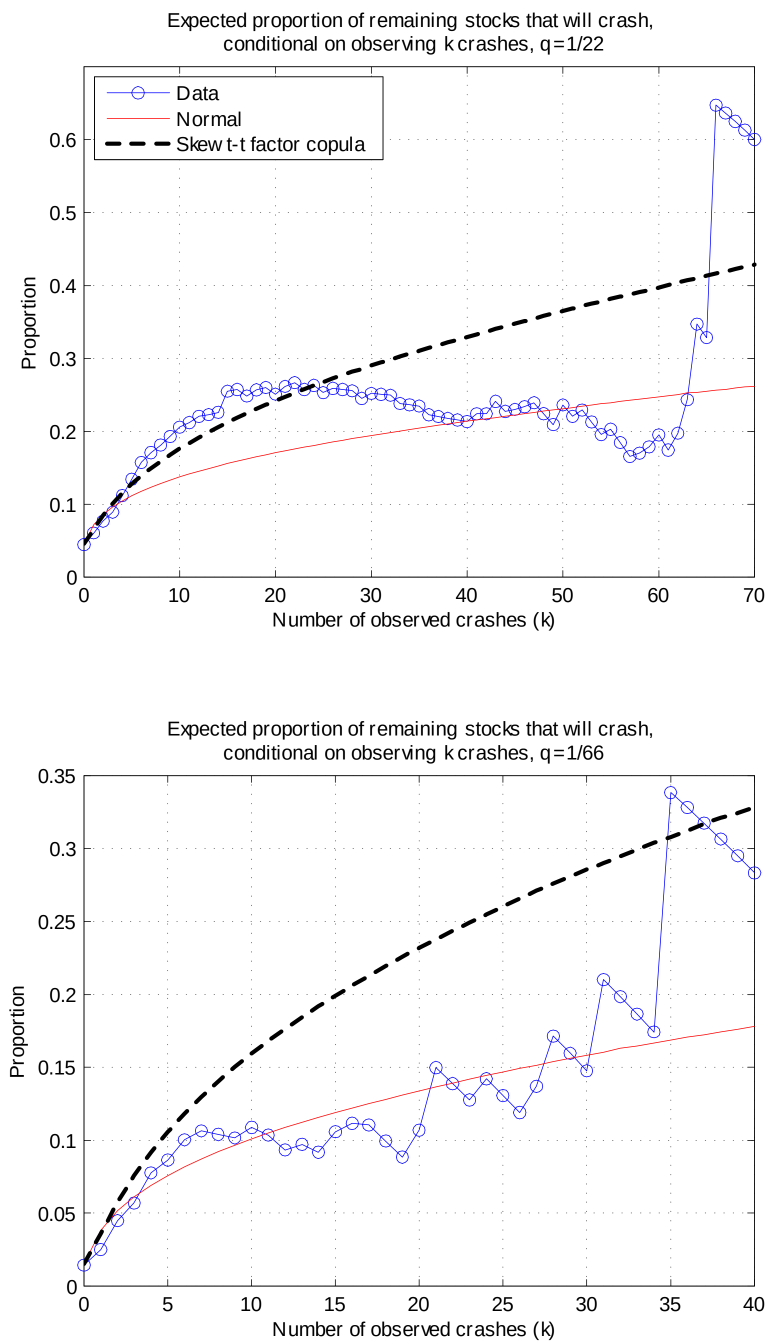


Figure 5: *Conditional on observing k out of 100 stocks crashing, this figure presents the expected proportion of the remaining (100-k) stocks that will crash. "Crash" events are defined as returns in the lower 1/22 (upper panel) and 1/66 (lower panel) tail. Note that the horizontal axes in these two panels are different, due to limited information in the joint tails.*