Simulated Method of Moments Estimation for Copula-Based Multivariate Models*

Dong Hwan Oh and Andrew J. Patton

Duke University

28 November 2011

.

*We thank Tim Bollerslev, George Tauchen, Yanqin Fan and seminar participants at Duke University, Erasmus University Rotterdam, HEC Montreal, McGill University, Monash University, Vanderbilt University, Humboldt-Copenhagen Financial Econometrics workshop, Econometric Society Australiasian Meetings and Econometric Society Asian Meetings 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.

Simulated Method of Moments Estimation for Copula-Based Multivariate Models

Abstract

This paper considers the estimation of the parameters of a copula via a simulated method of moments type approach. This approach is attractive when the likelihood of the copula model is not known in closed form, or when the researcher has a set of dependence measures or other functionals of the copula that are of particular interest. The proposed approach naturally also nests method of moments and generalized method of moments estimators. Drawing on results for simulation based estimation and on recent work in empirical copula process theory, we show the consistency and asymptotic normality of the proposed estimator, and obtain a simple test of over-identifying restrictions as a goodness-of-fit test. The results apply to both *iid* and time series data. We analyze the finite-sample behavior of these estimators in an extensive simulation study. We apply the model to a group of seven financial stock returns and find evidence of statistically significant tail dependence, and mild evidence that the dependence between these assets is stronger in crashes than booms.

Keywords: correlation, dependence, inference, method of moments, SMM

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

1 Introduction

Copula-based models for multivariate distributions are widely used in a variety of applications, including actuarial science and insurance (Embrechts, McNeil and Straumann, 2002; Rosenberg and Schuermann 2006), economics (Brendstrup and Paarsch 2007; Bonhomme and Robin 2009), epidemiology (Clayton 1978; Fine and Jiang 2000), finance (Cherubini, Luciano and Vecchiato 2004; Patton 2006a), geology and hydrology (Cook and Johnson 1981; Genest and Favre 2007), among many others. An important benefit they provide is the flexibility to specify the marginal distributions separately from the dependence structure, without imposing that they come from the same family of joint distributions.

While copulas provide a great deal of flexibility in theory, the search for copula models that work well in practice is an ongoing one. This search has spawned a number of new and flexible models, see Demarta and McNeil (2005), McNeil, Frey and Embrechts (2005), Smith, Min, Almeida and Czado (2010), Smith, Gan and Kohn (2011), and Oh and Patton (2011), among others. Some of these models are such that the likelihood of the copula is either not known in closed form, or is complicated to obtain and maximize, motivating the consideration of estimation methods other than MLE. Moreover, in many financial applications, the estimated copula model is used in pricing a derivative security, such as a collateralized debt obligation or a credit default swap (CDO or CDS), and it may be of interest to minimize the pricing error (the observed market price less the model-implied price of the security) in calibrating the parameters of the model. In some cases the mapping from the parameter(s) of the copula to dependence measures (such as Spearman's or Kendall's rank correlation, for example) or to the price of the derivative contract is known in closed form, thus allowing for method of moments or generalized method of moments (GMM) estimation. In general, however, this mapping is unknown, and an alternative estimation method is required. We consider a simple yet widely applicable simulation-based approach to address this problem.

This paper presents the asymptotic properties of a simulation-based estimator of the parameters of a copula model. We consider both *iid* and time series data, and we consider the case that the marginal distributions are estimated using the empirical distribution function (EDF). The estimation method we consider shares features with the simulated method of moments (SMM), see McFadden (1989) and Pakes and Pollard (1989), for example, however the presence of the EDF in the sample "moments" means that existing results on SMM are not directly applicable. We draw on well-known results on SMM estimators, see Newey and McFadden (1994) for example, and recent results from empirical process theory for copulas, see Fermanian, Radulović and Wegkamp (2004), Chen and Fan (2006) and Rémillard (2010), to show the consistency and asymptotic normality of simulation-based estimators of copula models. To the best of our knowledge, simulation-based estimation of copula models has not previously been considered in the literature. An extensive simulation study verifies that the asymptotic results provide a good approximation in finite samples. We illustrate the results with an application to a model of the dependence between the equity returns on seven financial firms during the recent crisis period.

In addition to maximum likelihood, numerous other estimation methods have been considered for copula-based multivariate models. These include: multi-stage maximum likelihood, also known as "inference functions for margins" in this literature (see Joe and Xu (1996) and Joe (2005) for *iid* data and Patton (2006b) for time series data); semi-parametric maximum likelihood (see Genest, Ghoudi and Rivest (1995) for *iid* data and Chen and Fan (2006) and Chen, Fan and Tsyrennikov (2006) for time series data); method of moments (see Genest (1987) and Genest and Rivest (1993) for *iid* data and Rémillard (2010) for time series data); minimum distance estimation, see Tsukahara (2005); and even "expert judgment" estimation, see Britton, Fisher and Whitley (1998). This paper contributes to this literature by considering the properties of a SMM-type estimator, for both *iid* and time series data, nesting GMM estimation of the copula parameter as a special case.

2 Simulation-based estimation of copula models

We consider the same class of data generating processes (DGPs) as Chen and Fan (2006) 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. As in those papers, and also earlier papers such as Genest and Rivest (1993) and Genest, Ghoudi and Rivest (1995), 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 with unknown parameter $\boldsymbol{\theta}_0$. The DGP we consider is:

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

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, which holds under mild conditions for many commonly-used models for multivariate time series, such as ARMA models, GARCH models, stochastic volatility models, etc. 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 (estimated) standardized residual $\{\hat{\boldsymbol{\eta}}_t \equiv \boldsymbol{\sigma}_t^{-1}(\hat{\boldsymbol{\phi}})[\mathbf{Y}_t - \boldsymbol{\mu}_t(\hat{\boldsymbol{\phi}})\}_{t=1}^{T}$ and simulations from the copula model, $\mathbf{C}(\cdot; \boldsymbol{\theta})$.

2.1 Definition of the SMM estimator

We will consider simulation from some parametric multivariate distribution, $\mathbf{F}_x(\boldsymbol{\theta})$, with marginal distributions $G_i(\boldsymbol{\theta})$, and copula $\mathbf{C}(\boldsymbol{\theta})$. This allows us to consider cases where it is possible to simulate directly from the copula model $\mathbf{C}(\boldsymbol{\theta})$ (in which case the G_i are all $Unif(0,1)$) and also cases where the copula model is embedded in some joint distribution with unknown marginal distributions, such as the factor copula models of Oh and Patton (2011).

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 functions of the marginal distributions (G_i) and contain no information on the copula. Measures like linear correlation contain information on the copula but are also affected by the marginal distributions. Dependence measures like Spearman's rank correlation and quantile dependence are purely functions of the copula and are unaffected by the marginal distributions, see Nelsen (2006) and Joe (1997) for example. Spearman's rank correlation and quantile dependence for the pair (η_i, η_j) are defined as:

$$\rho^{ij} \equiv 12E\left[F_i(\eta_i)F_j(\eta_j)\right] - 3 = 12\int\int uvdC_{ij}(u,v) - 3 \tag{2}$$

$$\tau_q^{ij} \equiv \begin{cases} P\left[F_i(\eta_i) \leq q | F_j(\eta_j) \leq q\right] = \frac{C_{ij}(q,q)}{q}, & q \in (0, 0.5] \\ P\left[F_i(\eta_i) > q | F_j(\eta_j) > q\right] = \frac{1-2q+C_{ij}(q,q)}{1-q}, & q \in (0.5, 1) \end{cases} \tag{3}$$

where C_{ij} is the copula of (η_i, η_j). The sample counterparts based on the estimated standardized residuals are defined as:

$$\hat{\rho}^{ij} \equiv \frac{12}{T}\sum_{t=1}^{T}\hat{F}_i(\hat{\eta}_{it})\hat{F}_j(\hat{\eta}_{jt}) - 3 \tag{4}$$

$$\hat{\tau}_q^{ij} \equiv \begin{cases} \frac{1}{Tq}\sum_{t=1}^{T} 1\{\hat{F}_i(\hat{\eta}_{it}) \leq q, \hat{F}_j(\hat{\eta}_{jt}) \leq q\}, & q \in (0, 0.5] \\ \frac{1}{T(1-q)}\sum_{t=1}^{T} 1\{\hat{F}_i(\hat{\eta}_{it}) > q, \hat{F}_j(\hat{\eta}_{jt}) > q\}, & q \in (0.5, 1) \end{cases} \tag{5}$$

where $\hat{F}_i(y) \equiv (T+1)^{-1} \sum_{t=1}^{T} 1\{\hat{\eta}_{it} \leq y\}$. We will denote the counterparts based on simulated data as $\tilde{\rho}^{ij}(\boldsymbol{\theta})$ and $\tilde{\tau}_q^{ij}(\boldsymbol{\theta})$.

Let $\tilde{\mathbf{m}}_S(\boldsymbol{\theta})$ be a $(m \times 1)$ vector of dependence measures computed using S simulations from $\mathbf{F}_x(\boldsymbol{\theta})$, $\{\mathbf{X}_s\}_{s=1}^{S}$, 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. Define the difference between these as

$$\mathbf{g}_{T,S}(\boldsymbol{\theta}) \equiv \hat{\mathbf{m}}_T - \tilde{\mathbf{m}}_S(\boldsymbol{\theta}) \tag{6}$$

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:

$$\begin{aligned} \hat{\boldsymbol{\theta}}_{T,S} &\equiv \arg\min_{\boldsymbol{\theta} \in \boldsymbol{\Theta}} Q_{T,S}(\boldsymbol{\theta}) \\ \text{where } \ Q_{T,S}(\boldsymbol{\theta}) &\equiv \mathbf{g}'_{T,S}(\boldsymbol{\theta}) \hat{\mathbf{W}}_T \mathbf{g}_{T,S}(\boldsymbol{\theta}) \end{aligned} \tag{7}$$

and $\hat{\mathbf{W}}_T$ 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., $m \geq p$). 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.

2.2 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 $\boldsymbol{\theta}$ since $\tilde{\mathbf{m}}_S(\boldsymbol{\theta})$ will be a number in a set of discrete values as θ varies on Θ, for example, $\left\{0, \frac{1}{Sq}, \frac{2}{Sq}, \ldots, \frac{S}{Sq}\right\}$ for a lower quantile dependence. This problem would vanish if, for the copula model being considered, we knew the mapping $\boldsymbol{\theta} \longmapsto \mathbf{m}_0(\boldsymbol{\theta}) \equiv \lim_{S\to\infty} \tilde{\mathbf{m}}_S(\boldsymbol{\theta})$ 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 (1) 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. First, 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

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

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) $\frac{1}{T}\sum_{t=1}^{T}\boldsymbol{\gamma}_{0t} \xrightarrow{p} \boldsymbol{\Gamma}_0$ *and* $\frac{1}{T}\sum_{t=1}^{T}\boldsymbol{\gamma}_{1kt} \xrightarrow{p} \boldsymbol{\Gamma}_{1k}$ *where* $\boldsymbol{\Gamma}_0$ *and* $\boldsymbol{\Gamma}_{1k}$ *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, sample rank correlation and quantile dependence converge in probability to their population counterparts, see Theorems 3 and 6 of Fermanian, Radulović and Wegkamp (2004) for the *iid* case, and combine with Corollary 1 of Rémillard (2010) for the time series case. (See Lemma 1 of the supplemental appendix for details.) 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:

$$\mathbf{g}_{T,S}(\boldsymbol{\theta}) \equiv \hat{\mathbf{m}}_T - \tilde{\mathbf{m}}_S(\boldsymbol{\theta}) \xrightarrow{p} \mathbf{g}_0(\boldsymbol{\theta}) \equiv \mathbf{m}_0(\boldsymbol{\theta}_0) - \mathbf{m}_0(\boldsymbol{\theta}), \text{ for } \forall \boldsymbol{\theta} \in \boldsymbol{\Theta} \text{ as } T, S \to \infty \quad (8)$$

We define the population objective function as

$$Q_0(\boldsymbol{\theta}) = \mathbf{g}_0(\boldsymbol{\theta})' \mathbf{W}_0 \mathbf{g}_0(\boldsymbol{\theta}) \quad (9)$$

where $\mathbf{W}_0$ is the probability limit of $\hat{\mathbf{W}}_T$. The convergence of $\mathbf{g}_{T,S}(\boldsymbol{\theta})$ and $\hat{\mathbf{W}}_T$ implies that

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

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 a 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 $C_{ij}(u_i, u_j; \boldsymbol{\theta})$ *of* $\mathbf{C}(\boldsymbol{\theta})$ *on* $(u_i, u_j) \in (0,1) \times (0,1)$ *is Lipschitz continuous on* $\boldsymbol{\Theta}$.

(iv) $\hat{\mathbf{W}}_T$ *is* $O_p(1)$ *and converges in probability to* $\mathbf{W}_0$, *a positive definite matrix.*

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

A sketch of all proofs is presented in the Appendix, and detailed proofs are in a supplemental appendix. Assumption 3(iii) is needed to prove the stochastic Lipschitz continuity of $\mathbf{g}_{T,S}(\boldsymbol{\theta})$ which is a sufficient condition for the stochastic equicontinuity of $\mathbf{g}_{T,S}(\boldsymbol{\theta})$. Many bivariate parametric copulas are easily shown to satisfy Assumption 3(iii). While Pakes and Pollard (1989) and McFadden (1989) show the consistency of SMM estimator for T, S diverging at the same rate, Proposition 1 shows that the copula parameter is consistent at any relative rate of T and S as long as both diverge. If we know the function $\mathbf{m}(\boldsymbol{\theta})$ in closed form, then GMM is feasible and is equivalent to our estimator with $S/T \to \infty$ as $T, S \to \infty$.

2.3 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) $\mathbf{g}_0(\boldsymbol{\theta})$ *is differentiable at* $\boldsymbol{\theta}_0$ *with derivative* $\mathbf{G}_0$ *such that* $\mathbf{G}_0'\mathbf{W}_0\mathbf{G}_0$ *is nonsingular.*

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

The first assumption above is standard, and the third assumption is standard in simulation-based estimation problems, see Newey and McFadden (1994) for example. The rate at which the o_p term vanishes in part (iii) turns out to depend on the smaller of T or S, as will become clear from the proposition below. The second 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 $\mathbf{G}_0'\mathbf{W}_0\mathbf{G}_0$ 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 based on three different relative divergence rates of T and S.

Proposition 2 *Suppose that Assumptions 1, 2, 3 and 4 hold. Then*

(i) If $S/T \to \infty$ *as* $T, S \to \infty$,

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

(ii) If $S/T \to k \in (0,\infty)$ as $T,S \to \infty$,

$$\sqrt{T}\left(\hat{\boldsymbol{\theta}}_{T,S} - \boldsymbol{\theta}_0\right) \xrightarrow{d} N\left(0, \left(1+\frac{1}{k}\right)\boldsymbol{\Omega}_0\right) \;\; \text{as } T,S \to \infty \tag{10}$$

(iii) If $S/T \to 0$ as $T,S \to \infty$,

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

where $\boldsymbol{\Omega}_0 = \left(\mathbf{G}_0'\mathbf{W}_0\mathbf{G}_0\right)^{-1}\mathbf{G}_0'\mathbf{W}_0\boldsymbol{\Sigma}_0\mathbf{W}_0\mathbf{G}_0\left(\mathbf{G}_0'\mathbf{W}_0\mathbf{G}_0\right)^{-1}$, *and* $\boldsymbol{\Sigma}_0 \equiv avar\left[\hat{\mathbf{m}}_T\right]$.

When S diverges faster than T, 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 $\boldsymbol{\Omega}_0 = \left(\mathbf{G}_0'\boldsymbol{\Sigma}_0^{-1}\mathbf{G}_0\right)^{-1}$ if $\mathbf{W}_0$ is the efficient weight matrix, $\boldsymbol{\Sigma}_0^{-1}$. However, when $S/T \to k$ as $T,S \to \infty$, the asymptotic variance of SMM estimator $\hat{\boldsymbol{\theta}}_{T,S}$ is $(1+1/k)\,\boldsymbol{\Omega}_0$, which incorporates efficiency loss from simulation error. When $S/T \to 0$ as $T,S \to \infty$, the convergence rate of $\hat{\boldsymbol{\theta}}_{T,S}$ becomes $\sqrt{S}$, and the asymptotic covariance matrix is the same as in the first case. In general, one would like to set S very large to minimize the impact of simulation error, however if the model is computationally costly to simulate, then the third case, where $S \ll T$, may be useful.

Chen and Fan (2006) 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 surprising result also holds for the SMM-type estimators proposed here.

The proof of the above proposition uses recent results for empirical copula processes presented in Fermanian, Radulović and Wegkamp (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, $\mathbf{v}_{T,S}(\boldsymbol{\theta}) = \sqrt{T}\left[\mathbf{g}_{T,S}(\boldsymbol{\theta}) - \mathbf{g}_0(\boldsymbol{\theta})\right]$. These are shown in Lemmas 6 and 7 in the supplemental appendix.

2.4 Consistent estimation of the asymptotic variance

The asymptotic variance of our estimator has the familiar form of standard GMM applications, however the components $\mathbf{\Sigma}_0$ and $\mathbf{G}_0$ require more care in their estimation than in standard applications. We suggest using an *iid* bootstrap to estimate $\mathbf{\Sigma}_0$:

1. 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.

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

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

$$\hat{\mathbf{\Sigma}}_{T,B} = \frac{T}{B}\sum_{b=1}^{B}\left(\hat{\mathbf{m}}_T^{(b)} - \hat{\mathbf{m}}_T\right)\left(\hat{\mathbf{m}}_T^{(b)} - \hat{\mathbf{m}}_T\right)' \tag{11}$$

For the estimation of $\mathbf{G}_0$, 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, Proposition 3 below shows that we need to let the step size go to zero, as usual, but *slower* than the inverse of the rate of convergence of the estimator (i.e., $1/\sqrt{T}$ for the first two cases and $1/\sqrt{S}$ for the third case). Let $\mathbf{e}_k$ denote the k-th unit vector whose dimension is the same as that of $\boldsymbol{\theta}$, and let $\varepsilon_{T,S}$ denote the step size. A two-sided numerical derivative estimator $\hat{\mathbf{G}}_{T,S}$ of $\mathbf{G}$ has k-th column

$$\hat{\mathbf{G}}_{T,S,k} = \frac{\mathbf{g}_{T,S}\left(\hat{\boldsymbol{\theta}}_{T,S}+\mathbf{e}_k\varepsilon_{T,S}\right) - \mathbf{g}_{T,S}\left(\hat{\boldsymbol{\theta}}_{T,S}-\mathbf{e}_k\varepsilon_{T,S}\right)}{2\varepsilon_{T,S}},\ k = 1, 2, ..., p \tag{12}$$

Combine this estimator with $\hat{\mathbf{W}}_T$ to form:

$$\hat{\mathbf{\Omega}}_{T,S,B} = \left(\hat{\mathbf{G}}'_{T,S}\hat{\mathbf{W}}_T\hat{\mathbf{G}}_{T,S}\right)^{-1}\hat{\mathbf{G}}'_{T,S}\hat{\mathbf{W}}_T\hat{\mathbf{\Sigma}}_{T,B}\hat{\mathbf{W}}_T\hat{\mathbf{G}}_{T,S}\left(\hat{\mathbf{G}}'_{T,S}\hat{\mathbf{W}}_T\hat{\mathbf{G}}_{T,S}\right)^{-1} \tag{13}$$

Proposition 3 *Suppose that all assumptions of Proposition 2 are satisfied, and that* $\varepsilon_{T,S} \to 0$*,* $\varepsilon_{T,S} \times \min\left(\sqrt{T}, \sqrt{S}\right) \to \infty$*,* $B \to \infty$ *as* T*,* $S \to \infty$*. Then* $\hat{\mathbf{\Sigma}}_{T,B} \xrightarrow{p} \mathbf{\Sigma}_0$*,* $\hat{\mathbf{G}}_{T,S} \xrightarrow{p} \mathbf{G}_0$ *and* $\hat{\mathbf{\Omega}}_{T,S,B} \xrightarrow{p} \mathbf{\Omega}_0$ *as* $T, S \to \infty$*.*

2.5 A test of overidentifying restrictions

If the number of moments used in estimation is greater than the number of copula parameters, then 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, leading to a non-standard limiting distribution.

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*

$$\begin{aligned} J_{T,S} &\equiv \min(T,S)\, \mathbf{g}_{T,S}\left(\hat{\boldsymbol{\theta}}_{T,S}\right)' \hat{\mathbf{W}}_T \mathbf{g}_{T,S}\left(\hat{\boldsymbol{\theta}}_{T,S}\right) \xrightarrow{d} \mathbf{u}'\mathbf{A}_0'\mathbf{A}_0\mathbf{u} \text{ as } T, S \to \infty \\ \text{where } \mathbf{u} &\sim N(0, \mathbf{I}) \end{aligned}$$

and $\mathbf{A}_0 \equiv \mathbf{W}_0^{1/2}\mathbf{\Sigma}_0^{1/2}\mathbf{R}_0$*,* $\mathbf{R}_0 \equiv \mathbf{I} - \mathbf{\Sigma}_0^{-1/2}\mathbf{G}_0\left(\mathbf{G}_0'\mathbf{W}_0\mathbf{G}_0\right)^{-1}\mathbf{G}_0'\mathbf{W}_0\mathbf{\Sigma}_0^{1/2}$*. If* $\hat{\mathbf{W}}_T = \hat{\mathbf{\Sigma}}_{T,B}^{-1}$*, then* $J_{T,S} \xrightarrow{d} \chi^2_{m-p}$ *as usual.*

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

1. Compute $\hat{\mathbf{R}}$ using $\hat{\mathbf{G}}_{T,S}$, $\hat{\mathbf{W}}_T$, and $\hat{\mathbf{\Sigma}}_{T,B}$.

2. Simulate $\mathbf{u}^{(k)} \sim iid\ N(0, \mathbf{I})$, for $k = 1, 2, ..., K$, where K is large.

3. For each simulation, compute $J_{T,S}^{(k)} = \mathbf{u}^{(k)\prime}\hat{\mathbf{R}}'\hat{\mathbf{\Sigma}}_{T,B}^{1/2\prime}\hat{\mathbf{W}}_T\hat{\mathbf{\Sigma}}_{T,B}^{1/2}\hat{\mathbf{R}}\mathbf{u}^{(k)}$

4. The sample $(1-\alpha)$ quantile of $\left\{J_{T,S}^{(k)}\right\}_{k=1}^{K}$ 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) and 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 $\hat{\mathbf{R}}$ need only be computed once, obtaining critical values for this test is simple and fast.

3 Simulation study

In this section we present a study of the finite sample properties of the simulation-based (SMM) estimator studied in the previous section. We consider two widely-known copula models, the Clayton and the Gaussian (or Normal) copulas, see Nelson (2006) for discussion, and the "factor copula" proposed in Oh and Patton (2011), outlined below. For the first two copulas a closed-form likelihood is available, and we contrast the finite-sample properties of the MLE with the SMM estimator. These two copulas also have closed-form cumulative distribution functions, and so quantile dependence (defined in equation 3) is also known in closed form. For the Clayton copula we have Kendall's rank correlation in closed form $(\rho_{Kendall} = \kappa/(2+\kappa))$ but not Spearman's rank correlation. For the Normal copula we have Spearman's rank correlation in closed form $\left(\rho_{Spearman} = 6/\pi \arcsin(\rho/2)\right)$ but not Kendall's rank correlation. This allows us to also compare GMM with SMM for these copulas, to quantify the loss in accuracy from having to resort to simulations.

The factor copula we consider is based on the following structure:

$$\begin{aligned}
\text{Let} \quad X_i &= Z + \varepsilon_i, \quad i = 1, 2, ..., N \\
\text{where } \ Z &\sim Skew\ t\left(0, \sigma^2, \nu^{-1}, \lambda\right), \ \varepsilon_i \sim iid\ t\left(\nu^{-1}\right), \ \text{ and } \ \varepsilon_i \perp\!\!\!\perp Z \ \forall\ i \qquad (14) \\
[X_1, ..., X_N]' &\equiv \mathbf{X} \sim \mathbf{F}_x = \mathbf{C}\left(G_x, ..., G_x\right)
\end{aligned}$$

where we use the skewed t distribution of Hansen (1994). We use the copula of $\mathbf{X}$ implied by the above structure as our "factor copula" model, and it is parameterized by $(\sigma^2, \nu^{-1}, \lambda)$. For the factor copula we have neither the likelihood nor any of the above dependence measures in closed form, and so simulation-based methods are required. For the simulation we set the parameters to generate rank correlation of around 1/2, and so set the Clayton copula parameter to 1, the Gaussian copula parameter to 1/2, and the factor copula parameters to $\sigma^2 = 1,\ \nu^{-1} = 1/4$ and $\lambda = -1/2$.

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 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, which is widely-used in time series applications:

$$\begin{aligned}
Y_{it} &= \phi_0 + \phi_1 Y_{i,t-1} + \sigma_{it}\eta_{it}, \quad t = 1, 2, ..., T \\
\sigma_{it}^2 &= \omega + \beta\sigma_{i,t-1}^2 + \alpha\sigma_{i,t-1}^2\eta_{i,t-1}^2 \qquad (15) \\
\boldsymbol{\eta}_t &\equiv [\eta_{1t}, ..., \eta_{Nt}] \sim iid \ \ \mathbf{F}_\eta = \mathbf{C}\left(\Phi, \Phi, ..., \Phi\right)
\end{aligned}$$

where Φ is the standard Normal distribution function and $\mathbf{C}$ can be Clayton, Gaussian, or the factor copula implied by equation (14). 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 models for the conditional mean and variance are estimated, and then the estimated standardized residuals are obtained:

$$\hat{\eta}_{it} = \frac{Y_{it} - \hat{\phi}_0 - \hat{\phi}_1 Y_{i,t-1}}{\hat{\sigma}_{it}}. \tag{16}$$

These residuals are used in a second stage to estimate the copula parameters. In all cases we consider a time series of length $T = 1,000,$ 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 use five dependence measures in estimation: Spearman's rank correlation, and the $0.05, 0.10, 0.90, 0.95$ quantile dependence measures, averaged across pairs of assets. We repeat each scenario 100 times, and in the results below we use the identity weight matrix for estimation. (Corresponding results based on the efficient weight matrix are comparable, and available in an online appendix to this paper.)

Table 1 reveals that for all three dimensions ($N = 2,$ 3 and 10) 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. 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 exchangeable nature of all three models.

Comparing the SMM estimator with the ML estimator, which is feasible for the Clayton copula and the Normal copula, we see that the SMM estimators suffer a loss in efficiency of around 40% for $N = 2$ to around 20% for $N = 10$. Some loss is of course expected, and this simulation indicates that the loss is moderate. Comparing the SMM estimator to the GMM estimator provides us with a measure of the loss in accuracy from having to estimate the population moment function via simulation. We find that this loss ranges from zero to 3%, and 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 that somewhat surprising asymptotic result is also relevant in finite samples.

[INSERT TABLES 1 AND 2 ABOUT HERE]

In Table 3 we present the finite-sample coverage probabilities of 95% confidence intervals 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 $\hat{\mathbf{G}}_{T,S}$. This step size, $\varepsilon_{T,S}$, must go to zero, but at a slower rate than $1/\sqrt{T}$. Ignoring constants, our simulation sample size of $T = 1,000$ suggests setting $\varepsilon_{T,S} > 0.001$, which is much larger than standard step sizes used in computing numerical derivatives. (For example, the default in many functions in MATLAB is a step size of around 6×10^{-6}, which is an optimal choice in certain applications, see Judd (1998) for example.) We study the impact of the choice of step size by considering a range of values from 0.0001 to 0.1. Table 3 shows that when the step size is set to 0.01 or 0.1 the finite-sample coverage rates are close to their nominal levels. However if the step size is chosen too small (0.001 or smaller) then the coverage rates are much lower than nominal levels. For example, setting $\varepsilon_{T,S} = 0.0001$ (which is still 16 times larger than the default setting in MATLAB) we find coverage rates as low as 2% 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.

Table 3 also presents 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 $\hat{\mathbf{G}}_{T,S}$, and we again compute $\hat{\mathbf{G}}_{T,S}$ using a step size of $\varepsilon_{T,S} = 0.1, 0.01, 0.001$ and 0.0001. The rejection rates are close to their nominal levels 95% for the all three copula models.

These simulation results provide support for the proposed estimation method: for empirically realistic parameter values and sample size, the estimator is approximately unbiased, with estimated confidence intervals that have coverage close to their nominal level when the step size for the numerical derivative is chosen in line with our theoretical results.

[INSERT TABLE 3 ABOUT HERE]

4 Application to the dependence between financial firms

This section considers models for the dependence between seven large financial firms. We use daily stock return data over the period January 2001 to December 2010, a total of $T = 2515$ trade days, on Bank of America, Bank of New York, Citigroup, Goldman Sachs, J.P. Morgan, Morgan Stanley and Wells Fargo. Summary statistics for these returns are presented in Table S1 of the supplemental appendix, and indicate that all series are positively skewed and leptokurtotic, with kurtosis ranging from 16.0 (J.P. Morgan) to 119.8 (Morgan Stanley).

To capture the impact of time-varying conditional means and variances in each of these series, we estimate the following autoregressive, conditionally heteroskedastic models:

$$
\begin{aligned}
r_{it} &= \phi_{0i} + \phi_{1i} r_{i,t-1} + \phi_{2i} r_{m,t-1} + \varepsilon_{it}, \quad \varepsilon_{it} = \sigma_{it}\eta_{it} \\
\text{where} \quad \sigma_{it}^2 &= \omega_i + \beta_i \sigma_{i,t-1}^2 + \alpha_{1i}\varepsilon_{i,t-1}^2 + \gamma_{1i}\varepsilon_{i,t-1}^2 \cdot \mathbf{1}_{[\varepsilon_{i,t-1} \leq 0]} \\
&\quad + \alpha_{2i}\varepsilon_{m,t-1}^2 + \gamma_{2i}\varepsilon_{m,t-1}^2 \cdot \mathbf{1}_{[\varepsilon_{m,t-1} \leq 0]}
\end{aligned} \tag{17}
$$

where r_{it} is the return on one of these seven firms and r_{mt} is the return on the S&P 500 index. We include the lagged market index return in both the mean and variance specifications to capture any influence of lagged information in the model for a given stock, and in the model for the market index itself we set $\phi_1 = \alpha_1 = \gamma_1 = 0$. Estimated parameters from these models are presented in Table S2 of the supplemental appendix, and are consistent with the values found in the empirical finance literature, see Bollerslev, Engle and Nelson (1994) for example. From these models we obtain the estimated standardized residuals, $\hat{\eta}_{it}$, which are used in the estimation of the dependence structure.

In Table 4 we present measures of dependence between these seven firms. The upper panel reveals that rank correlation between their standardized residuals is 0.63 on average, and ranges from 0.55 to 0.76. The lower panel of Table 4 presents measures of dependence in the tails between these series. The upper triangle presents the average of the 1% and 99% quantile dependence measures presented in equation (5), and we see substantial dependence here, with values ranging between 0.16 and 0.40. The lower triangle presents the difference between the 90% and 10% quantile dependence measures, as a gauge of the degree of asymmetry in the dependence structure. These differences are mostly negative (14 out of 21), indicating greater dependence during crashes than during booms.

Table 5 presents the estimation results for three different copula models of these series. The first model is the well-known Clayton copula, the second is the Normal copula and the third is a "factor copula" as proposed by Oh and Patton (2011). The first copula allows for lower tail dependence, but imposes that upper tail dependence is zero. The second copula implies zero tail dependence in both directions. The third copula allows for tail dependence in both tails, and allows the degree of dependence to differ across positive and negative realizations.

[INSERT TABLES 4 AND 5 ABOUT HERE]

For the first two copulas a closed-form likelihood is available, and so maximum likelihood estimation is possible. For all three copulas we also implement the SMM-type method proposed in Section 2 using five dependence measures: Spearman's rank correlation, and the $0.05, 0.10, 0.90, 0.95$ quantile dependence measures, averaged across pairs of assets. 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 $B = 1,000$ bootstraps used to estimate $\Sigma_{T,S}$, and $\varepsilon_{T,S} = 0.1$ used as the step size to compute $\hat{\mathbf{G}}_{T,S}$.

The parameter estimates for the Normal copula are similar for ML and SMM, while they are quite different for the Clayton copula. This may be explained by the results of the test of over-identifying restrictions: the Clayton copula is strongly rejected, while the Normal copula is less strongly rejected (p-value of 0.043). The factor copula is not rejected using this test. The improvement in fit from the factor copula appears to come from its ability to capture tail dependence: the parameter that governs tail dependence (ν^{-1}) is significantly greater than zero, while the parameter that governs asymmetric dependence (λ) is not significantly different from zero.

Figure 1 sheds some further light on the relative performance of these copula models. This figure compares the empirical quantile dependence function with those implied by the three copula models. An *iid* bootstrap with $B = 1,000$ replications is used to construct pointwise confidence intervals for the sample quantile dependence estimates. We see here that the Clayton copula is "too asymmetric" relative to the data, while both the Normal and the factor copula models appear to provide a reasonable fit.

[INSERT FIGURE 1 ABOUT HERE]

5 Conclusion

This paper presents the asymptotic properties of a new simulation-based estimator of the parameters of a copula model, which matches measures of rank dependence implied by the model to those observed in the data. The estimation method shares features with the simulated method of moments (SMM), see McFadden (1989) and Newey and McFadden (1994), for example, however the use of rank dependence measures as "moments" means that existing results on SMM cannot be used. We extend well-known results on SMM estimators using recent work in empirical process theory for copula estimation, see Fermanian, Radulović and Wegkamp (2004), Chen and Fan (2006) and Rémillard (2010), to show the consistency and asymptotic normality of SMM-type estimators of copula models. To the best of our knowledge, simulation-based estimation of copula models has not previously been considered in the literature. We also provide a method for obtaining a consistent estimate of the asymptotic covariance matrix, and a test of the over-identifying restrictions. Our results apply to both *iid* and time series data, and an extensive simulation study verifies that the asymptotic results provide a good approximation in finite samples. We illustrate the results with an application to a model of the dependence between the equity returns on seven financial firms, and find evidence of statistically significant tail dependence, and some evidence that the dependence between these assets is stronger in crashes than booms.

Appendix: Sketch of proofs

Detailed proofs are available in a supplemental appendix to this paper.

Proof of Proposition 1. First note that: (a) $Q_0(\boldsymbol{\theta})$ is uniquely minimized at $\boldsymbol{\theta}_0$ by Assumption 3(i) and positive definite $\mathbf{W}_0$ of Assumption 3(iv), (b) $\boldsymbol{\Theta}$ is compact by Assumption 3(ii); (c) $Q_0(\boldsymbol{\theta})$ consists of linear combinations of rank correlations and quantile dependence measures that are functions of pair-wise copula functions, so $Q_0(\boldsymbol{\theta})$ is continuous by Assumption 3(iii). The main part of the proof requires establishing that $Q_{T,S}$ uniformly converges in probability to Q_0, which we show using five lemmas in the supplemental appendix: Pointwise convergence of $\mathbf{g}_{T,S}(\boldsymbol{\theta})$ to $\mathbf{g}_0(\boldsymbol{\theta})$ and stochastic Lipschitz continuity of $\mathbf{g}_{T,S}(\boldsymbol{\theta})$ is shown using results from Fermanian, Wegkamp and Radulović (2004) and Rémillard (2010), combined with Assumption 3(iii). This is sufficient for the stochastic equicontinuity of $\mathbf{g}_{T,S}$ and for the uniform convergence in probability of $\mathbf{g}_{T,S}$ to $\mathbf{g}_0$ by Lemmas 2.8 and 2.9 of Newey and McFadden (1994). Using the triangle and Cauchy-Schwarz inequalities this implies that $Q_{T,S}$ uniformly converges in probability to Q_0. We have thus verified that the conditions of Theorem 2.1 of Newey and McFadden (1994) hold, and we have $\hat{\boldsymbol{\theta}} \xrightarrow{p} \boldsymbol{\theta}_0$ as claimed. ■

Proof of Proposition 2. We prove this proposition by verifying the five conditions of Theorem 7.2 of Newey and McFadden (1994) for our problem: (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 $\mathbf{G}_0$ such that $\mathbf{G}_0'\mathbf{W}_0\mathbf{G}_0$ is nonsingular by Assumption 4(ii). (iii) $\boldsymbol{\theta}_0$ is an interior point of $\boldsymbol{\Theta}$ by Assumption 4(i). (iv) This part requires showing the asymptotic normality of $\sqrt{T}\mathbf{g}_{T,S}(\boldsymbol{\theta}_0)$. We will present the result only for $S/T \to k \in (0, \infty)$. The results for the cases that $S/T \to 0$ or $S/T \to \infty$ are similar. In Lemma 6 of the supplemental appendix we show that $\sqrt{T}(\hat{\mathbf{m}}_T - \mathbf{m}_0(\boldsymbol{\theta}_0)) \xrightarrow{d} N(0, \boldsymbol{\Sigma}_0)$ as $T \to \infty$ and $\sqrt{S}(\tilde{\mathbf{m}}_S(\boldsymbol{\theta}_0) - \mathbf{m}_0(\boldsymbol{\theta}_0)) \xrightarrow{d} N(0, \boldsymbol{\Sigma}_0)$ as $S \to \infty$ using Theorem 3 and Theorem 6 of Fermanian, Radulović and Wegkamp (2004) and

Corollary 1, Proposition 2 and Proposition 4 of Rémillard (2010). This implies that

$$\sqrt{T}\mathbf{g}_{T,S}(\boldsymbol{\theta}_0) = \underbrace{\sqrt{T}\left(\hat{\mathbf{m}}_T - \mathbf{m}_0(\boldsymbol{\theta}_0)\right)}_{\xrightarrow{d} N(0,\boldsymbol{\Sigma}_0)} - \underbrace{\sqrt{\frac{T}{S}}}_{\to 1/\sqrt{k}} \underbrace{\sqrt{S}\left(\tilde{\mathbf{m}}_S(\boldsymbol{\theta}_0) - \mathbf{m}_0(\boldsymbol{\theta}_0)\right)}_{\xrightarrow{d} N(0,\boldsymbol{\Sigma}_0)}$$

and so $\sqrt{T}\mathbf{g}_{T,S}(\boldsymbol{\theta}_0) \xrightarrow{d} N(\mathbf{0}, (1+1/k)\boldsymbol{\Sigma}_0)$ as $T, S \to \infty$. (v) This part requires showing that $\sup_{\|\theta-\theta_0\|<\delta} \sqrt{T}\left\|\boldsymbol{g}_{T,S}(\boldsymbol{\theta}) - \boldsymbol{g}_{T,S}(\boldsymbol{\theta}_0) - \boldsymbol{g}_0(\boldsymbol{\theta})\right\| / \left[1 + \sqrt{T}\|\boldsymbol{\theta} - \boldsymbol{\theta}_0\|\right] \xrightarrow{p} 0$. The main part of this proof involves showing the stochastic equicontinuity of $\mathbf{v}_{T,S}(\boldsymbol{\theta}) = \sqrt{T}\left[\boldsymbol{g}_{T,S}(\boldsymbol{\theta}) - \boldsymbol{g}_0(\boldsymbol{\theta})\right]$. This is shown in Lemma 7 of the supplemental appendix by showing that $\{\mathbf{g}_{\cdot,\cdot}(\boldsymbol{\theta}) : \theta \in \Theta\}$ is a type II class of functions in Andrews (1994), and then using that paper's Theorem 1. ■

Proof of Proposition 3. If $\boldsymbol{\mu}_t$ and $\boldsymbol{\sigma}_t$ are known constant, or if $\boldsymbol{\phi}_0$ is known, then the consistency of $\hat{\boldsymbol{\Sigma}}_{T,B}$ follows from Theorems 5 and 6 of Fermanian, Radulović and Wegkamp (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, Radulović and Wegkamp (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

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

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 (11) is a consistent estimator of $\mathbf{\Sigma}_0$. Consistency of the numerical derivatives $\hat{\mathbf{G}}_{T,S}$ can be established using a similar approach to Theorem 7.4 of Newey and McFadden (1994), and since $\hat{\mathbf{W}}_T \xrightarrow{p} \mathbf{W}_0$ by Assumption 3(iv), we thus have $\hat{\mathbf{\Omega}}_{T,S,B} \xrightarrow{p} \mathbf{\Omega}_0$. ■

Proof of Proposition 4. We consider only the case where $S/T \to \infty$ or $S/T \to k > 0$. The case for $k = 0$ is analogous. A Taylor expansion of $\mathbf{g}_0\left(\hat{\boldsymbol{\theta}}_{T,S}\right)$ around $\boldsymbol{\theta}_0$ yields

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

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

$$\sqrt{T}\mathbf{g}_0\left(\hat{\boldsymbol{\theta}}_{T,S}\right) = \mathbf{G}_0 \cdot \sqrt{T}\left(\hat{\boldsymbol{\theta}}_{T,S} - \boldsymbol{\theta}_0\right) + o_p\left(1\right) \tag{18}$$

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

$$\sqrt{T}\mathbf{g}_{T,S}\left(\hat{\boldsymbol{\theta}}_{T,S}\right) = \sqrt{T}\mathbf{g}_{T,S}\left(\boldsymbol{\theta}_0\right) + \hat{\mathbf{G}}_{T,S} \cdot \sqrt{T}\left(\hat{\boldsymbol{\theta}}_{T,S} - \boldsymbol{\theta}_0\right) + \mathbf{R}_{T,S}\left(\hat{\boldsymbol{\theta}}_{T,S}\right) \tag{19}$$

where the remaining term is captured by $\mathbf{R}_{T,S}\left(\hat{\boldsymbol{\theta}}_{T,S}\right)$. Combining equations (18) and (19) we obtain

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

The stochastic equicontinuity of $\mathbf{v}_{T,S}\left(\boldsymbol{\theta}\right) = \sqrt{T}\left[\mathbf{g}_{T,S}\left(\boldsymbol{\theta}\right) - \mathbf{g}_0\left(\boldsymbol{\theta}\right)\right]$ is established in the proof of Proposition 2, which implies (see proof of Proposition 2) that

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

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

$$\sqrt{T}\mathbf{g}_{T,S}\left(\hat{\boldsymbol{\theta}}_{T,S}\right) = \sqrt{T}\mathbf{g}_{T,S}\left(\boldsymbol{\theta}_0\right) + \hat{\mathbf{G}}_{T,S} \cdot \sqrt{T}\left(\hat{\boldsymbol{\theta}}_{T,S} - \boldsymbol{\theta}_0\right) + o_p\left(1\right) \tag{20}$$

The remainder of the proof is the same as in standard GMM applications, see Hall (2005) for example. ■

References

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

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

[3] Bollerslev, T., R.F. Engle, and D.B. Nelson, 1994, ARCH Models, in in R.F. Engle and D. McFadden (eds.), *Handbook of Econometrics*, 4, 2959-3038.

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

[5] Brendstrup, B. and H.J. Paarsch, 2007, Semiparametric Identification and Estimation in Multi-Object English Auctions, *Journal of Econometrics*, 141, 84-108.

[6] Britton, E., P. Fisher, and J. Whitley, 1998, The Inflation Report Projections: Understanding the Fan Chart, *Bank of England Quarterly Bulletin*, 30-37.

[7] Chen, X. and Y. Fan, 2006, Estimation and Model Selection of Semiparametric Copula-based Multivariate Dynamic Models under Copula Misspecification, *Journal of Econometrics*, 135, 125-154.

[8] Chen, X., Y. Fan, and V. Tsyrennikov, 2006, Efficient Estimation of Semiparametric Multivariate Copula Models, *Journal of the American Statistical Association*, 101(475), 1228-1240.

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

[10] Clayton, D.G., 1978, A model for association in bivariate life tables and its application in epidemiological studies of familial tendency in chronic disease incidence, *Biometrika*, 65, 141-151.

[11] Cook, R.D. and M.E Johnson, 1981, A family of multivariate distributions for modelling non-elliptically symmetric multivariate data, *Journal of the Royal Statistical Society – B*, 43(2), 210-218.

[12] Demarta, S. and A.J. McNeil, 2005, The t Copula and Related Copulas, *International Statistical Review*, 73, 111-129.

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

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

[15] Fine, J.P. and H. Jiang, 2000, On association in a copula with time transformations, *Biometrika*, 87(3), 559-571.

[16] Genest, C., 1987, Frank's Family of Bivariate Distributions, *Biometrika*, 74, 549-555.

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

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

[19] Genest, C., K. Ghoudi, and L.-P. Rivest, 1995, A Semiparametric Estimation Procedure of Dependence Parameters in Multivariate Families of Distributions, *Biometrika*, 82(3), 543-552.

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

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

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

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

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

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

[26] McFadden, D., 1989, A Method of Simulated Moments for Estimation of Discrete Response Models without Numerical Integration, *Econometrica*, 47, 995-1026.

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

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

[29] 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.

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

[31] Pakes, A. and D. Pollard, 1989, Simulation and the Asymptotics of Optimization Estimators, *Econometrica*, 47, 1027-1057.

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

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

[34] Rémillard, B., 2010, Goodness-of-fit Tests for Copulas of Multivariate Time Series, working paper.

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

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

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

[38] Smith, M., Q. Gan, and R. Kohn, 2011, Modeling Dependence using Skew t Copulas: Bayesian Inference and Applications, *Journal of Applied Econometrics*, forthcoming.

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

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

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

Table 1: Simulation results for iid data

	Clayton				*Normal*			*Factor copula*		
	MLE	GMM	SMM	SMM*	MLE	GMM	SMM	SMM		
	κ	κ	κ	κ	ρ	ρ	ρ	σ^2	ν^{-1}	λ
True	1.00	1.00	1.00	1.00	0.5	0.5	0.5	1.00	0.25	-0.50
					$N = 2$					
Bias	0.001	-0.014	-0.006	-0.004	0.004	-0.001	-0.001	0.016	-0.012	-0.089
St dev	0.085	0.119	0.122	0.110	0.024	0.034	0.034	0.152	0.123	0.199
Median	1.011	0.982	0.991	0.998	0.503	0.497	0.498	0.985	0.243	-0.548
90-10%	0.216	0.308	0.293	0.294	0.063	0.086	0.087	0.374	0.363	0.540
					$N = 3$					
Bias	0.015	0.008	0.006	0.008	0.003	-0.004	-0.005	0.032	-0.002	-0.057
St dev	0.061	0.090	0.092	0.091	0.020	0.025	0.026	0.124	0.111	0.157
Median	1.013	1.003	0.999	0.998	0.503	0.497	0.499	1.031	0.256	-0.542
90-10%	0.155	0.226	0.219	0.216	0.049	0.064	0.068	0.297	0.293	0.395
					$N = 10$					
Bias	0.008	0.007	0.008	0.004	0.003	-0.002	-0.002	0.026	0.001	-0.011
St dev	0.050	0.068	0.066	0.059	0.014	0.017	0.017	0.093	0.070	0.082
Median	1.005	1.002	1.005	0.999	0.504	0.498	0.499	1.013	0.255	-0.508
90-10%	0.132	0.198	0.177	0.152	0.035	0.039	0.045	0.248	0.186	0.168

Notes: This table presents the results from 100 simulations of Clayton, the Normal copula, and a factor copula. In the SMM and GMM estimation all three copulas use five dependence measures, including four quantile dependence measures ($q = 0.05, 0.10, 0.90, 0, 95$). The Normal and factor copulas also use Spearman's rank correlation, while the Clayton copula uses either Kendall's (GMM and SMM) or Spearman's (SMM*) rank correlation. The marginal distributions of the data are assumed to be *iid* $N(0, 1)$. Problems of dimension $N = 2$, 3 and 10 are considered, the sample size is $T = 1,000$ 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 of the estimated parameters. The third and fourth rows present the median and the difference between the 90^{th} and 10^{th} percentiles of the distribution of estimated parameters.

Table 2: Simulation results for AR-GARCH data

	Clayton				*Normal*			*Factor copula*		
	MLE	GMM	SMM	SMM*	MLE	GMM	SMM	SMM		
	κ	κ	κ	κ	ρ	ρ	ρ	σ^2	ν^{-1}	λ
True	1.00	1.00	1.00	1.00	0.5	0.5	0.5	1.00	0.25	-0.50
					$N = 2$					
Bias	-0.005	-0.029	-0.028	-0.020	0.003	-0.001	-0.001	0.015	-0.012	-0.073
St dev	0.087	0.124	0.124	0.108	0.024	0.035	0.036	0.155	0.121	0.188
Median	0.998	0.977	0.975	0.982	0.503	0.497	0.499	0.995	0.235	-0.558
90-10%	0.228	0.327	0.340	0.267	0.061	0.084	0.090	0.411	0.346	0.509
					$N = 3$					
Bias	0.006	-0.007	0.002	-0.008	0.003	-0.005	-0.006	0.013	-0.020	-0.052
St dev	0.060	0.087	0.088	0.080	0.020	0.026	0.026	0.121	0.106	0.148
Median	1.005	0.991	0.994	0.981	0.502	0.497	0.499	1.005	0.238	-0.521
90-10%	0.145	0.205	0.213	0.195	0.050	0.065	0.068	0.311	0.297	0.357
					$N = 10$					
Bias	-0.004	-0.002	-0.004	-0.003	0.002	-0.003	-0.004	-0.004	-0.016	-0.012
St dev	0.049	0.067	0.064	0.059	0.014	0.016	0.017	0.085	0.079	0.071
Median	0.995	0.996	0.987	0.988	0.503	0.497	0.497	0.990	0.238	-0.508
90-10%	0.134	0.179	0.170	0.152	0.034	0.041	0.045	0.210	0.209	0.165

Notes: This table presents the results from 100 simulations of Clayton, the Normal copula, and a factor copula. In the SMM and GMM estimation all three copulas use five dependence measures, including four quantile dependence measures ($q = 0.05, 0.10, 0.90, 0, 95$). The Normal and factor copulas also use Spearman's rank correlation, while the Clayton copula uses either Kendall's (GMM and SMM) or Spearman's (SMM*) rank correlation. The marginal distributions of the data are assumed to follow AR(1)-GARCH(1,1) processes, as described in Section 3. Problems of dimension $N = 2$, 3 and 10 are considered, the sample size is $T = 1,000$ 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 of the estimated parameters. The third and fourth rows present the median and the difference between the 90^{th} and 10^{th} percentiles of the distribution of estimated parameters.

Table 3: Simulation results on coverage rates

	Clayton		Normal		Factor copula			
	κ	J	ρ	J	σ^2	ν^{-1}	λ	J
				$N=2$				
$\varepsilon_{T,S}$								
0.1	91	98	94	98	94	100	95	98
0.01	46	99	92	98	94	99	96	100
0.001	2	99	76	98	76	79	74	99
0.0001	1	99	21	98	54	75	57	97
				$N=3$				
$\varepsilon_{T,S}$								
0.1	97	99	89	97	99	100	96	99
0.01	63	98	88	97	99	96	95	100
0.001	11	98	83	98	92	84	93	100
0.0001	2	100	38	99	57	70	61	99
				$N=10$				
$\varepsilon_{T,S}$								
0.1	96	99	87	97	97	97	95	98
0.01	88	99	87	96	96	97	97	97
0.001	18	100	87	98	97	95	88	97
0.0001	0	98	71	97	73	85	81	98

Notes: This table presents the results from 100 simulations of Clayton copula, the Normal copula, and a 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 3. Problems of dimension $N = 2$, 3 and 10 are considered, the sample size is $T = 1,000$ and the number of simulations used is $S = 25 \times T$. The rows of each panel contain the step size, $\varepsilon_{T,S}$, used in computing the matrix of numerical derivatives, $\hat{\mathbf{G}}_{T,S}$. The numbers in column $\kappa, \rho, \sigma^2, \nu^{-1}$, and λ present the percentage of simulations for which the 95% confidence interval based on the estimated covariance matrix contained the true parameter. The numbers in column J present the percentage of simulations for which the test statistic of over-identifying restrictions test described in Section 2 was smaller than its computed critical value under 95% confidence level.

Table 4: Sample dependence statistics

	Bank of America	Bank of N.Y.	Citi Group	Goldman Sachs	JP Morgan	Morgan Stanley	Wells Fargo
BoA		0.586	0.691	0.556	0.705	0.602	0.701
BoNY	0.551		0.574	0.578	0.658	0.592	0.595
Citi	0.685	0.558		0.608	0.684	0.649	0.626
Goldman	0.564	0.565	0.609		0.655	0.759	0.548
JPM	0.713	0.633	0.694	0.666		0.667	0.683
Morgan S	0.604	0.587	0.650	0.774	0.676		0.578
Wells F	0.715	0.593	0.636	0.554	0.704	0.587	
BoA		0.219	0.239	0.219	0.398	0.298	0.358
BoNY	-0.048		0.179	0.199	0.159	0.219	0.199
Citi	-0.045	-0.004		0.199	0.318	0.219	0.199
Goldman	-0.068	0.000	0.032		0.239	0.378	0.199
JPM	-0.024	-0.056	-0.012	0.012		0.239	0.358
Morgan S	-0.060	-0.020	-0.064	-0.036	-0.008		0.219
Wells F	0.020	-0.052	0.044	-0.028	0.024	0.000	

Notes: This table presents measures of dependence between the seven financial firms under analysis. The upper panel presents Spearman's rank correlation (upper triangle) and linear correlation (lower triangle), and the lower panel presents the difference between the 10% tail dependence measures (lower triangle) and average 1% upper and lower tail dependence (upper triangle). All dependence measures are computed using the standardized residuals from the models for the conditional mean and variance.

Table 5: Estimation results for daily returns on seven stocks

	Clayton		Normal		Factor
	MLE	SMM	MLE	SMM	SMM
	κ	κ	ρ	ρ	$\sigma^2, \nu^{-1}, \lambda$
Estimate	0.907	1.274	0.650	0.682	2.019
Std err	0.028	0.048	0.007	0.010	0.077
Estimate	–	–	–	–	0.088
Std err	–	–	–	–	0.034
Estimate	–	–	–	–	-0.015
Std err	–	–	–	–	0.035
$Q_{SMM} \times 100$	–	19.820	–	0.240	0.040
J_{pval}	–	0.000	–	0.043	0.139

Notes: This table presents estimation results for various copula models applied to seven daily stock returns in the financial sector over the period January 2001 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.

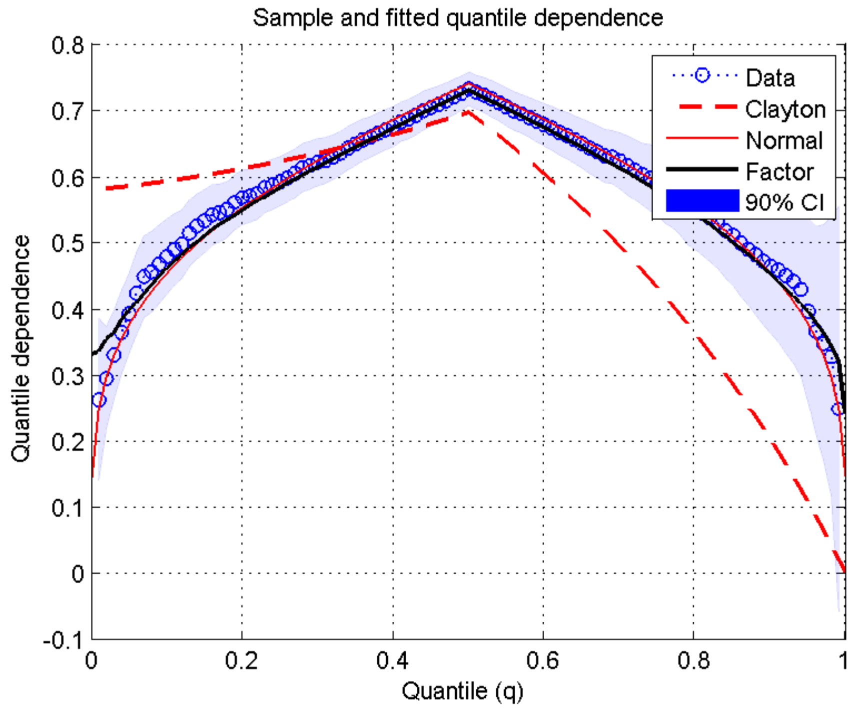

Figure 1: *This figure plots the probability of both variables being less than their q quantile (q<0.5) or greater than the q quantile (q>0.5). For the data this is averaged across all pairs, and a bootstrap 90% (pointwise) confidence interval is presented.*

.