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

Dong Hwan Oh and Andrew J. Patton

Duke University

30 May 2012

.

*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 specification 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. We describe these here and contrast them with the SMM approach proposed in this paper. 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) is one of the most widely-used estimation methods. The "maximization by parts" algorithm of Song, Fan and Kalbfleisch (2005) is an iterative method that improves the efficiency of multi-stage MLE, and attains full efficiency under some conditions. Like MLE, both of these methods only apply when the marginal distributions are parametric. When the marginal distribution models are correctly specified this improves the efficiency of the estimator, relative to the proposed SMM approach using nonparametric margins, however it introduces the possibility of mis-specified marginal distributions, which can have deleterious effects on the copula parameter estimates, see Kim, Silvapulle and Silvapulle (2007).

Semi-parametric maximum likelihood (see Genest, Ghoudi and Rivest (1995) for *iid* data and Chen and Fan (2006), Chan, Chen, Chen, Fan and Peng (2009) and Chen, Fan and Tsyrennikov (2006) for time series data) is also a widely-used estimation method and has a number of attractive features. Most importantly, with respect to SMM approach proposed here, it yields fully efficient estimates of the copula parameters, whereas SMM generally does not. Semi-parametric MLE requires, of course, the copula likelihood and for some more complicated models the likelihood can be cumbersome to derive or to compute, e.g. the "stochastic copula" model of Hafner and Manner (2012) or the high dimension factor copula model of Oh and Patton (2011). In such applications it may be desirable to avoid the likelihood and use a simpler SMM approach.

A long-standing estimator of the copula parameter is the method of moments (MM) estimator (see Genest (1987) and Genest and Rivest (1993) for *iid* data and Rémillard (2010) for time series data). This estimator exploits the known one-to-one mapping between the parameters of certain copulas and certain measures of dependence. For example, a Clayton copula with parameter κ implies Kendall's tau of $\kappa/(\kappa+2)$, yielding a simple MM estimator of the parameter of this copula as $\hat{\kappa} = 2\hat{\tau}/(1-\hat{\tau})$. MM estimators usually have the benefit of being very fast to compute. The SMM estimator proposed in this paper is a generalization of MM in two directions. Firstly, it allows the consideration of over-identified models: For some copulas we have more implied dependence measures than unknown parameters (e.g., for the Normal copula we have both Kendall's tau and Spearman's rank correlation in closed form). By treating this as a GMM estimation problem we can draw on the information in all available dependence measures. Secondly, we allow for dependence measures that are not known closed-form functions of the copula parameters, and use simulations to obtain the mapping, making this SMM rather than GMM.

Other, less-widely used, estimation methods considered in the literature include minimum distance estimation, see Tsukahara (2005), and "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 and MM estimation of the copula parameter as special cases.

2 Simulation-based estimation of copula models

We consider the same class of data generating processes (DGPs) as Chen and Fan (2006), Chan, Chen, Chen, Fan and Peng (2009) 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:

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

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, quantile dependence, and Kendall's tau for the pair $\left(\eta_i, \eta_j\right)$ are defined as:

$$
\begin{aligned}
\rho^{ij} &\equiv 12E\left[F_i\left(\eta_i\right)F_j\left(\eta_j\right)\right]-3 = 12\int\int uvdC_{ij}\left(u,v\right)-3 &(2)\\
\lambda_q^{ij} &\equiv \begin{cases} P\left[F_i\left(\eta_i\right)\leq q | F_j\left(\eta_j\right)\leq q\right] = \frac{C_{ij}(q,q)}{q}, & q\in(0,0.5] \\ P\left[F_i\left(\eta_i\right)> q | F_j\left(\eta_j\right)> q\right] = \frac{1-2q+C_{ij}(q,q)}{1-q}, & q\in(0.5,1)\end{cases} &(3)\\
\tau^{ij} &\equiv 4E\left[C_{ij}\left(F_i\left(\eta_i\right),F_j\left(\eta_j\right)\right)\right]-1 &(4)
\end{aligned}
$$

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:

$$
\begin{aligned}
\hat{\rho}^{ij} &\equiv \frac{12}{T}\sum_{t=1}^{T}\hat{F}_i\left(\hat{\eta}_{it}\right)\hat{F}_j\left(\hat{\eta}_{jt}\right)-3 &(5)\\
\hat{\lambda}_q^{ij} &\equiv \begin{cases} \frac{1}{Tq}\sum_{t=1}^{T}1\{\hat{F}_i\left(\hat{\eta}_{it}\right)\leq q, \hat{F}_j\left(\hat{\eta}_{jt}\right)\leq q\}, & q\in(0,0.5] \\ \frac{1}{T(1-q)}\sum_{t=1}^{T}1\{\hat{F}_i\left(\hat{\eta}_{it}\right)> q, \hat{F}_j\left(\hat{\eta}_{jt}\right)> q\}, & q\in(0.5,1)\end{cases} &(6)\\
\hat{\tau}^{ij} &\equiv \frac{4}{T}\sum_{t=1}^{T}\hat{C}_{ij}\left(\hat{F}_i\left(\hat{\eta}_{it}\right),\hat{F}_j\left(\hat{\eta}_{jt}\right)\right)-1 &(7)
\end{aligned}
$$

where $\hat{F}_i\left(y\right)\equiv\left(T+1\right)^{-1}\sum_{t=1}^{T}1\{\hat{\eta}_{it}\leq y\}$, and $\hat{C}_{ij}\left(u,v\right)\equiv\left(T+1\right)^{-1}\sum_{t=1}^{T}1\{\hat{F}_i\left(\hat{\eta}_{it}\right)\leq u, \hat{F}_j\left(\hat{\eta}_{jt}\right)\leq v\}$. Counterparts based on simulations are denoted $\tilde{\rho}^{ij}\left(\boldsymbol{\theta}\right)$, $\tilde{\lambda}_q^{ij}\left(\boldsymbol{\theta}\right)$ and $\tilde{\tau}^{ij}\left(\boldsymbol{\theta}\right)$.

Let $\tilde{\mathbf{m}}_S\left(\boldsymbol{\theta}\right)$ be a $(m\times 1)$ vector of dependence measures computed using S simulations from $\mathbf{F}_x\left(\boldsymbol{\theta}\right)$, $\{\mathbf{X}_s\}_{s=1}^{S}$, 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

$$
\mathbf{g}_{T,S}\left(\boldsymbol{\theta}\right)\equiv\hat{\mathbf{m}}_T-\tilde{\mathbf{m}}_S\left(\boldsymbol{\theta}\right) \tag{8}
$$

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:

$$\hat{\boldsymbol{\theta}}_{T,S} \equiv \underset{\boldsymbol{\theta}\in\boldsymbol{\Theta}}{\arg\min}\ Q_{T,S}(\boldsymbol{\theta}) \tag{9}$$

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

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. The supplemental appendix presents details on the computation of the objective function.

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 $\boldsymbol{\theta}$ varies on $\boldsymbol{\Theta}$, 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 equation (1), or if $\boldsymbol{\phi}_0$ is known, then Assumption 1 is sufficient to prove Proposition 1 below, using the results of Fermanian, Radulović and Wegkamp (2004). If, however, estimated 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 $\dot{\boldsymbol{\mu}}_t\left(\boldsymbol{\phi}\right) = \frac{\partial \boldsymbol{\mu}_t(\boldsymbol{\phi})}{\partial \boldsymbol{\phi}'}$, $\dot{\boldsymbol{\sigma}}_{kt}\left(\boldsymbol{\phi}\right) = \frac{\partial [\boldsymbol{\sigma}_t(\boldsymbol{\phi})]_{\text{k-th column}}}{\partial \boldsymbol{\phi}'}$, $k = 1, \ldots, N$. 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:

$$\mathbf{g}_{T,S}(\boldsymbol{\theta}) \equiv \hat{\mathbf{m}}_T - \tilde{\mathbf{m}}_S(\boldsymbol{\theta}) \xrightarrow{p} \mathbf{g}_0(\boldsymbol{\theta}) \equiv \mathbf{m}_0(\boldsymbol{\theta}_0) - \mathbf{m}_0(\boldsymbol{\theta}), \text{ for } \forall \boldsymbol{\theta} \in \boldsymbol{\Theta} \text{ as } T, S \to \infty \quad (10)$$

We define the population objective function as

$$Q_0(\boldsymbol{\theta}) = \mathbf{g}_0(\boldsymbol{\theta})' \mathbf{W}_0 \mathbf{g}_0(\boldsymbol{\theta}) \quad (11)$$

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 the 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 $\mathbf{g}_{T,S}(\boldsymbol{\theta})$, and can easily be shown to be satisfied for many bivariate parametric copulas. Assumption 3(ii) requires compactness of the parameter space, a common assumption, and is aided by having outside information (such as constraints from economic arguments) that allow the researcher to bound the set of plausible parameters. 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 $T, S \to \infty$.

We focus on weak consistency of our estimator because it suffices for our asymptotic distribution theory, presented below. A corresponding strong consistency result, i.e., $\hat{\boldsymbol{\theta}}_{T,S} \xrightarrow{a.s.} \boldsymbol{\theta}_0$, may be obtained by drawing on recent work by Bouzebda and Zari (2011). The key is to show uniform strong convergence of the sample objective function, from which strong consistency of the estimator easily follows, see Newey and McFadden (1994) for example. Uniform strong consistency of the objective function can be shown by combining minor changes in the above assumptions (e.g., $\hat{\mathbf{W}}_T$ must converge *a.s.* to $\mathbf{W}_0$) with pointwise strong convergence of the objective function, which can be obtained using results of Bouzebda and Zari (2011).

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

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 $o_p(1/T + 1/S) = o_p\left(\min(T,S)^{-1}\right)$. 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.

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

$$\frac{1}{\sqrt{1/T + 1/S}}\left(\hat{\boldsymbol{\theta}}_{T,S} - \boldsymbol{\theta}_0\right) \xrightarrow{d} N(0, \boldsymbol{\Omega}_0) \text{ as } T, S \to \infty \tag{12}$$

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* $\boldsymbol{\Sigma}_0 \equiv avar[\hat{\mathbf{m}}_T]$.

The rate of convergence is thus shown to equal $\min(T,S)^{1/2}$. In general, one would like to set S very large to minimize the impact of simulation error and obtain a $\sqrt{T}$ convergence rate, however if the model is computationally costly to simulate, then the result for $S \ll T$ may be useful. When S and T diverge at different rates the asymptotic variance of $\min(T,S)^{1/2}\left(\hat{\boldsymbol{\theta}}_{T,S}-\boldsymbol{\theta}_0\right)$ is simply $\boldsymbol{\Omega}_0$. When S and T diverge at the same rate, say $S/T \to k \in (0,\infty)$, the asymptotic variance of $\sqrt{T}\left(\hat{\boldsymbol{\theta}}_{T,S}-\boldsymbol{\theta}_0\right)$ is $(1+1/k)\,\boldsymbol{\Omega}_0$, which incorporates efficiency loss from simulation error. As usual we find that $\boldsymbol{\Omega}_0=\left(\mathbf{G}_0'\boldsymbol{\Sigma}_0^{-1}\mathbf{G}_0\right)^{-1}$ if $\mathbf{W}_0$ is the efficient weight matrix, $\boldsymbol{\Sigma}_0^{-1}$.

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.

Chen and Fan (2006), Chan, Chen, Chen, Fan and Peng (2009) 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. In applications based on *parametric* models for the marginal distributions, the asymptotic covariance matrix of the copula parameter is more complicated. In such cases, the model is fully parametric and the estimation approach here is a form of two-stage GMM (or SMM). In the absence of simulations, this can be treated using existing methods, see White (1994) and Gouriéroux, Monfort and Renault (1996) for example. If simulations are used in the copula estimation step, then the lemmas presented in the appendix can be combined with existing results on two-stage GMM to obtain the limiting distribution. This is straightforward but requires some additional detailed notation, and so is not presented here.

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:

$$\hat{\mathbf{\Sigma}}_{T,B} = \frac{T}{B} \sum_{b=1}^{B} \left(\hat{\mathbf{m}}_T^{(b)} - \hat{\mathbf{m}}_T\right) \left(\hat{\mathbf{m}}_T^{(b)} - \hat{\mathbf{m}}_T\right)' \tag{13}$$

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., $1/\min\left(\sqrt{T}, \sqrt{S}\right)$). 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

$$\hat{\mathbf{G}}_{T,S,k} = \frac{\mathbf{g}_{T,S}\left(\hat{\boldsymbol{\theta}}_{T,S}+\mathbf{e}_k\varepsilon_{T,S}\right) - \mathbf{g}_{T,S}\left(\hat{\boldsymbol{\theta}}_{T,S}-\mathbf{e}_k\varepsilon_{T,S}\right)}{2\varepsilon_{T,S}}, \ k = 1, 2, ..., p \tag{14}$$

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

$$\hat{\mathbf{\Omega}}_{T,S,B} = \left(\hat{\mathbf{G}}_{T,S}'\hat{\mathbf{W}}_T\hat{\mathbf{G}}_{T,S}\right)^{-1} \hat{\mathbf{G}}_{T,S}'\hat{\mathbf{W}}_T\hat{\mathbf{\Sigma}}_{T,B}\hat{\mathbf{W}}_T\hat{\mathbf{G}}_{T,S} \left(\hat{\mathbf{G}}_{T,S}'\hat{\mathbf{W}}_T\hat{\mathbf{G}}_{T,S}\right)^{-1} \tag{15}$$

Proposition 3 *Suppose that all assumptions of Proposition 2 are satisfied, and that* $\varepsilon_{T,S} \rightarrow 0,\ \varepsilon_{T,S} \times \min\left(\sqrt{T}, \sqrt{S}\right) \rightarrow \infty,\ B \rightarrow \infty$ *as* $T,\ S \rightarrow \infty$. *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* $T, S \rightarrow \infty$.

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, which can be used as a specification test of the model. 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 3 are satisfied and that the number of moments* (m) *is greater than the number of copula parameters* (p). *Then*

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

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 $J_{T,S}^{(k)} = \mathbf{u}^{(k)\prime}\hat{\mathbf{R}}^{\prime}\hat{\mathbf{\Sigma}}_{T,B}^{1/2\prime}\hat{\mathbf{W}}_T\hat{\mathbf{\Sigma}}_{T,B}^{1/2}\hat{\mathbf{R}}\mathbf{u}^{(k)}$

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.

2.6 SMM under model mis-specification

All of the above results hold under the assumption that the copula model is correctly specified. In the event that the specification test proposed in the previous section rejects a model as mis-specified, one is led directly to the question of whether these results, or extensions of them, hold for mis-specified models.

In the literature on GMM, there are two common ways to define mis-specification. Newey (1985) defines a form of "local" mis-specification (where the degree of mis-specification vanishes in the limit), and in that case it is simple to show that the asymptotic behavior of the SMM estimator does not change at all except the mean of limit distribution. Hall and Inoue (2003) consider "non-local" mis-specification. Formally, a model is said to be mis-specified if there is no value of $\boldsymbol{\theta} \in \mathbf{\Theta}$ which satisfies $\mathbf{g}_0(\boldsymbol{\theta}) = \mathbf{0}$. As Hall and Inoue (2003) note, mis-specification is only a concern when the model is over-identified, and so in this section we assume $m > p$. The absence of a parameter that satisfies the population moment conditions means that we must instead consider a "pseudo-true" parameter:

Definition 1 *The pseudo-true parameter is* $\boldsymbol{\theta}_* (\mathbf{W}_0) \equiv \arg\min_{\boldsymbol{\theta} \in \boldsymbol{\Theta}} \mathbf{g}_0' (\boldsymbol{\theta}) \mathbf{W}_0 \mathbf{g}_0 (\boldsymbol{\theta})$.

While the true parameter, $\boldsymbol{\theta}_0$, when it exists, is determined only by the population moment condition $\mathbf{g}_0 (\boldsymbol{\theta}_0) = \mathbf{0}$, the pseudo-true parameter depends on the moment condition and also on the weight matrix $\mathbf{W}_0$, and thus we denote it as $\boldsymbol{\theta}_* (\mathbf{W}_0)$. With the additional assumptions below, the consistency of the SMM estimator under mis-specification can be proven.

Assumption 5 *(i) (Non-local mis-specification)* $\|\mathbf{g}_0 (\boldsymbol{\theta})\| > 0$ *for all* $\boldsymbol{\theta} \in \boldsymbol{\Theta}$

(ii) (Identification) There exists $\boldsymbol{\theta}_* (\mathbf{W}_0) \in \boldsymbol{\Theta}$ *such that* $\mathbf{g}_0 (\boldsymbol{\theta}_* (\mathbf{W}_0))' \mathbf{W}_0 \mathbf{g}_0 (\boldsymbol{\theta}_* (\mathbf{W}_0)) < \mathbf{g}_0 (\boldsymbol{\theta})' \mathbf{W}_0 \mathbf{g}_0 (\boldsymbol{\theta})$ *for all* $\boldsymbol{\theta} \in \boldsymbol{\Theta} \backslash \{\boldsymbol{\theta}_* (\mathbf{W}_0)\}$

Proposition 5 *Suppose Assumption 1, 2, 3(ii)-3(iv), and 5 holds. Then* $\hat{\boldsymbol{\theta}}_{T,S} \xrightarrow{p} \boldsymbol{\theta}_* (\mathbf{W}_0)$ *as* $T, S \to \infty$

The above proposition shows that, under mis-specification, the SMM estimator $\hat{\boldsymbol{\theta}}_{T,S}$ converges in probability to the pseudo true parameter $\boldsymbol{\theta}_* (\mathbf{W}_0)$ rather than the true parameter $\boldsymbol{\theta}_0$. This extends existing results for GMM under mis-specification in Hall (2000) and Hall and Inoue (2003), as it is established under the discontinuity of the moment functions.

While consistency of $\hat{\boldsymbol{\theta}}_{T,S}$ under mis-specification is easily obtained, establishing the limit distribution of $\hat{\boldsymbol{\theta}}_{T,S}$ is not straightforward. A key contribution of Hall and Inoue (2003) is to show that the limit distribution of GMM (with smooth, differentiable moment functions) depends on the limit distribution of the weight matrix, not merely the probability limit of the weight matrix. In SMM applications, it is possible to show that the limit distribution will additionally depend on the limit distribution of the numerical derivative matrix, denoted $\hat{\mathbf{G}}_{T,S}$ above. Some results on the statistical properties of numerical derivatives are presented in Hong, Mahajan and Nekipelov (2010), but this remains a relatively unexplored topic. In addition to incorporating the dependence on the distribution of $\hat{\mathbf{G}}_{T,S}$, under mis-specification one needs an alternative approach to establish the stochastic equicontinuity of the objective function, which is required for a Taylor series expansion of the population objective function to be used to obtain the limit distribution of the estimator. We leave the interesting problem of the limit distribution of $\hat{\boldsymbol{\theta}}_{T,S}$ under mis-specification for future research.

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. A closed-form likelihood is available for the first two copulas, while the third copula requires a numerical integration step to obtain the likelihood (details on this are presented in the supplemental appendix). In all cases we contrast the finite-sample properties of the MLE with the SMM estimator. The first 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 tau in closed form ($\tau = \kappa / (2 + \kappa)$) but not Spearman's rank correlation, see Nelsen (2006). For the Normal copula we have both Spearman's rank correlation in closed form ($\rho_S = 6/\pi \arcsin(\rho/2)$) and Kendall's tau ($\tau = 2/\pi \arcsin(\rho)$), see Nelsen (2006) and Demarta and McNeil (2005). 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:

$$
\begin{aligned}
\text{Let} \quad X_i &= Z + \varepsilon_i, \quad i = 1, 2, ..., N \\
\text{where} \;\; Z &\sim Skew\ t\left(0, \sigma^2, \nu^{-1}, \lambda\right), \; \varepsilon_i \sim iid\ t\left(\nu^{-1}\right), \;\; \text{and} \;\; \varepsilon_i \perp\!\!\!\perp Z \; \forall \; i \qquad (16) \\
[X_1, ..., X_N]' &\equiv \mathbf{X} \sim \mathbf{F}_x = \mathbf{C}\left(G_x, ..., G_x\right)
\end{aligned}
$$

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 $\left(\sigma^2, \nu^{-1}, \lambda\right)$. For the factor copula we have none 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:

$$
\begin{aligned}
Y_{it} &= \phi_0 + \phi_1 Y_{i,t-1} + \sigma_{it}\eta_{it}, \quad t = 1, 2, ..., T \\
\sigma_{it}^2 &= \omega + \beta\sigma_{i,t-1}^2 + \alpha\sigma_{i,t-1}^2\eta_{i,t-1}^2 \qquad (17) \\
\boldsymbol{\eta}_t &\equiv [\eta_{1t}, ..., \eta_{Nt}]' \sim iid \;\; \mathbf{F}_\eta = \mathbf{C}\left(\Phi, \Phi, ..., \Phi\right)
\end{aligned}
$$

where Φ is the standard Normal distribution function and $\mathbf{C}$ can be Clayton, Gaussian, or the factor copula implied by equation (16). 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:

$$\hat{\eta}_{it} = \frac{Y_{it} - \hat{\phi}_0 - \hat{\phi}_1 Y_{i,t-1}}{\hat{\sigma}_{it}}. \tag{18}$$

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, $\hat{\mathbf{W}}_T = \hat{\mathbf{\Sigma}}_{T,B}^{-1}$, are comparable, and available in the supplemental appendix to this paper.) We also report the computation times (per simulation) for each estimation.

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, we see that the SMM estimators suffer a loss in efficiency of around 50% for $N = 2$ and around 20% for $N = 10$. The loss is greatest for the ν^{-1} parameter of the factor copula, and is moderate and similar for the remaining parameters. Some loss is of course expected, and this simulation indicates that the loss is moderate overall. 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 $\varepsilon_{T,S} > 0.03$, 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 $\mathbf{W} = \mathbf{I}$ 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.

[INSERT TABLE 3 ABOUT HERE]

We finally consider the properties of the estimator under model mis-specification. In Table 4 we consider two scenarios: one where the true copula is Clayton but the model is Normal, and one where the true copula is Normal but the model is Clayton. The pseudo-true parameters for these two scenarios are not known in closed form, and we use a simulation of 10 million observations to estimate it. The pseudo-true parameters vary across the dimension of the model, and we report them in the top row of each panel of Table 4. The remainder of Table 4 has the same structure as Tables 1 and 2. Similar to those tables, in this mis-specified case we see that the estimated parameters are centered on the pseudo-true values, with the average estimated bias being small relative to the standard deviation. These mis-specified scenarios also provide some insight into the power of the specification test based on over-identifying restrictions. We find that for all three dimensions and for both *iid* and AR-GARCH data, the J-test rejected the null of correct specification across all 100 simulations, indicating this test has power to detect model mis-specification.

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, and the test for model mis-specification has finite-sample rejection frequencies that are close to their nominal levels when the model is correctly specified, and has good power to reject mis-specified models.

[INSERT TABLE 4 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 S4 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:

$$\begin{aligned} r_{it} &= \phi_{0i} + \phi_{1i} r_{i,t-1} + \phi_{2i} r_{m,t-1} + \varepsilon_{it}, \quad \varepsilon_{it} = \sigma_{it}\eta_{it} \\ \text{where} \quad \sigma_{it}^2 &= \omega_i + \beta_i \sigma_{i,t-1}^2 + \alpha_{1i}\varepsilon_{i,t-1}^2 + \gamma_{1i}\varepsilon_{i,t-1}^2 \cdot \mathbf{1}_{[\varepsilon_{i,t-1} \leq 0]} \\ &\quad + \alpha_{2i}\varepsilon_{m,t-1}^2 + \gamma_{2i}\varepsilon_{m,t-1}^2 \cdot \mathbf{1}_{[\varepsilon_{m,t-1} \leq 0]} \end{aligned} \tag{19}$$

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 S5 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 5 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 5 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 (6), 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 6 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 5 AND 6 ABOUT HERE]

For all three copulas we implent the SMM estimator proposed in Section 2, with the identity weight matrix and the efficient weight matrix, 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. We also implement the MLE for comparison. 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 and factor copula models 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 (with a p-value of less than 0.001 for both choices of weight matrix), while the Normal is less strongly rejected (p-values of 0.043 and 0.001). The factor copula is not rejected using this test for either choice of weight matrix. 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.

Given that our sample period spands the financial crisis, one may wonder whether the copula is constant throughout the period. To investigate this, we implement the copula structrual break test proposed by Rémillard (2010). This test uses a Kolmogorov-Smirnov type test statistic to compare the empirical copula before and after a given point in the sample, and then searches across all dates in the sample. We implement this test using 1000 simulations for the "multiplier" method, and find a p-value of 0.001, indicating strong evidence of a change in the copula over this period. Running this test on just the last two years of our sample period (January 2009 to December 2010) results in a p-value of 0.191, indicating no evidence of a change in the copula over this sub-period. We re-estimate our three copula models using data from this sub-period, and present the results in the lower panel of Table 6. The estimated parameters all indicate a (slight) increase in dependence in this sub-sample relative to the full sample. Perhaps the largest change is in the ν parameter of the factor copula, which goes from around 8.8 (across the three estimation methods) to around 4.4, indicating a substantial increase in the degree of tail dependence between these firms. The results of the specification tests over this sub-sample are comparable to the full sample results: the Clayton copula is strongly rejected, the Normal copula is rejected but less strongly, and the factor copula is not rejected, using either weight matrix.

Figure 1 sheds some further light on the relative performance of these copula models, over the full sample. 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 the 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

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

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, Radulović and Wegkamp with those of Rémillard (2010): For simplicity, assume that only one dependence measure is used. Let $\hat{\rho}_{ij}$ and $\hat{\rho}_{ij}^{(b)}$ be the sample rank correlations 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, $\ddot{\rho}_{ij}$ and $\ddot{\rho}_{ij}^{(b)}$, 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 ρ as ρ_0. Theorem 5 of Fermanian, Radulović and Wegkamp (2004) shows that

$$\sqrt{T}\left(\ddot{\rho}_{ij} - \rho_0\right) = \sqrt{T}\left(\ddot{\rho}_{ij}^{(b)} - \ddot{\rho}_{ij}\right) + o_p(1)$$

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

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

Combining those three equations, we obtain

$$\sqrt{T}\left(\hat{\rho}_{ij} - \rho_0\right) = \sqrt{T}\left(\hat{\rho}_{ij}^{(b)} - \hat{\rho}_{ij}\right) + o_p(1), \text{ as } T, B \to \infty$$

and so equation (13) 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 $\sqrt{T}\left\|\hat{\boldsymbol{\theta}}_{T,S} - \boldsymbol{\theta}_0\right\| = O_p(1)$

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

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

$$\sqrt{T}\mathbf{g}_{T,S}\left(\hat{\boldsymbol{\theta}}_{T,S}\right) = \sqrt{T}\mathbf{g}_{T,S}\left(\boldsymbol{\theta}_0\right) + \hat{\mathbf{G}}_{T,S} \cdot \sqrt{T}\left(\hat{\boldsymbol{\theta}}_{T,S} - \boldsymbol{\theta}_0\right) + \mathbf{R}_{T,S}\left(\hat{\boldsymbol{\theta}}_{T,S}\right) \tag{21}$$

where the remaining term is captured by $\mathbf{R}_{T,S}\left(\hat{\boldsymbol{\theta}}_{T,S}\right)$. Combining equations (20) and (21) we obtain

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

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

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

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

$$\sqrt{T}\mathbf{g}_{T,S}\left(\hat{\boldsymbol{\theta}}_{T,S}\right) = \sqrt{T}\mathbf{g}_{T,S}(\boldsymbol{\theta}_0) + \hat{\mathbf{G}}_{T,S} \cdot \sqrt{T}\left(\hat{\boldsymbol{\theta}}_{T,S} - \boldsymbol{\theta}_0\right) + o_p(1) \tag{22}$$

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

Proof of Proposition 5. Lemma 1, 2, 3 and 4 are not affected by mis-specification. Lemma 5 (i) is replaced by Assumption 5 (ii). Therefore, $\hat{\boldsymbol{\theta}}_{T,S} \xrightarrow{p} \boldsymbol{\theta}_*(\mathbf{W}_0)$. ■

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, 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, 63-92.

[5] Bouzebda, S. and T. Zari, 2011, Strong Approximation of Empirical Copula Processes by Gaussian Processes, preprint available at http://arxiv.org/pdf/0811.3330.pdf

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

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

[8] 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.

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

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

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

[12] 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.

[13] 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, 210-218.

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

[15] 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.

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

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

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

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

[20] 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.

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

[22] Gouriéroux, C., A. Monfort, and E. Renault, 1996, Two-stage generalized moment method with applications to regressions with heteroscedasticity of unknown form, *Journal of Statistical Planning and Inference,* 50, 37-63.

[23] Hafner, C. M. and H. Manner, 2012, Dynamic Stochastic Copula Models: Estimation, Inference and Applications, *Journal of Applied Econometrics*, 27, 269-295.

[24] Hall, A.R., 2000, Covariance matrix estimation and the power of the overidentifying restrictions test, *Econometrica*, 68, 1517-1528.

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

[26] Hall, A.R. and A. Inoue, 2003, The large sample behaviour of the generalized method of moments estimator in misspecified models, *Journal of Econometrics,* 114, 361-394.

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

[28] Hong, H., A. Mahajan and D. Nekipelov, 2010, Extremum Estimation and Numerical Derivatives, working paper, Stanford University.

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

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

[31] 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.

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

[33] Kim, G., M. J. Silvapulle and P. Silvapulle, 2007, Comparison of Semiparametric and Parametric Methods for Estimating Copulas, *Computational Statistics & Data Analysis*, 51, 2836-2850.

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

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

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

[37] Newey, W.K., 1985, Generalized Method of Moments Specification Testing, *Journal of Econometrics*, 29, 229-256.

[38] 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.

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

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

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

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

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

[44] 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.

[45] Rothenberg, T.J., 1971, Identification of Parametric Models, *Econometrica*, 39, 577-591.

[46] 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, 1467-1479.

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

[48] Song, P., Y. Fan and J. Kalbfleisch, 2005, Maximization by Parts in Likelihood Inference, *Journal of the American Statistical Association*, 100, 1145-1158.

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

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

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

[52] 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	MLE			SMM		
	κ	κ	κ	κ	ρ	ρ	ρ	σ^2	ν^{-1}	λ	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	1.00	0.25	-0.50
							$N = 2$						
Bias	0.001	-0.014	-0.006	-0.004	0.004	-0.001	-0.001	0.026	-0.002	-0.026	0.016	-0.012	-0.089
St dev	0.085	0.119	0.122	0.110	0.024	0.034	0.034	0.135	0.045	0.144	0.152	0.123	0.199
Median	1.011	0.982	0.991	0.998	0.503	0.497	0.498	1.027	0.251	-0.500	0.985	0.243	-0.548
90-10%	0.216	0.308	0.293	0.294	0.063	0.086	0.087	0.331	0.118	0.355	0.374	0.363	0.540
Time	0.061	0.060	512	49.5	0.021	0.292	0.483		254			103	
							$N = 3$						
Bias	0.015	0.008	0.006	0.008	0.003	-0.004	-0.005	0.014	-0.001	-0.012	0.032	-0.002	-0.057
St dev	0.061	0.090	0.092	0.091	0.020	0.025	0.026	0.120	0.028	0.109	0.124	0.111	0.157
Median	1.013	1.003	0.999	0.998	0.503	0.497	0.499	1.001	0.250	-0.502	1.031	0.256	-0.542
90-10%	0.155	0.226	0.219	0.216	0.049	0.064	0.068	0.297	0.073	0.222	0.297	0.293	0.395
Time	0.113	0.091	1360	56.2	0.023	0.293	0.815		263			136	
							$N = 10$						
Bias	0.008	0.007	0.008	0.004	0.003	-0.002	-0.002	0.011	0.000	0.006	0.026	0.001	-0.011
St dev	0.050	0.068	0.066	0.059	0.014	0.017	0.017	0.092	0.016	0.063	0.093	0.070	0.082
Median	1.005	1.002	1.005	0.999	0.504	0.498	0.499	1.005	0.248	-0.494	1.013	0.255	-0.508
90-10%	0.132	0.198	0.177	0.152	0.035	0.039	0.045	0.240	0.034	0.166	0.248	0.186	0.168
Time	0.409	0.998	22289	170	0.475	0.331	3.140		396			341	

Table 2: Simulation results for AR-GARCH data

	Clayton				*Normal*			*Factor copula*					
	MLE	GMM	SMM	SMM*	MLE	GMM	SMM	MLE			SMM		
	κ	κ	κ	κ	ρ	ρ	ρ	σ^2	ν^{-1}	λ	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	1.00	0.25	-0.50
							$N = 2$						
Bias	-0.005	-0.029	-0.028	-0.020	0.003	-0.001	-0.001	0.021	-0.009	-0.029	0.015	-0.012	-0.073
St dev	0.087	0.124	0.124	0.108	0.024	0.035	0.036	0.137	0.046	0.150	0.155	0.121	0.188
Median	0.998	0.977	0.975	0.982	0.503	0.497	0.499	1.021	0.245	-0.503	0.995	0.235	-0.558
90-10%	0.228	0.327	0.340	0.267	0.061	0.084	0.090	0.343	0.118	0.382	0.411	0.346	0.509
Time	0.026	0.059	525	52	0.030	0.299	0.505		234			95	
							$N = 3$						
Bias	0.006	-0.007	0.002	-0.008	0.003	-0.005	-0.006	0.007	-0.007	-0.011	0.013	-0.020	-0.052
St dev	0.060	0.087	0.088	0.080	0.020	0.026	0.026	0.118	0.028	0.110	0.121	0.106	0.148
Median	1.005	0.991	0.994	0.981	0.502	0.497	0.499	0.997	0.243	-0.502	1.005	0.238	-0.521
90-10%	0.145	0.205	0.213	0.195	0.050	0.065	0.068	0.315	0.074	0.224	0.311	0.297	0.357
Time	0.127	0.108	1577	73	0.022	0.288	1.009		232			119	
							$N = 10$						
Bias	-0.004	-0.002	-0.004	-0.003	0.002	-0.003	-0.004	0.005	-0.006	0.008	-0.004	-0.016	-0.012
St dev	0.049	0.067	0.064	0.059	0.014	0.016	0.017	0.091	0.015	0.063	0.085	0.079	0.071
Median	0.995	0.996	0.987	0.988	0.503	0.497	0.497	1.002	0.243	-0.493	0.990	0.238	-0.508
90-10%	0.134	0.179	0.170	0.152	0.034	0.041	0.045	0.240	0.037	0.169	0.210	0.209	0.165
Time	0.292	1.059	24549	171	1.099	0.392	3.437		430			309	

Notes to Table 1: 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. The last row in each panel presents the average time in seconds to compute the estimator.

Notes to Table 2: 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. The last row in each panel presents the average time in seconds to compute the estimator.

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: Simulation results for mis-specified models

	iid		*AR-GARCH*	
True copula	*Normal*	*Clayton*	*Normal*	*Clayton*
Model	Clayton	Normal	Clayton	Normal
		$N = 2$		
Pseudo-true	0.542	0.599	0.543	0.588
Bias	-0.013	0.111	-0.007	0.046
St dev	0.050	0.173	0.035	0.120
Median	0.526	0.659	0.539	0.617
90-10%	0.130	0.433	0.091	0.265
Time	4	72	1	70
J test prob.	0	0	0	0
		$N = 3$		
Pseudo-true	0.543	0.599	0.542	0.607
Bias	0.003	0.077	-0.002	0.006
St dev	0.039	0.164	0.027	0.088
Median	0.544	0.629	0.540	0.609
90-10%	0.107	0.432	0.072	0.198
Time	5	90	1	86
J test prob.	0	0	0	0
		$N = 10$		
Pseudo-true	0.544	0.602	0.544	0.603
Bias	0.001	0.059	-0.001	0.047
St dev	0.033	0.118	0.016	0.116
Median	0.546	0.622	0.540	0.618
90-10%	0.086	0.307	0.043	0.314
Time	20	206	4	207
J test prob.	0	0	0	0

Notes: This table presents the results from 100 simulations when the true copula and the model are different (i.e., the model is mis-specified). The parameters of the copula models are estimated using SMM based on rank correlation and four quantile dependence measures ($q = 0.05, 0.10, 0.90, 0, 95$). The marginal distributions of the data are assumed to be either *iid* $N(0,1)$ or 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 pseudo-true parameter is estimated using 10 million observations. The last row in each panel presents the proportion of tests of over-identifying restrictions that are smaller than the 95% critical value.

Table 5: 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 6: Estimation results for daily returns on seven stocks

	Panel A: Full sample (2001-2010)								
	Clayton			*Normal*			*Factor*		
	MLE	SMM	SMM-opt	MLE	SMM	SMM-opt	MLE	SMM	SMM-opt
	κ	κ	κ	ρ	ρ	ρ	$\sigma^2, \nu^{-1}, \lambda$	$\sigma^2, \nu^{-1}, \lambda$	$\sigma^2, \nu^{-1}, \lambda$
Estimate	0.907	1.274	1.346	0.650	0.682	0.659	1.995	2.019	1.955
Std err	0.028	0.048	0.037	0.007	0.010	0.008	0.020	0.077	0.069
Estimate	–	–	–	–	–	–	0.159	0.088	0.115
Std err	–	–	–	–	–	–	0.010	0.034	0.033
Estimate	–	–	–	–	–	–	-0.021	-0.015	-0.013
Std err	–	–	–	–	–	–	0.032	0.035	0.034
$Q_{SMM} \times 100$	–	19.820	19.872	–	0.240	0.719	–	0.040	0.187
J_{pval}	–	0.000	0.000	–	0.043	0.001	–	0.139	0.096
Time	0.7	344	360	0.5	6	6	1734	801	858

Panel B: Sub-sample (2009-2010)

	Clayton			*Normal*			*Factor*		
	MLE	SMM	SMM-opt	MLE	SMM	SMM-opt	MLE	SMM	SMM-opt
	κ	κ	κ	ρ	ρ	ρ	$\sigma^2, \nu^{-1}, \lambda$	$\sigma^2, \nu^{-1}, \lambda$	$\sigma^2, \nu^{-1}, \lambda$
Estimate	1.053	1.649	1.581	0.660	0.768	0.701	2.165	2.567	2.575
Std err	0.072	0.152	0.094	0.015	0.370	0.192	0.018	0.249	0.256
Estimate	–	–	–	–	–	–	0.167	0.266	0.297
Std err	–	–	–	–	–	–	0.062	0.081	0.078
Estimate	–	–	–	–	–	–	-0.095	0.033	0.050
Std err	–	–	–	–	–	–	0.062	0.070	0.067
$Q_{SMM} \times 100$	–	25.295	29.702	–	1.198	4.092	–	0.039	0.152
J_{pval}	–	0.000	0.000	–	0.055	0.000	–	0.748	0.858
Time	0.5	53	57	0.4	1	1.4	316	150	156

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 (Panel A) and January 2009 to December 2010 (Panel B). 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. Estimates labeled "SMM" are estimated using the identity weight matrix; estimates labeled "SMM-opt" are estimated using the efficient weight matrix.

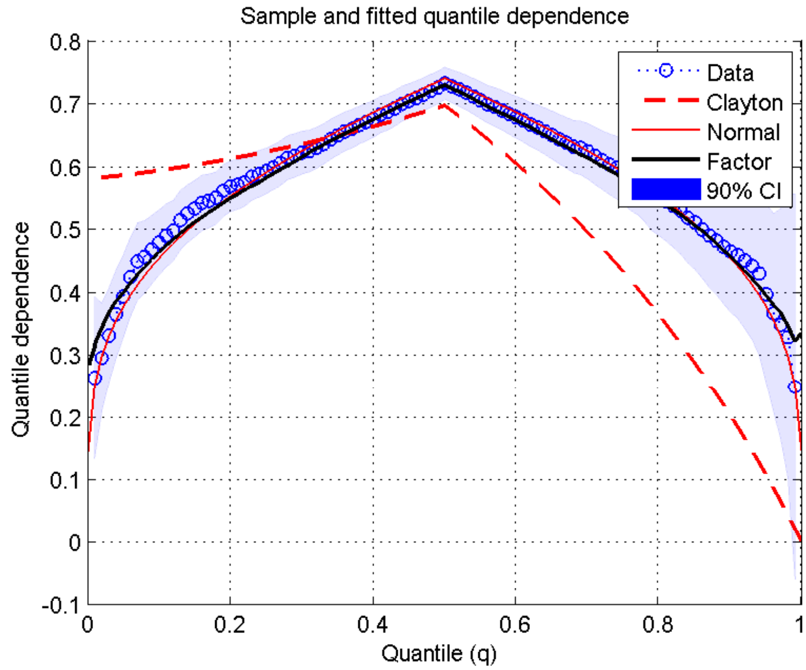

Figure 1: *This figure plots the probability of both variables being less than their q quantile (q<0.5) or greater than the q quantile (q>0.5). For the data this is averaged across all pairs, and a bootstrap 90% (pointwise) confidence interval is presented.*

.