ISSN 1440-771X

MONASH University

Australia

Department of Econometrics and Business Statistics

http://www.buseco.monash.edu.au/depts/ebs/pubs/wpapers/

Estimating the Error Distribution in Multivariate Heteroscedastic Time Series Models

Gunky Kim, Mervyn J Silvapulle and Paramsothy Silvapulle

June 2007

Working Paper 8/07

Estimating the error distribution in multivariate heteroscedastic time series models

By Gunky Kim, Mervyn J. Silvapulle, and Paramsothy Silvapulle

Department of Econometrics and Business Statistics, Monash University, P.O. Box 197, Caulfield East, Melbourne, Australia 3145.

gunky.kim@buseco.monash.edu.au, mervyn.silvapulle@buseco.monash.edu.au, param.silvapulle@buseco.monash.edu.au

Abstract

A semiparametric method is studied for estimating the dependence parameter and the joint distribution of the error term in a class of multivariate time series models when the marginal distributions of the errors are unknown. This method is a natural extension of Genest *et al.* (1995a) for independent and identically distributed observations. The proposed method first obtains $\sqrt{n}$-consistent estimates of the parameters of each univariate marginal time-series, and computes the corresponding residuals. These are then used to estimate the joint distribution of the multivariate error terms, which is specified using a copula. Our developments and proofs make use of, and build upon, recent elegant results of Koul and Ling (2006) and Koul (2002) for these models. The rigorous proofs provided here also lay the foundation and collect together the technical arguments that would be useful for other potential extensions of this semiparametric approach. It is shown that the proposed estimator of the dependence parameter of the multivariate error term is asymptotically normal, and a consistent estimator of its large sample variance is also given so that confidence intervals may be constructed. A large scale simulation study was carried out to compare the estimators particularly when the error distributions are unknown, which is almost always the case in practice. In this simulation study, our proposed semiparametric method performed better than the well-known parametric methods. An example on exchange rates is used to illustrate the method.

Some key words: Association; Copula; Estimating Equation; Pseudolikelihood; Semiparametric. (JEL CLASSIFICATION: C13, C14)

1 Introduction

The multivariate normal distribution continues to play a central role for modeling multivariate processes. There are several reasons that justifies such an important role for the normal distribution. For example, often the parameter of interest is a regression coefficient, and its maximum likelihood estimator is usually $\sqrt{n}$-consistent even if the assumption that the error term is normally distributed is incorrect. In these settings, the error distribution is treated as a nuisance parameter, and valid inferences could be made on the unknown finite dimensional parameters. By contrast, there is a range of statistical inference problems where the distribution of the error term is also of interest. In such cases, one does not wish to assume that the error distribution is normal, because the shape of the normal distribution is highly restrictive and hence is likely to be inappropriate for estimating the true error distribution.

Estimation of the joint distribution of the error term and learning about the interdependence among its components are important topics in statistical inference. This paper provides a rigorous treatment of a method for estimating the joint distribution of the error term in a class of time series models that includes nonlinear Generalized Autoregressive Conditional Heteroscedasticity [GARCH] models.

Let us consider an example to indicate the nature of the topic studied in this paper. The following example is a slightly modified version of that studied by Patton (2006). Let $E_{1,t}$ and $E_{2,t}$ denote the DM-USDollar and Yen-USDollar exchange rates respectively. Let $Y_{1t} = logE_{1,t} - logE_{1,t-1}$ and $Y_{2t} = logE_{2,t} - logE_{2,t-1}$. Thus, Y_{1t} and Y_{2t} can be seen as measures of returns of the two investments. We consider the model

$$\begin{aligned} Y_{1t} &= x_{1t}\beta_1 + \epsilon_{1t}, \quad & \epsilon_{1t} &= \sqrt{h_{1t}}\eta_{1t}, \quad & h_{1t} &= \alpha_{11} + \alpha_{12}\epsilon_{1,t-1}^2 + \alpha_{13}h_{1,t-1}, \\ Y_{2t} &= x_{2t}\beta_2 + \epsilon_{2t}, \quad & \epsilon_{2t} &= \sqrt{h_{2t}}\eta_{2t}, \quad & h_{2t} &= \alpha_{21} + \alpha_{22}\epsilon_{2,t-1}^2 + \alpha_{23}h_{2,t-1} \end{aligned} \tag{1}$$

where x_{1t} and x_{2t} are some exogenous variables and $\{(\eta_{1t}, \eta_{2t}) : t = 1, \ldots\}$ are *iid*. Thus, after accounting for the effect of the exogenous variables, η_{1t} and η_{2t} represent the undiversifiable risks, in other words, those that are not under the investor's control. For managing the risks associated with exchange rate fluctuations involving these currencies, the main quantities of interest are functions of the joint distribution of (η_{1t}, η_{2t}). For example, one such quantity is the probability of declines in both exchange rates. More generally, quantities of interest include, $\text{pr}(Y_{1t} \le a_1 \text{ and } Y_{2t} \le a_2 \mid x_t, \mathcal{F}_{t-1})$ and $\text{pr}(Y_{1t} \le a_1 \mid Y_{2t} \le a_2, x_t, \mathcal{F}_{t-1})$ where a_1 and a_2 are given numbers and $\mathcal{F}_{t-1}$ is the information set up to time $(t-1)$. Clearly estimation of these quantities require not only estimates of the α and β parameters in (1), but also the joint distribution of (η_{1t}, η_{2t}). Therefore, this inference problem is different from the standard ones arising in empirical studies involving the GARCH models where one's main interest involves only the parameter (α, β) in (1).

This type of inference problems, where estimates of the mean and variance functions together with that of the error distribution are required, abound in the financial statistics and risk management literature. In such cases, the typical joint distribution of the error term (η_{1t}, η_{2t}) is skewed and has long tails. Therefore, in the formulation of such inference problems, it is important that the joint distribution of (η_{1t}, η_{2t}) is not assumed to belong to a small parametric family such as the normal. In this area, *copulas* have been emerging as the natural objects for modeling the joint distribution of (η_{1t}, η_{2t}). First, it would be helpful to recall some elementary facts about copulas.

Let $H(x_1, \ldots, x_k)$ denote the joint cumulative distribution function of a random vector $(X_1, \ldots, X_k)$ with continuous marginals $F_p(x_p) = \mathrm{pr}(X_p \leq x_p)$. Then, H has the unique representation $H(x_1, \ldots, x_k) = C\{F_1(x_1), \ldots, F_k(x_k)\}$, where $C(u_1, \ldots, u_k)$ is the joint cumulative distribution of $(U_1, \ldots, U_k)$ and $U_p = F_p(X_p)$ is distributed uniformly on $[0, 1]$, $p = 1, \ldots, k$ (Sklar (1959)). The function C is called the *copula* of $(X_1, \ldots, X_k)$. There has been a substantial interest in the recent literature on copulas for studying multivariate observations. Two of the reasons for such increased interest includes the flexibility it offers because it can represent practically any shape for the joint distribution, and its ability to separate the intrinsic measures of association among $\{X_1, \ldots, X_k\}$ from their marginal distributions. For an excellent discussion of this see Genest and Favre (2007); for more comprehensive book-length treatments see (Joe (1997), Cherubini *et al.* (2004), and Nelsen (2006).

It is possible that the distribution functions H, $F_1, \ldots, F_k$, and C may belong to parametric families, for example, $H(x_1, \ldots, x_k; \alpha_1, \ldots, \alpha_k, \theta) = C\{F_1(x_1; \alpha_1), \ldots, F_k(x_k; \alpha_k); \theta\}$. In this case, θ is called the dependence parameter or association parameter. This helps to separate the parameters of the marginal distributions from their intrinsic association, which is nicely captured by θ. An attractive feature of this approach is that the copula C and the association parameter θ are invariant under continuous and monotonically increasing transformations of the marginal variables. Hence copulas have an advantage when the interest centers on the intrinsic association among the marginals (Joe (1997), Wang and Ding (2000), Oakes and Wang (2003)).

For the rest of this section, we shall restrict to bivariate copulas for simplicity of notation, but the results would extend in an obvious way to higher dimensions. If the joint *cdf* of (X_1, X_2) is $C(F_1(x_1), F_2(x_2); \theta)$, then the joint *pdf* of (X_1, X_2) is $c(F_1(x_1), F_2(x_2); \theta) f_1(x_1) f_2(x_2)$ where $c(\cdot)$ is the pdf corresponding to $C(\cdot)$. If F_1 and F_2 are also specified by parametric families, then the joint *pdf* would take the form $c(F_1(x_1; \alpha_1), F_2(x_2; \alpha_2); \theta) f_1(x_1; \alpha_1) f_2(x_2; \alpha_2)$, and maximum likelihood can be applied to estimate $(\alpha_1, \alpha_2, \theta)$. However, the more flexible Inference Function for Margin [IFM] has emerged as preferable to the method of maximum likelihood. In IFM, the parameters $\{\alpha_1, \alpha_2\}$ are estimated in the first stage, and then θ is estimated in the second stage; see Joe (1997) for an excellent account. If the marginal distributions were known, then, the loglikelihood would take the form

$$\Sigma_{i=1}^n \log c\{F_1(x_{1i}), F_2(x_{2i}); \theta\} + \Sigma_{i=1}^n \log\{f_1(x_{1i}) f_2(x_{2i})\}.$$

Since the marginal distributions are usually unknown, Genest *et al.* (1995a) developed an elegant theory for estimating θ by the semiparametric method that maximizes the pseudo-likelihood,

$$\Sigma_{i=1}^n \log c\{\hat{F}_1(x_{1i}, \hat{F}_2(x_{2i}); \theta\} + \Sigma_{i=1}^n \log\{\hat{f}_1(x_{1i}) \hat{f}_2(x_{2i}\}$$

which is obtained by simply substituting the empirical estimates for the marginal distributions. The possibility of the same method was also indicated in broad terms by Oakes (1994).

Genest *et al.* (1995a) proved that the asymptotic distribution of the resulting semiparametric estimator, which we denote by $\tilde{\theta}$, is normal. Kim *et al.* (2007) reported the results of an extensive simulation study to show that this method is better than its competitors. This method was further extended by Kim *et al.* (2005) to the multivariate regression model, $Y_{pi} = x'_{pi}\beta_p + \varepsilon_{pi}, (i = 1, \ldots, n, p = 1, 2)$, where the copula of (ϵ_1, ϵ_2) was assumed to have a known parametric form, denoted $C(u_1, u_2; \theta)$. They showed that if $\tilde{\beta}_p$, a preliminary $\sqrt{n}$-consistent estimator of β_p, is available for every p, then the method in Genest *et al.* (1995a) can be adapted to estimate the copula parameter and hence the joint distribution of $(\varepsilon_1, \varepsilon_2)$. In this method, the *cdf* of ε_p is estimated by the *cdf* of the residuals $\{\tilde{\varepsilon}_{p1}, \ldots, \tilde{\varepsilon}_{pn}\}$, where $\tilde{\varepsilon}_{pi} = Y_{pi} - x'_{pi}\tilde{\beta}_p$. Now, the semiparametric estimator of the copula parameter θ is $argmax_\theta \Sigma_{i=1}^n \log c\{\hat{F}_1(\tilde{\varepsilon}_{1i}), \hat{F}_2(\tilde{\varepsilon}_{2i}); \theta\}$.

It is clear for these discussions that the essence of the approach in Genest *et al.* (1995a) and Oakes (1994) is likely to be adaptable in more general settings involving multivariate regression and time series models where $C(u_1, u_2; \theta)$ is the copula of the vector of error terms, provided suitable preliminary estimates of the marginal distributions are available. In this paper, we study one such extension. In particular, our objective of is to provide a complete and rigorous proofs for the extension of the foregoing approach to a class of multivariate time series models and to provide simulation results to show that the proposed method is better than its fully parametric competitors. The proofs are nontrivial, and for this purpose we use recent results of Koul (2002) and Koul and Ling (2006). Their regularity conditions and results play a crucial role in our entire proofs.

Semiparametric estimation of the copula parameter has been studied in detail in Bagdonavicius *et al.* (2006) for the case when the marginal variable satisfy a generalized regression model, and essentially the same result of this paper is also used in Chen and Fan (2006) without proofs. Copulas have been used in a very wide range of areas and the literature is quite extensive indeed. The areas include survival analysis, analysis of current status data, censored data and finance (Bandeen-Roche and Liang (2002), Wang (2003), Wang and Ding (2000), Shih and Louis (1995), and Cherubini *et al.* (2004)). Joe (1997) and Nelsen (2006) provide comprehensive and authoritative accounts of statistical inference in copulas and dependence measures using copulas. Hutchinson and Lai (1990) provides an extensive range of practical examples where copulas are useful.

The plan for the rest of this paper is as follows. In the next section, we state the estimation method more formally. Section 3 presents simulation results to illustrate the superiority of the semiparametric method when the marginal distributions are unknown. Section 4 illustrates the method using a data example. Section 5 concludes. The proofs are given in appendix.

2 Main Results

Let $\{y_i : i = 0, \pm 1, \pm 2, \ldots\}$ denote a vector time series, where $y_i = (y_{1i}, \ldots, y_{ki})'$, and let the data generating process be

$$y_{pi} = \mu_{pi}(\alpha_{p1}) + \varepsilon_{pi} \qquad \text{and} \qquad \varepsilon_{pi} = \sqrt{h_{pi}(\alpha_p)}\eta_{pi}, \tag{2}$$

where $\alpha_p = (\alpha'_{p1}, \alpha'_{p2})'$, $\alpha_{p1} \in \Omega_{p1}$, $\alpha_{p2} \in \Omega_{p2}$, Ω_{p1} and Ω_{p2} are open subsets in Euclidean spaces, $\mu_{pi} : \Omega_{p1} \to \mathbb{R}$ and $h_{pi} : \Omega_p \to \mathbb{R}^+$ are known and twice continuously differentiable functions that may depend on past observations and covariates, and the standardized random variables $\{(\eta_{1i}, \ldots, \eta_{ki})\}$ are iid. Throughout, we shall assume that the series $\{y_{pi}\}$ is strictly stationary and ergodic, and that y_{p0} is independent of all previous observations. Let $\Omega_p = \Omega_{p1} \times \Omega_{p2}$, $\alpha = (\alpha_1^{'}, \ldots, \alpha_k^{'})'$ and α^0 denote the true parameter value of α.

Let F_p be the cumulative distribution function of $\{\eta_{pi}\}$ and let f_p denote the corresponding density function. Let $C(u;\theta)$ and $c(u;\theta)$ denote the copula of η and the corresponding density function, respectively, where $u = (u_1, \ldots, u_k)'$ and $\eta = (\eta_1, \ldots, \eta_k)'$. Let $\mathbb{F} = (F_1, \ldots, F_k)$, $\mathbb{F}(\eta_i) = (F_1(\eta_{1i}), \ldots, F_k(\eta_{ki}))$, and $\mathbb{F}(\eta) = (F_1(\eta_1), \ldots, F_k(\eta_k))$. For simplicity, we shall treat the case when θ is a scalar. The main results for the vector case are stated later.

Let us temporarily suppose that F_p and f_p are completely known. Now, the loglikelihood takes the form,

$$\ell^*(\theta, \alpha) = L^*(\theta; \alpha, \mathbb{F}) + B(\alpha, f_1, \ldots, f_k)$$

where $L^*(\theta;\alpha,\mathbb{F}) = \sum \log c\{\mathbb{F}(\eta_i);\theta\}$ and $B(\alpha, f_1, \ldots, f_k) = \sum \log\{f_1(\eta_{1i}) \ldots f_k(\eta_{ki})\}$. By standard results, the maximum likelihood estimator of (θ, α), which is $Argmax\ \ell^*(\theta, \alpha)$, is consistent and asymptotically normal provided the model is correctly specified.

Now, let us relax the temporary assumption and suppose that $(F_1, \ldots, F_k)$ is unknown, which is usually the case in practice. In this case we modify the foregoing method for estimating θ. Let $F_{pn}(t) = (n+1)^{-1}\Sigma_{i=1}^{n} I(\eta_{pi} \le t)$, the rescaled *empirical distribution function* [edf] of $\{\eta_{pi} : i = 1, \ldots, n\}$. The difference between F_{pn} and the usual empirical distribution function is that the denominator is $n+1$ instead of n.

We assume that the time series model (2) can be estimated for each margin separately, and let $\tilde{\alpha}_p$ denote a $\sqrt{n}-$consistent estimator of α_p^0, $(p = 1, \ldots, k)$. Now, the residuals corresponding to $\{\eta_{pi}\}$ can be estimated by

$$\tilde{\eta}_{pi} = [y_{pi} - \mu_{pi}(\tilde{\alpha}_{p1})]/\sqrt{h_{pi}(\tilde{\alpha}_p)}. \tag{3}$$

Let $\tilde{F}_{pn}(t) = (n+1)^{-1}\Sigma_{i=1}^{n} I(\tilde{\eta}_{pi} \le t)$ and $\tilde{\mathbb{F}}_n(\eta_i) = (\tilde{F}_{1n}(\eta_{1i}), \ldots, \tilde{F}_{1n}(\eta_{1i}))$; thus, $\tilde{F}_{pn}(t)$ is the *edf* of the residuals $\{\tilde{\eta}_{pi}\}$. Now, a natural generalization of the approach in Genest *et al.* (1995a) and Oakes (1994) for iid observations suggests that an estimate $\tilde{\theta}$ of θ is given by

$$\tilde{\theta} = \text{argmax}_\theta\ L(\theta) \quad \text{where} \quad L(\theta) = \Sigma \log c\{\tilde{\mathbb{F}}_n(\tilde{\eta}_i);\theta\}. \tag{4}$$

Since $\tilde{\mathbb{F}}_n$ is expected to be close to $\mathbb{F}$ for large n, we would expect $L(\theta)$ to be close to the loglikelihood, $\Sigma \log c\{\mathbb{F}(\eta_i);\theta\}$, except for a constant term and hence one would expect that $\tilde{\theta}$ is likely to be a reasonable estimator of θ.

We will show that $\tilde{\theta}$ is consistent and asymptotically normal, propose a consistent estimator of its large sample variance, and evaluate these by simulation studies.

Theorem 1. *Suppose that the regularity condition A in the Appendix is satisfied. Let $\tilde{\alpha}$ be a $\sqrt{n}$-consistent estimator of α and let $\tilde{\theta}$ denote the estimator of θ defined in (4). Let $\ell(\theta, u) = \log c(u_1, \ldots, u_k; \theta)$, $\ell_\theta(\theta, u) = (\partial/\partial\theta)\ell(\theta, u)$, and $\ell_{\theta,p}(\theta, u) = (\partial^2/\partial\theta\partial u_p)\ell(\theta, u_1, \ldots, u_k)$. Then, $\tilde{\theta}$ is a consistent estimator of θ_0 and the asymptotic distribution of $n^{1/2}(\tilde{\theta} - \theta_0)$ is $N(0, \nu^2)$, where $\nu^2 = \sigma^2/\gamma^2$, $\gamma = -E\big[l_{\theta,\theta}\{\theta_0, \mathbb{F}(\eta)\}\big]$,*

$$\begin{aligned} \sigma^2 &= var\big[l_\theta\{\theta_0, \mathbb{F}(\eta)\} + W_1(\eta_1) \ldots + W_k(\eta_k)\big], \\ \text{and} \qquad W_p(\eta_p) &= \int I\big(F_p(\eta_p) \le u_p\big) l_{\theta,p}(\theta_0, u) dC(u; \theta_0). \end{aligned} \tag{5}$$

Further, a consistent estimator $\tilde{\nu}^2$ of ν^2 is given by $\tilde{\nu}^2 = \tilde{\sigma}^2/\tilde{\gamma}^2$, where

$$\begin{aligned} \tilde{\gamma} = -n^{-1}\Sigma_{i=1}^n l_{\theta,\theta}\{\tilde{\theta}, \tilde{\mathbb{F}}_n(\tilde{\eta}_i)\}, &\quad \tilde{\sigma}^2 = \textit{ Sample variance of } \{\tilde{T}_1(\tilde{\theta}), \ldots, \tilde{T}_n(\tilde{\theta})\}, \\ \tilde{T}_i(\theta) = l_\theta\{\theta, \tilde{\mathbb{F}}_n(\tilde{\eta}_i)\} + n^{-1}\Sigma_{p=1}^k W_p(\tilde{\eta}_{pi}, \theta), &\; i = 1, \ldots, n, \\ \text{and} \quad \tilde{W}_p(\tilde{\eta}_{pi}, \theta) = n^{-1}\Sigma_{j=1}^n I\big(\tilde{F}_{pn}(\tilde{\eta}_{pi}) \le \tilde{F}_{pn}(\tilde{\eta}_{pj})\big)&\; l_{\theta,p}\{\theta, \tilde{\mathbb{F}}_n(\tilde{\eta}_j)\}. \end{aligned} \tag{6}$$

Genest *et al.* (1995a) showed that the foregoing semiparametric method for *iid* observations is fully efficient for the independent copula. Since ν^2 is the same as that for the iid setting, it follows that the foregoing result of Genest *et al.* (1995a) holds for the time series model in (2).

Theorem 1 holds when θ is a vector parameter as well. This is stated below. The proof is essentially the same as that for the case when θ is a scalar.

Theorem 2. *Let the setting be as in Theorem 1 except that θ is a vector. Let $\phi(u; \theta)$ denote $(\partial/\partial\theta)\log\{c(u; \theta)\}$. Then, $n^{1/2}(\tilde{\theta} - \theta_0)$ converges in distribution to $N(0, V)$ where $V = B^{-1}\Sigma B^{-\prime}$, $B = \int(\partial/\partial\theta')\phi(u; \theta_0)dC(u; \theta_0)$, and Σ is the covariance matrix of the m-dimensional random vector whose jth component is*

$$\phi_j\{\mathbb{F}(\eta); \theta_0\} + \Sigma_{p=1}^k \int I\{F_p(\eta_p) \le u_p)(\partial/\partial u_p)\phi_j(u; \theta_0)dC(u; \theta_0), \; j = 1, \ldots, m.$$

An estimator of the covariance matrix V is $\hat{V} = \hat{B}^{-1}\hat{\Sigma}\hat{B}^{-\prime}$ where $\hat{\Sigma}$ is the sample covariance matrix of $\{\tilde{T}_1^\phi(\tilde{\theta}), \ldots, \tilde{T}_n^\phi(\tilde{\theta})\}$, $\tilde{T}_i^\phi(\theta) = \phi\{\tilde{\mathbb{F}}_n(\tilde{\eta}_i); \theta\} + \tilde{W}_1^\phi(\tilde{\eta}_{1i}, \theta) + \ldots + \tilde{W}_k^\phi(\tilde{\eta}_{ki}, \theta)$, $\tilde{W}_p^\phi(\tilde{\eta}_{pi}, \theta) = n^{-1}\Sigma_{j=1}^n I(\tilde{\eta}_{pi} \le \tilde{\eta}_{pj})(\partial/\partial u_p)\phi\{\tilde{\mathbb{F}}_n(\tilde{\eta}_j), \theta\}$, and $\hat{B} = n^{-1}\Sigma_{i=1}^n(\partial/\partial\theta')\phi\{\tilde{\mathbb{F}}_n(\tilde{\eta}_i); \tilde{\theta}\}$.

3 Simulation Study

Since the main results on the properties of the proposed semiparametric method are asymptotic, a large scale simulation study was carried out to compare its properties with other competing ones, namely MLE and IFM.

Design of the study:

The models for the two margins are: $y_{1t} = h_{1t}\eta_{1t},\ h_{1t} = 0.2 + 0.1y_{1,t-1} + 0.15h_{1,t-1}$ and $y_{2t} = h_{2t}\eta_{2t},\ h_{2t} = 0.1 + 0.1y_{2,t-1} + 0.15y_{2,t-1} + 0.1h_{1,t-1}$.

The following copulas were considered in the study.

(1) *Ali-Mikhail-Haq [AMH] Family of copulas:* $C(u,v;\theta) = uv/\{1-\theta(1-u)(1-v)\}$.

(2) *Frank copula:* $C(u,v;\theta) = -\theta^{-1}\log\left([1+(e^{-\theta u}-1)(e^{-\theta v}-1)]/(e^{-\theta}-1)\right)$

(3) *Gumbel copula:* $C(u,v;\theta) = \exp-\left((-\log u)^{\theta} + (-\log v)^{\theta}\right)^{\frac{1}{\theta}}$

(4) *Joe copula:* $C(u,v;\theta) = 1-\left((1-u)^{\theta}+(1-v)^{\theta}-(1-u)^{\theta}(1-v)^{\theta}\right)^{\frac{1}{\theta}}$

(5) *Plackett copula:*

$C(u,v;\theta) = [1+(\theta-1)(u+v)-\{\{(1+(\theta-1)(u+v)\}^2-4\theta(\theta-1)uv\}^{\frac{1}{2}}]/\{2(\theta-1)\}$.

(6) *Gaussian copula:*

$C(u,v;\theta) = \int_{-\infty}^{\Phi^{-1}(u)}\int_{-\infty}^{\Phi^{-1}(v)}\{2\pi(1-\theta^2)^{1/2}\}^{-1}\exp[\{2(1-\theta^2)\}^{-1}(2rst-s^2-t^2)]^2 dsdt$

(7) *Student-t copula:*

$C(u,v;\theta) = \int_{-\infty}^{t_\nu^{-1}(u)}\int_{-\infty}^{t_\nu^{-1}(v)}\{2\pi(1-\theta^2)^{1/2}\}^{-1}[1+\{(s^2+t^2-2\theta st)/\{\nu(1-\theta^2)\}]^{-(\nu+2)/2}dsdt$

Error Distributions: The following pairs of marginal distributions were considered for the error distribution: (1) η_1 and η_2 are normally distributed, (2) $\eta_1 \sim t_r$ and $\eta_2 \sim t_r$, (3) $\eta_1 \sim t_r$ and $\eta_2 \sim$ skew t_r with skewness $= 0.5$, and (4) $\eta_1 \sim t_r$ and $\eta_2 \sim \chi^2_5$. The values 3 and 8 were considered for r in t_r.

The IFM and ML methods in this simulation assume that the marginal error distributions are all normal. Therefore, the first pair of error distributions corresponds to the correct specification of the marginal distributions, while each of the other three leads to a misspecification of the model. A skewed-t_r distribution has tails that are of the same order as that for t_r but the probability masses on either sides of the origin are different, leading to skewness. Since the semiparametric method estimates each marginal distribution nonparametrically, it is meant to be used when the sample size is moderate to large. This is not a concern because data sets used for modeling GARCH type models are usually very large. For such models, 100 observations is of course very small. In this study, we considered sample sizes 100 and 500. This captures a broad range of realistic settings. All the computations were programmed in MATLAB Version 7.0.4 and optimizations were performed using the procedure "fmincon.m" in the "Optimization Toolbox (3.0.2).

The values for the copula parameter θ were chosen so that the corresponding Kendall's tau takes the values 0.2, 0.5 and 0.8, except for the Ali-Mikhail-Haq family because its tau value cannot be as high as 0.5. Therefore, for this copula we chose the τ values -0.1, 0.1 and 0.2. The foregoing seven copulas cover a very wide range of distributional shapes for the errors, and the error distributions also cover a very broad range. Thus, the design of the simulation covers a broad range of realistic settings.

Results:

The differences between the performance of IFM and ML estimators were small, with the IFM performing slightly better. Since IFM is now well accepted in the copula literature, we evaluate the performance of the semiparametric method relative to the IFM method.

Each marginal distribution is correctly specified as normal:

The results are given in Table 1 under the heading N-N. Since the marginal distributions and the copula are correctly specified, there is no misspecification for the IFM method. Consequently, as expected, the IFM estimator performs better than the semiparametric estimator. However, the differences between these estimators are small. These results show that, even if the error distribution is known, the use of the semiparametric method in Theorem 1, which ignores the fact that the error distribution is known, does not suffer significant loss of efficiency.

Each marginal distribution is incorrectly specified as Normal:

Table 2 provides means of the simulated estimates of θ when the two marginal error distributions are T_3 and χ^2_5. The same table also provides standard deviations of the simulated estimates of θ. Table 2 shows that if the error distribution is misspecified then the distribution of the IFM estimator may be centered at a value that is quite different from the true value of θ. Tables 1 and 2 show quite clearly that (i) the IFM estimator is highly nonrobust against misspecification of the marginal distributions, and (ii) the distribution of the semiparametric estimator is centered around the true value of θ and is far superior to the IFM estimator of θ. The very large values for relative MSE in Table 1 reflect the fact that misspecification of the marginal distribution may result in the IFM estimator being inconsistent, and consequently, turns out to be substantially worse than the semiparametric estimator.

The coverage rates of a large sample 95% confidence interval based on a normal approximation for the large sample distribution of $\tilde{\theta}$ given in Theorem 1, are provided in Table 3. This table shows that the coverage rates are close to 95% in most cases. The coverage rates tend to drop for some copulas when the parameter is close to the boundary. Overall, these results show that the semiparametric method offers a reliable and easy to compute large sample confidence interval for θ.

In summary, the semiparametric estimator $\tilde{\theta}$ is better than the parametric ML and IFM estimators, and Theorem 1 provides the main results for implementing this semiparametric method for statistical inference.

4 An Example

In this section we briefly discuss an example to illustrate the semiparametric method proposed in section 2. This is a simpler version of the model developed and studied in detail by Patton (2006), to which we refer the readers for detailed discussions about the practical aspects of the problem. Let Y_{1t} and Y_{2t} denote the log difference of DM-USD and Yen-USD exchange rates respectively, as defined in the Introduction. We consider the model

$$\begin{aligned} Y_{1t} &= \mu_1 + \epsilon_{1t}, \quad & \epsilon_{1t} &= \sqrt{h_{1t}}\eta_{1t}, \quad & h_{1t} &= \alpha_{11} + \alpha_{12}\epsilon^2_{1,t-1} + \alpha_{13}h_{1,t-1}, \\ Y_{2t} &= \mu_2 + \epsilon_{2t}, \quad & \epsilon_{2t} &= \sqrt{h_{2t}}\eta_{2t}, \quad & h_{2t} &= \alpha_{21} + \alpha_{22}\epsilon^2_{2,t-1} + \alpha_{23}h_{2,t-1} \end{aligned} \tag{7}$$

We assume that $\{\eta_{1t}, \eta_{2t}\}$ are *iid*, and let $C(u_1, u_2; \theta)$ denote their copula. The main purpose of the methodology introduced in the earlier sections was to estimate θ, and to estimate the joint distribution of $\{\eta_{1t}, \eta_{2t}\}$ when its marginal distributions are unknown.

We use the data for the period Jan 1991 - Jan 1999, which is the period prior to the introduction of the Euro. The total number of observations is 2046. Thus, we have a reasonably large number of observations. We estimated several copulas. The estimates and their standard errors are given in Table 4. To assess the goodness of fit of the estimated copulas, we adopted a method similar to that in Patton (2006). We applied a chi-square goodness of fit with the unit square, the support of the copula, divided into 100 cells. To this end, we adopted the method in Junker and May (2005). The cells were formed by gridlines parallel to the two axes. Each cell with small count was merged with a neighbouring one. The expected count for each cell was estimated by substituting the estimates for the copula parameters. Since such an estimation involves a nonparametric estimator namely the empirical distribution functions of the marginals, it is not clear that the chi-square statistic would be approximately chi-square distributed. In any case, we computed the p-values corresponding to a chi-square distribution. These are given in Table 4. Since the sample size is large, it is likely that the p-values in Table 4 are reasonably reliable. Despite this caveat, these p-values can be compared among themselves to rank the models in terms of the goodness of fit. The indications are that the Gaussian copula provides the best fit.

For the Gaussian copula, the estimated distribution of (η_{1t}, η_{2t}), conditional on $\mathcal{F}_{t-1}$ is $C(\tilde{F}_{1n}(\tilde{\eta}_1), \tilde{F}_{2n}(\tilde{\eta}_2); 0.53)$ where $\tilde{F}_{1n}$ and $\tilde{F}_{2n}$ are the empirical distribution functions of the residuals $\{\eta_{1t}\}$ and $\{\eta_{2t}\}$ respectively, and $C(u_1, u_2; \theta)$ is the Gaussian copula. Based on this estimated joint distribution, we can write down the estimated joint distribution of (y_{1t}, y_{2t}) conditional on the history up to time $t-1$. This estimated joint distribution can then be used for estimating various quantities of interest. To illustrate this, we consider the following two quantities: $A(c) = P\{E_{1,T} \leq cE_{1,T-1}, E_{2,T} \leq cE_{2,T-1} \mid \mathcal{F}_{T-1}\}$ and $B(c) = P\{E_{1,T} \leq cE_{1,T-1} \mid E_{2,T} \leq cE_{2,T-1}, \mathcal{F}_{T-1}\}$, where $c > 0$. Thus, $A(1)$ is the probability that both exchange rates fall at time T given the history up to that time, and $A(0.9)$ is the probability that both exchange rates fall to 90% of the value on the previous day, given the history up to that time. Similarly, $B(1)$ is the probability that the first exchange rate falls at time T given that the second has fallen and the history up to that time. Substituting directly into the estimated distribution of (y_{1t}, y_{2t}), we obtain the estimates 0.333, 0.247, and 0.106 for $A(c)$ for $c = 1$, 0.9 and 0.7 respectively. Similarly, estimates of $B(c)$ for $c = 1$, 0.9 and 0.7 are 0.498, 0.425, and 0.268 respectively. Note that these estimates are obtained without assuming any parametric functional form for the marginal error distributions. Theorem 1 says that this is a valid method. This is in contrast to other more familiar methods that would assume particular parametric forms such as bivariate normal or bivariate t-distributions. Such a method would not be valid if the assumed distribution is incorrect, which is most likely to be the case in practice. This exemplifies the importance of the semiparametric method proposed in this paper.

5 Conclusion

We developed a semiparametric method for estimating the dependence parameter and the joint distribution of the error terms in a class of multivariate nonlinear GARCH models. The method has two stages. In the first stage, parameters of the GARCH model are estimated separately for each margin. In the second stage, the empirical distribution of the residuals is used as an estimate of the true error distribution for each margin, which in turn is substituted in the likelihood function to estimate the copula parameter. The nonparametric part of the semiparametric method estimates the error distribution of each margin by the empirical of the residuals. Consequently, the method does not require knowledge of the marginal error distributions. This is a very appealing feature of this method.

We showed that the proposed semiparametric estimator is asymptotically normal. It turns out that the form of the asymptotic variance is a natural extension of that obtained by Genest *et al.* (1995a) for the case when the observations are independent and identically distributed. This helped us to use their results and construct consistent estimates for the asymptotic variance and confidence interval for the dependence parameter. Simulation results showed that our semiparametric estimator performs significantly better than the parametric ones when the true error distribution deviates from that assumed by the parametric methods, maximum likelihood and inference function for margins. This is important because the true error distribution is usually unknown in practice. We conclude that the semiparametric method proposed here is better than its main competitors, namely maximum likelihood and inference function for margins, for estimating the joint distribution of the error terms in linear and nonlinear GARCH models.

6 Appendix

As in the text, the index p refers to the pth component. For notational simplicity, we provide the proof for the bivariate case and hence $p = 1$ or 2. However, only minor changes to notation are needed for the higher dimensional case. For example, the loglikelihood $l(\theta, u_1, u_2)$ for the bivariate case would need to be written as $l(\theta, u_1, \ldots, u_k)$ for the multivariate case. For simplicity of notation/expression, we shall avoid writing 'for every p' or '$p = 1, 2$', as far as possible.

Let $H(\theta, u_1, u_2)$ denote a derivative of $l(\theta, u_1, u_2)$ up to third order in θ and second order in (u_1, u_2). Let (U_1, U_2) denote a random vector with the same distribution as that of $(F_1(\eta_1), F_2(\eta_2))$ so that $(U_1, U_2) \sim C(u_1, u_2; \theta_0)$. For any function $g(x)$, let $\dot{g}(x) = (\partial/\partial x)g(x)$ and let $\|g\| = \sup_x |g(x)|$. To simplify notation, we shall write $\mu_{pi}(\alpha_p)$ for $\mu_{pi}(\alpha_{p1})$ so that μ_{pi} is treated as a function of the same vector parameter α_p which is defined as $(\alpha'_{p1}, \alpha'_{p2})'$. Thus, we have $\dot{\mu}_{pi}(\alpha_p) = ((\partial/\partial\alpha'_{p1})\mu_{pi}(\alpha_{p1}), 0')'$ and $\dot{h}_{pi}(\alpha_p) = ((\partial/\partial\alpha'_{p1})h_{pi}(\alpha_{p2}), (\partial/\partial\alpha'_{p2})h_{pi}(\alpha_{p2}))'$.

Let

$$a_{np,i} = \dot{\mu}_{pi}(\alpha_p^0)/\sqrt{h_{pi}(\alpha_p^0)} \quad \text{and} \quad b_{np,i} = \dot{h}_{pi}(\alpha_p^0)/\{2h_{pi}(\alpha_p^0)\}. \tag{8}$$

The variables $a_{np,i}$ and $b_{np,i}$ may be seen as standardized forms or explanatory variables for the mean and the variance functions, respectively, of y_{pi}.

Now, let us introduce the following regularity conditions.

Condition A:

(A.1): The distribution function F_p has continuously differentiable density, denoted by f_p and it satisfies $\sup_x |xf_p(x)| < \infty$ and $\sup_x |\dot{f}_p(x)| < \infty$.

(A.2): There exist a function $G(u_1, u_2)$ such that $|H(\theta, u_1, u_2)| \leq G(u_1, u_2)$ in a small neighbourhood of θ_0, and $E\{G^2(U_1, U_2)\} < \infty$.

(A.3): Let $\Psi(\theta, u_1, u_2)$ denote $H(\theta, u_1, u_2)$ or $G(u_1, u_2)$. Then, for any given θ, there exist $k(u_1, u_2; \theta)$ and $\varepsilon_\theta > 0$ such that $E\{k^2(u_1, u_2; \theta_0)\} < \infty$ and satisfies $|\Psi(\theta, u_1 + d_1, u_2 + d_2) - \Psi(\theta, u_1, u_2)| \leq k(u_1, u_2; \theta)(|d_1| + |d_2|)$, for any u_1, u_2, and $|d_j| \leq \varepsilon_\theta$.

(A.4): The conditions of Proposition A.1 in Genest *et al.* (1995a) are satisfied.

(A.5): $\max_{1\leq i\leq n} n^{-1/2}\|a_{np,i}\| = o_p(1)$ and $\max_{1\leq i\leq n} n^{-1/2}\|b_{np,i}\| = o_p(1)$. This is the same as condition (8.3.10) on page 384 in Koul (2002).

(A.6): This is the same as (8.3.2) and (8.3.3) on page 381 of Koul (2002).

$$\sup n^{1/2} \mid \mu_{pi}(t) - \mu_{pi}(s) - (t-s)'\dot{\mu}_{pi}(s) \mid / \{h_{pi}(\alpha_p^0)\}^{1/2} = o_p(1),$$
$$\sup\ n^{1/2} \mid \{h_{pi}(t)\}^{1/2} - \{h_{pi}(s)\}^{1/2} - (t-s)'\dot{h}_{pi}(s) \mid / \{h_{pi}(\alpha_p^0)\}^{1/2} = o_p(1),$$

where, the supremum is taken over $1 \leq i \leq n$, and over all t and s in the parameter space for α_p satisfying $n^{1/2}\|t - s\| \leq K$, for some $K < \infty$.

(A.7): For $t, s \in \Omega_p$ and $\forall \epsilon > 0$, $\exists \delta > 0$, and an $n_1 \ni \forall 0 < b < \infty, \forall ||s|| < b$, $\forall n > n_1$,

$$P\Big(n^{-1/2}\Sigma_{i=1}^n\Big\{\sup_{||t-s||<\delta}\frac{|\mu_{pi}(t) - \mu_{pi}(s)|}{\sqrt{h_{pi}(\alpha_p^0)}} + \sup_{||t-s||<\delta}\frac{|\sqrt{h_{pi}(t)} - \sqrt{h_{pi}(s)}|}{\sqrt{h_{pi}(\alpha_p^0)}}\Big\} \leq \epsilon\Big) > 1 - \epsilon.$$

This is the same as condition (4.12) of Koul and Ling (2005).

(A.8): $n^{-1}\Sigma_{i=1}^n\Big\{\|a_{np,i}\|^2 + \|b_{np,i}\|^2\Big\} = O_p(1)$. This is similar to condition (4.11) of Koul and Ling (2005).

(A.9) Let $\bar{a}_{np} = n^{-1}\Sigma_{i=1}^n a_{np,i}$ and $\bar{b}_{np} = n^{-1}\Sigma_{i=1}^n b_{pn,i}$. Then $n^{-1}\Sigma_{j=1}^n(a_{np,j} - \bar{a}_{np})g(\eta_j) \xrightarrow{p} 0$ and $n^{-1}\Sigma_{j=1}^n(b_{np,j} - \bar{b}_{np})g(\eta_j) \xrightarrow{p} 0$, where $g(.)$ is a given function such that $E[g^2(\eta)] < \infty$.

Conditions (A.1) - (A.4) do not involve the time-series aspects of the model. They were also used in Kim *et al.* (2005) for the linear regression case with iid errors. The conditions

(A.5)-(A.8) are taken from earlier work by Koul (2002) and Koul and Ling (2006). Condition (A.9) is a mild one. For example, if μ_{pi} is a function of past values of the time series which is strictly stationary and ergodic, then the summand forms a strictly stationary and ergodic process with mean zero, and hence (A.9) would be satisfied (see Taniguchi and Kakizawa (2000), Theorems 1.3.3 - 1.3.5).

In what follows, we shall assume that Condition A is satisfied, as in Theorem 1. Now we provide a proof of Theorem 1 by establishing several lemmas. For $t \in \Omega_p$, the parameter space of α_p, let

$$\begin{aligned} u_{npi}(t) &= \{h_{pi}(\alpha_p^0)\}^{-1/2}[\mu_{pi}(\alpha_p^0 + n^{-1/2}t) - \mu_{pi}(\alpha_p^0)], \\ \text{and} \quad v_{npi}(t) &= \{h_{pi}(\alpha_p^0)\}^{-1/2} h_{pi}(\alpha_p^0 + n^{-1/2}t)^{1/2} - 1. \end{aligned} \tag{9}$$

Lemma 1. *Let $\bar{\alpha}_p, \tilde{\alpha}_p$, and α_p^* be $\sqrt{n}-$consistent estimators of α_p^0, and let $\{\bar{\eta}_i\}, \{\tilde{\eta}_i\}$, and $\{\eta_i^*\}$ be the corresponding residuals so that $y_{pi} = \mu_{pi}(\bar{\alpha}_p) + \sqrt{h_{pi}(\bar{\alpha}_p)}\ \bar{\eta}_{pi}$, $y_{pi} = \mu_{pi}(\tilde{\alpha}_p) + \sqrt{h_{pi}(\tilde{\alpha}_p)}\ \tilde{\eta}_{pi}$, and $y_{pi} = \mu_{pi}(\alpha_p^*) + \sqrt{h_{pi}(\alpha_p^*)}\ \eta_{pi}^*$. Let $\tilde{t} = n^{1/2}(\tilde{\alpha}_p - \alpha_p^0)$. Then*

$$\sup_{1\le i\le n} \|u_{npi}(\tilde{t}) - (\tilde{\alpha}_p - \alpha_p^0)' a_{np,i}\| = o_p(n^{-1/2}), \ \sup_{1\le i\le n} \|v_{npi}(\tilde{t}) - (\tilde{\alpha}_p - \alpha_p^0)' b_{np,i}\| = o_p(n^{-1/2})$$

$$\sup_{1\le i\le n} |u_{npi}(\tilde{t})|, \quad \sup_{1\le i\le n} |v_{npi}(\tilde{t})| = o_p(1)$$

$$\sup_{1\le i\le n} |(\tilde{\eta}_{pi} - \eta_{pi}) f_p(\eta_{pi})| = o_p(1), \ \sup_{1\le i\le n} |(\tilde{\eta}_{pi} - \eta_{pi}) f_p(\tilde{\eta}_{pi})| = o_p(1)$$

$$\sup_{1\le i\le n} |(\tilde{\eta}_{pi} - \bar{\eta}_{pi}) f_p(\bar{\eta}_{pi})| = o_p(1), \ \sup_{1\le i\le n} |(\tilde{\eta}_{pi} - \eta_{pi}) f_p(\bar{\eta}_{pi})| = o_p(1).$$

$$\sup_{1\le i\le n} |(\tilde{\eta}_{pi} - \bar{\eta}_{pi}) f_p(\eta_{pi}^*)| = o_p(1), \ \sup_{1\le i\le n} |F_p(\tilde{\eta}_{pi}) - F_p(\eta_{pi})| = o_p(1).$$

Proof. The first part follows from Condition (A.6). Now, the second part follows from Condition (A.5). To prove the third part, let $C_{npi} = \{h_{pi}(\tilde{\alpha}_p)/h_{pi}(\alpha_p^0)\}^{1/2}$. By substituting directly from the definitions, it may be verified that

$$\tilde{\eta}_{pi} - \eta_{pi} = -C_{npi}^{-1}[u_{npi}(\tilde{t}) + v_{npi}(\tilde{t})\eta_{pi}] \text{ and } \sup_{1\le i\le n} \left|C_{npi}^{-1}\right| = 1 + o_p(1). \tag{10}$$

Now, we have

$$\begin{aligned} \sup_{1\le i\le n} |(\tilde{\eta}_{pi} - \eta_{pi}) f_p(\eta_{pi})| &\le \sup_{1\le i\le n} \left|C_{npi}^{-1}\right| \sup_{1\le i\le n} \left|u_{npi}(\tilde{t})\right| \ ||f_p|| \\ &+ \sup_{1\le i\le n} \left|C_{npi}^{-1}\right| \sup_{1\le i\le n} \left|v_{npi}(\tilde{t})\right| \sup_{1\le i\le n} |\eta_{pi} f_p(\eta_{pi})| = o_p(1). \end{aligned} \tag{11}$$

The rest of the proofs follow by similar arguments and from the following identities with $\bar{t} = n^{1/2}(\bar{\alpha}_p - \alpha_0)$:

$$\tilde{\eta}_{pi} - \eta_{pi} = -[u_{npi}(\tilde{t}) + v_{npi}(\tilde{t})\tilde{\eta}_{pi}], \tag{12}$$

$$\tilde{\eta}_{pi} - \bar{\eta}_{pi} = -C_{npi}^{-1}[\{u_{npi}(\tilde{t}) - u_{npi}(\bar{t})\} + \{v_{npi}(\tilde{t}) - v_{npi}(\bar{t})\}\bar{\eta}_{pi}]. \tag{13}$$

□

Lemma 2. $\sup_{1\le i\le n}\Big|n^{1/2}\Big\{F_{pn}(\tilde{\eta}_{pi}) - F_{pn}(\eta_{pi})\Big\} - n^{1/2}\big[F_p(\tilde{\eta}_{pi}) - F_p(\eta_{pi})\big]\Big| = o_p(1).$

Proof. Let $W(t) = \Sigma_{i=1}^n[I\{F_p(\eta_{pi}) \le t\} - t]$. Then, we have

$$\sup_{1\le i\le n}\Big|n^{1/2}\Big\{F_{pn}(\tilde{\eta}_{pi}) - F_{pn}(\eta_{pi})\Big\} - n^{1/2}\big[F_p(\tilde{\eta}_{pi}) - F_p(\eta_{pi})\big]\Big|$$
$$= \sup_{1\le i\le n}\Big|W\{F_p(\tilde{\eta}_{pi})\} - W\{F_p(\eta_{pi})\}\Big| \le \sup_{|t-s|<\delta}|W(t) - W(s)|$$

with arbitrary large probability, by Lemma 1. Now, the desired result follows from Theorem 2.2.1 of Koul (2002). □

Lemma 3. *Let* $\tilde{t} = n^{1/2}(\tilde{\alpha}_p - \alpha_p^0)$. *Then*

$$\sup_{x\in\mathbb{R}}\Big|\sqrt{n}\Big\{\tilde{F}_{pn}(x) - F_{pn}(x) - n^{-1}\Sigma_{j=1}^n\Big[F_p\big(x + xv_{npj}(\tilde{t}) + u_{npj}(\tilde{t})\big) - F_p(x)\Big]\Big\}\Big| = o_p(1),$$

$$\sup_{1\le i\le n}\Big|\sqrt{n}\Big\{\tilde{F}_{pn}(\tilde{\eta}_{pi}) - F_{pn}(\tilde{\eta}_{pi}) - n^{-1}\Sigma_{j=1}^n\Big[F_p\big(\tilde{\eta}_{pi} + \tilde{\eta}_{pi}v_{npj}(\tilde{t}) + u_{npj}(\tilde{t})\big) - F_p(\tilde{\eta}_{pi})\Big]\Big\}\Big| = o_p(1).$$

Proof. To prove this, we shall apply Lemma 4.1 Koul and Ling (2005) with their $\ell_{ni}(t) = 1$. By (A.6) and (A.8), we have that

$$n^{-1/2}\Sigma_{i=1}^n\Big|\Big|\Big\{|v_{npi}(\tilde{t})| + |u_{npi}(\tilde{t})|\Big\} =$$
$$\le n^{1/2}\|\tilde{\alpha}_p - \alpha_p^0\|\ \Big[n^{-1}\Sigma_{i=1}^n\Big\{\|a_{np,i}\| + \|b_{np,i}\|\Big\}\Big] + o_p(1) = O_p(1).$$

Now, $I(\tilde{\eta}_{pj} \le x) = I[\{y_{p,j} - \mu_{p,j}(\tilde{\alpha}_p)\}/h_{p,j}(\tilde{\alpha}_p)^{1/2} \le x] = I\big(\eta_{pj} \le x + x\ v_{npj}(\tilde{t}) + u_{npj}(\tilde{t})\big)$. Therefore, by Lemma 4.1 of Koul and Ling (2006), we have

$$\begin{aligned}&\sup_{x\in\mathbb{R}}\Big|\sqrt{n}\Big\{\tilde{F}_{pn}(x) - F_{pn}(x) - n^{-1}\Sigma_{j=1}^n\Big[F_p\big(x + xv_{npj}(\tilde{t}) + u_{npj}(\tilde{t})\big) - F_p(x)\Big]\Big\}\Big|\\ &= \sup_{x\in\mathbb{R}}\Big|\tilde{\mathcal{U}}(x,\tilde{t}) - \mathcal{U}^*(x,\tilde{t})| = o_p(1),\end{aligned} \tag{14}$$

where $\tilde{\mathcal{U}}(x,t) = n^{-1/2}\Sigma_{i=1}^n I\big(\eta_{ni} \le x + xv_{ni}(t) + u_{ni}(t)\big) - n^{-1/2}\Sigma_{i=1}^n H\big(\eta_{ni} \le x + xv_{ni}(t) + u_{ni}(t)\big)$, and $\mathcal{U}^*(x,t) = n^{-1/2}\Sigma_{i=1}^n\Big[I(\eta_{ni} \le x) - H(x)\Big]$. □

Let

$$\delta_{pi} = \tilde{F}_{pn}(\tilde{\eta}_{pi}) - F_{pn}(\eta_{pi}), \qquad \text{and} \qquad \delta_{pi}^* = \tilde{F}_{pn}(\tilde{\eta}_{pi}) - F_p(\eta_{pi}). \tag{15}$$

Lemma 4. $\sup_{1\le i\le n}|\delta_{pi}| = o_p(1)$ *and* $\sup_{1\le i\le n}|\delta^*_{pi}| = o_p(1)$, *for every* p.

Proof. By adding and subtracting several terms, we have that

$$\begin{aligned}\delta_{pi} = \tilde F_{pn}(\tilde\eta_{pi}) - F_{pn}(\tilde\eta_{pi}) - n^{-1}\Sigma_{j=1}^n\Big[F_p\big(\tilde\eta_{pi}+\tilde\eta_{pi}v_{npj}(\tilde t)+u_{npj}(\tilde t)\big) - F_p(\tilde\eta_{pi})\Big]\\ + \Big\{F_{pn}(\tilde\eta_{pi}) - F_{pn}(\eta_{pi})\Big\} - \big[F_p(\tilde\eta_{pi}) - F_p(\eta_{pi})\big]\\ + n^{-1}\Sigma_{j=1}^n\Big[F_p\big(\tilde\eta_{pi}+\tilde\eta_{pi}v_{npj}(\tilde t)+u_{npj}(\tilde t)\big) - F_p(\tilde\eta_{pi})\Big] + \big[F_p(\tilde\eta_{pi}) - F_p(\eta_{pi})\big].\end{aligned} \tag{16}$$

Expanding $F_p\big(\tilde\eta_{pi}+\tilde\eta_{pi}v_{npj}(\tilde t)+u_{npj}(\tilde t)\big)$ about $F_p(\eta_{pi})$ and then adding and subtracting terms, we have the following for some $\bar\eta_{pij}$ between $\tilde\eta_{pi}+\tilde\eta_{pi}v_{npj}(\tilde t)+u_{npj}(\tilde t)$ and η_{pi}:

$$\begin{aligned}&\sup_{1\le i\le n}\Big|n^{-1}\Sigma_{j=1}^n\Big[F_p\big(\tilde\eta_{pi}+\tilde\eta_{pi}v_{npj}(\tilde t)+u_{npj}(\tilde t)\big) - F_p(\eta_{pi})\Big]\Big|\\ &\quad\le \sup_{1\le i,j\le n}\Big|\big(\tilde\eta_{pi}-\bar\eta_{pij}\big)\ f_p(\bar\eta_{pij})\Big| + \sup_{1\le i,j\le n}\Big|\big(\bar\eta_{pij}-\eta_{pi}\big)\ f_p(\bar\eta_{pij})\Big|\\ &\quad+ \sup_{1\le j\le n}|v_{npj}(\tilde t)|\Big(\sup_{1\le i,j\le n}\Big|\big(\tilde\eta_{pi}-\bar\eta_{pij}\big)\ f_p(\bar\eta_{pij})\Big| + \sup_{1\le i,j\le n}\Big|\bar\eta_{pij}\ f_p(\bar\eta_{pij})\Big|\Big)\\ &\quad+ \sup_{1\le j\le n}|u_{npj}(\tilde t)|\ \|f_p\| \qquad = o_p(1).\end{aligned} \tag{17}$$

Now, by Lemmas 2 and 3, we have $\sup_{1\le i\le n}|\delta_{pi}| = o_p(1)$. The second part follows from $\sup_{1\le i\le n}|\delta^*_{pi}| \le \sup_{1\le i\le n}|\delta_{pi}| + \sup_{1\le i\le n}|F_{pn}(\eta_{pi}) - F_p(\eta_{pi})| = o_p(1)$. □

The proofs of the following Lemmas 5-7 are the same as for the case of linear regression with iid errors. They are given in Kim *et al.* (2005), and hence are not given here.

Lemma 5. *Let* $\Psi(\theta,u_1,u_2)$ *and* $G(u_1,u_2)$ *be the functions defined in (A.3). Also, let* $\{d^n_{pi}\}$ *be a sequence of random variables such that* $\sup_{1\le i\le n}|d^n_{pi}| = o_p(1)$. *Then, for any given* θ,

$$n^{-1}\Sigma_{i=1}^n\Big|\Psi\Big(\theta, F_1(\eta_{1i})+d^n_{1i}, F_2(\eta_{2i})+d^n_{2i}\Big) - \Psi\Big(\theta, F_1(\eta_{1i}), F_2(\eta_{2i})\Big)\Big| = o_p(1), \tag{18}$$

$$n^{-1}\Sigma_{i=1}^n\Big|\Psi\Big(\theta, F_{1n}(\eta_{1i})+d^n_{1i}, F_{2n}(\eta_{2i})+d^n_{2i}\Big) - \Psi\Big(\theta, F_{1n}(\eta_{1i}), F_{2n}(\eta_{2i})\Big)\Big| = o_p(1), \tag{19}$$

$$n^{-1}\Sigma_{i=1}^n\Big|G\Big(F_1(\eta_{1i})+d^n_{1i}, F_2(\eta_{2i})+d^n_{2i}\Big) - G\Big(F_1(\eta_{1i}), F_2(\eta_{2i})\Big)\Big| = o_p(1). \tag{20}$$

Lemma 6. $\sup_{1\le j\le n}\Big|n^{-1}\Sigma_{i=1}^n\big(I\{\tilde\eta_{pi}\le\tilde\eta_{pj}\} - I\{\eta_{pi}\le\eta_{pj}\}\big)\Big| = o_p(1)$ *and* $n^{-1}\Sigma_{j=1}^n n^{-1}\Sigma_{i=1}^n\big(I\{\tilde\eta_{pi}\le\tilde\eta_{pj}\} - I\{\eta_{pi}\le\eta_{pj}\}\big)^2 = o_p(1)$.

Let $\tilde W_p(\tilde\eta_{pi},\theta)$ and $\tilde T_i(\theta_0)$ be as in Theorem 1. Further, let

$$\begin{aligned}\hat W_p(\eta_{pi},\theta) &= n^{-1}\Sigma_{j=1}^n I\big(\eta_{pi}\le\eta_{pj}\big)\ l_{\theta,p}\{\theta, F_1(\eta_{1j}), F_2(\eta_{2j})\}\\ T_i(\theta) &= l_\theta\{\theta, F_1(\eta_{1i}), F_2(\eta_{2i})\} + \hat W_1(\eta_{1i},\theta) + \hat W_2(\eta_{2i},\theta).\end{aligned} \tag{21}$$

Lemma 7. *Let $\tilde{T}_i(\theta)$ and $T_i(\theta)$ be as in Theorem 1 and (21) respectively. Then, there exists an open neighbourhood* $\mathbf{N}$ *of* θ_0 *such that* $\sup_{\theta \in \mathbf{N}(\theta_0)} n^{-1}\Sigma_{i=1}^n G_{in}(\theta) = O_p(1)$, *where* $G_{in}(\theta)$ *is any one of the following four expressions:* $\{\tilde{T}_i(\theta)\}^2$, $\{(\partial/\partial\theta)\tilde{T}_i(\theta)\}^2$, $\{T_i(\theta)\}^2$, $\{(\partial/\partial\theta)T_i(\theta)\}^2$. *Further,* $n^{-1}\Sigma_{i=1}^n \big(\tilde{T}_i(\theta_0) - T_i(\theta_0)\big)^2 = o_p(1)$.

Now, to prove the asymptotic normality of the semiparametric estimator, we first expand the loglikelihood about the true value. Recall that $\tilde{\theta}$ denotes the point at which $L(\theta)$ of (4) has a global maximum. Then, by Taylor expansion around the true copula parameter θ_0, we have

$$0 = \frac{\partial}{\partial\theta}L(\tilde{\theta}) = \frac{\partial}{\partial\theta}L(\theta_0) + (\tilde{\theta}-\theta_0)\frac{\partial^2}{\partial\theta^2}L(\theta_0) + \frac{1}{2}(\tilde{\theta}-\theta_0)^2\frac{\partial^3}{\partial\theta^3}L(\theta^*), \tag{22}$$

where $\theta^* \in [\theta_0, \tilde{\theta}]$. The proof of the consistency of $\tilde{\theta}$ follows by arguments very similar to those used for the MLE as in, for example, Lehmann (1983). Now, solving (22) for $(\tilde{\theta}-\theta_0)$, we obtain

$$\sqrt{n}(\tilde{\theta}-\theta_0) = A_n/\{B_n + C_n\}, \tag{23}$$

$$\begin{aligned} \text{where} \quad A_n &= n^{-1/2}\Sigma_{i=1}^n l_\theta\{\theta_0, \tilde{F}_{1n}(\tilde{\eta}_{1i}), \tilde{F}_{2n}(\tilde{\eta}_{2i})\}, \\ B_n &= -n^{-1}\Sigma_{i=1}^n l_{\theta,\theta}\{\theta_0, \tilde{F}_{1n}(\tilde{\eta}_{1i}), \tilde{F}_{2n}(\tilde{\eta}_{2i}), \\ \text{and} \quad C_n &= -(2n)^{-1/2}\Sigma_{i=1}^n (\tilde{\theta}-\theta_0) l_{\theta,\theta,\theta}\{\theta^*, \tilde{F}_{1n}(\tilde{\eta}_{1i}), \tilde{F}_{2n}(\tilde{\eta}_{2i})\}. \end{aligned} \tag{24}$$

We will show that A_n converges in distribution and $B_n + C_n$ converges in probability. From which, we deduce that the asymptotic distribution of $\sqrt{n}(\tilde{\theta}-\theta_0)$ is normal. Using Taylor expansion around the empirical d.f's $F_{1n}(\eta_{1i})$ and $F_{2n}(\eta_{2i})$, A_n in (23) can be expressed as $A_n = \Sigma_{k=1}^6 A_{nk}$, where, for some $0 \le |c_{1i}|,\ |c_{2i}| \le 1$,

$$\begin{aligned} A_{n1} &= n^{-1/2}\Sigma_{i=1}^n l_\theta\{\theta_0, F_{1n}(\eta_{1i}), F_{2n}(\eta_{2i})\} \\ A_{n2} &= n^{-1/2}\Sigma_{i=1}^n \delta_{1i}\ l_{\theta,1}\{\theta_0, F_{1n}(\eta_{1i}), F_{2n}(\eta_{2i})\} \\ A_{n3} &= n^{-1/2}\Sigma_{i=1}^n \delta_{2i}\ l_{\theta,2}\{\theta_0, F_{1n}(\eta_{1i}), F_{2n}(\eta_{2i})\} \\ A_{n4} &= n^{-1/2}\Sigma_{i=1}^n \delta_{1i}\ \delta_{2i}\ l_{\theta,1,2}\{\theta_0, F_{1n}(\eta_{1i}) + c_{1i}\delta_{1i}, F_{2n}(\eta_{2i}) + c_{2i}\delta_{2i}\} \\ A_{n5} &= n^{-1/2}\Sigma_{i=1}^n \frac{1}{2}(\delta_{1i})^2\ l_{\theta,1,1}\{\theta_0, F_{1n}(\eta_{1i}) + c_{1i}\delta_{1i}, F_{2n}(\eta_{2i}) + c_{2i}\delta_{2i}\} \\ A_{n6} &= n^{-1/2}\Sigma_{i=1}^n \frac{1}{2}(\delta_{2i})^2\ l_{\theta,2,2}\{\theta_0, F_{1n}(\eta_{1i}) + c_{1i}\delta_{1i}, F_{2n}(\eta_{2i}) + c_{2i}\delta_{2i}\}. \end{aligned} \tag{25}$$

We will show that $A_n = A_{n1} + o_p(1)$, from which it follows that the asymptotic distribution of A_n is determined A_{n1}.

Lemma 8. *For $j \in \{2,3\}$, $A_{nj} = o_p(1)$.*

Proof. The expressions A_{n2} and A_{n3} are identical except that they are evaluated for each of the two margins. Therefore, it suffices to to show that $A_{n2} = o_p(1)$. By (16), one has $|A_{n2}| = \left| n^{-1}\Sigma_{i=1}^n n^{1/2}\delta_{1i}\ l_{\theta,1}\{\theta_0, F_{1n}(\eta_{1i}), F_{2n}(\eta_{2i})\} \right| \le \Sigma_{k=1}^3 A_{n2}^{(k)}$, where

$$\begin{aligned} A_{n2}^{(1)} &= \Big| n^{-1/2}\Sigma_{i=1}^n \Big\{ \tilde{F}_{1n}(\tilde{\eta}_{1i}) - F_{1n}(\tilde{\eta}_{1i}) - n^{-1}\Sigma_{j=1}^n \Big[F_1\big(\tilde{\eta}_{1i} + \tilde{\eta}_{1i} v_{n1j}(\tilde{t}) + u_{n1j}(\tilde{t})\big) - F_1(\tilde{\eta}_{1i}) \Big] \Big\} \\ &\qquad \times l_{\theta,1}\{\theta_0, F_{1n}(\eta_{1i}), F_{2n}(\eta_{2i})\} \Big|, \\ A_{n2}^{(2)} &= \Big| n^{-1/2}\Sigma_{i=1}^n \Big\{ \big[F_{1n}(\tilde{\eta}_{1i}) - F_{1n}(\eta_{1i})\big] - \big[F_1(\tilde{\eta}_{1i}) - F_1(\eta_{1i})\big] \Big\} l_{\theta,1}\{\theta_0, F_{1n}(\eta_{1i}), F_{2n}(\eta_{2i})\} \Big|, \\ A_{n2}^{(3)} &= \Big| n^{-1/2}\Sigma_{i=1}^n \Big\{ n^{-1}\Sigma_{j=1}^n \Big[F_1\big(\tilde{\eta}_{1i} + \tilde{\eta}_{1i} v_{n1j}(\tilde{t}) + u_{n1j}(\tilde{t})\big) - F_1(\tilde{\eta}_{1i}) \Big] + \big[F_1(\tilde{\eta}_{1i}) - F_1(\eta_{1i})\big] \Big\} \\ &\qquad \times l_{\theta,1}\{\theta_0, F_{1n}(\eta_{1i}), F_{2n}(\eta_{2i})\} \Big|. \end{aligned}$$

We will show that $A_{n2}^{(k)} = o_p(1)$, for $k = 1, 2, 3$. First, since $n^{-1}\Sigma_{i=1}^n |l_{\theta,1}\{\theta_0, F_1(\eta_{1i}), F_2(\eta_{2i})\}| = O_p(1)$, it follows from Lemma 3 that $A_{n2}^{(1)} = o_p(1)$. Similarly, $A_{np}^{(1)} = o_p(1)$. To show that $A_{n2}^{(3)} = o_p(1)$, note that, by arguments similar to those for (17),

$$A_{n2}^{(3)} = \Big| n^{-3/2}\Sigma_{i,j=1}^{n,n} \Big(\tilde{\eta}_{1i} - \eta_{1i} + \tilde{\eta}_{1i} v_{n1j}(\tilde{t}) + u_{n1j}(\tilde{t}) \Big)\ f_1(\bar{\eta}_{1ij})\ l_{\theta,1}\{\theta_0, F_{1n}(\eta_{1i}), F_{2n}(\eta_{2i})\} \Big|. \quad (26)$$

Let $a_{np,j}, b_{np,j}, \bar{a}_{np}$, and $\bar{b}_{np}$ be defined as in (8), $\bar{a}_{np} = n^{-1}\Sigma_{i=1}^n a_{np,i}$ and $\bar{b}_{np} = n^{-1}\Sigma_{i=1}^n b_{pn,i}$.

By (12), we have that

$$\begin{aligned} \tilde{\eta}_{1i} - \eta_{1i} + \tilde{\eta}_{1i} v_{n1j}(\tilde{t}) + u_{n1j}(\tilde{t}) &= u_{n1j}(\tilde{t}) - u_{n1i}(\tilde{t}) + \big\{ v_{n1j}(\tilde{t}) - v_{n1i}(\tilde{t}) \big\} \tilde{\eta}_{1i} \\ &= (\tilde{\alpha}_1 - \alpha_1^0)'[(a_{n1,j} - a_{n1,i}) + (b_{n1,j} - b_{n1,i})\tilde{\eta}_{1i}]. \end{aligned} \quad (27)$$

Therefore, $A_{n2}^{(3)} \le |B_{n1}| + |B_{n2}|$, where

$$\begin{aligned} B_{n1} &= n^{-3/2}\Sigma_{i=1}^n \Sigma_{j=1}^n (\tilde{\alpha}_1 - \alpha_1^0)'(a_{n1,j} - a_{n1,i}) f_1(\bar{\eta}_{1ij})\ l_{\theta,1}\{\theta_0, F_{1n}(\eta_{1i}), F_{2n}(\eta_{2i})\}, \\ B_{n2} &= n^{-3/2}\Sigma_{i=1}^n \Sigma_{j=1}^n (\tilde{\alpha}_1 - \alpha_1^0)'(b_{n1,j} - b_{n1,i})\tilde{\eta}_{1i} f_1(\bar{\eta}_{1ij}) l_{\theta,1}\{\theta_0, F_{1n}(\eta_{1i}), F_{2n}(\eta_{2i})\}. \end{aligned}$$

We will show that $|B_{n1}| = o_p(1)$ and $|B_{n2}| = o_p(1)$. By (A.3), there exist $|s_{in}| < 1$ such that

$$\begin{aligned} l_{\theta,1}\{\theta_0, F_{1n}(\eta_{1i}), F_{2n}(\eta_{2i})\} &= l_{\theta,1}\{\theta_0, F_1(\eta_{1i}), F_2(\eta_{2i})\} \\ &\quad + s_{in} k(F_1(\eta_{1i}), F_2(\eta_{2i}); \theta_0)\big(|F_{1n}(\eta_{1i}) - F_1(\eta_{1i})| + |F_{2n}(\eta_{2i}) - F_2(\eta_{2i})|\big). \end{aligned} \quad (28)$$

Then, one has $B_{n1} = B_{n1}^{(1)} + B_{n1}^{(2)} + B_{n1}^{(3)} + B_{n1}^{(4)}$, where,

$$|B_{n1}^{(1)}| = \Big| n^{-1}\Sigma_{i=1}^{n}\sqrt{n}(\tilde{\alpha}_1 - \alpha_1^0)'(\bar{a}_{n1} - a_{n1,i})f_1(\eta_{1i})\ l_{\theta,1}\{\theta_0, F_1(\eta_{1i}), F_2(\eta_{2i})\}\Big|,$$

$$|B_{n1}^{(2)}| = \Big| n^{-1}\Sigma_{i=1}^{n}\sqrt{n}(\tilde{\alpha}_1 - \alpha_1^0)'(\bar{a}_{n1} - a_{n1,i})f_1(\eta_{1i}) \times s_{in}k(F_1(\eta_{1i}), F_2(\eta_{2i});\theta_0)\big(|F_{1n}(\eta_{1i}) - F_1(\eta_{1i})| + |F_{2n}(\eta_{2i}) - F_2(\eta_{2i})|\big)\Big|,$$

$$|B_{n1}^{(3)}| = \Big| n^{-2}\Sigma_{i,j=1}^{n,n}\sqrt{n}(\tilde{\alpha}_1 - \alpha_1^0)'(a_{n1,j} - a_{n1,i}) \times \Big(f_1(\bar{\eta}_{1ij}) - f_1(\eta_{1i})\Big)\ l_{\theta,1}\{\theta_0, F_1(\eta_{1i}), F_2(\eta_{2i})\}\Big|,$$

$$|B_{n1}^{(4)}| = \Big| n^{-2}\Sigma_{i,j=1}^{n,n}\sqrt{n}(\tilde{\alpha}_1 - \alpha_1^0)'(a_{n1,j} - a_{n1,i})\Big(f_1(\bar{\eta}_{1ij}) - f_1(\eta_{1i})\Big) \times s_{in}k(F_1(\eta_{1i}), F_2(\eta_{2i});\theta_0)\big(|F_{1n}(\eta_{1i}) - F_1(\eta_{1i})| + |F_{1n}(\eta_{2i}) - F_1(\eta_{2i})|\big)\Big|.$$

Since $(\bar{a}_{n1} - a_{n1,i})f_1(\eta_{1i})l_{\theta,1}\{\theta_0, F_1(\eta_{1i}), F_2(\eta_{2i})\}$ is a strictly stationary and ergodic process with mean zero, for $n = 1, 2, \ldots$, it follows that $B_{n1}^{(1)} = o_p(1)$. By the Cauchy-Schwarz inequality, we have

$$|B_{n1}^{(2)}| \leq ||n^{1/2}(\tilde{\alpha}_1 - \alpha_1^0)||\Big(n^{-1}\Sigma_{i=1}^{n}(\bar{a}_{n1} - a_{n1,i})'(\bar{a}_{n1} - a_{n1,i})\Big)^{1/2} \times \Big(n^{-1}\Sigma_{i=1}^{n}\big|k(F_1(\eta_{1i}), F_2(\eta_{2i});\theta_0)\big|^2\Big)^{1/2} \times \sup_{1\leq i\leq n}\big(|F_{1n}(\eta_{1i}) - F_1(\eta_{1i})| + |F_{1n}(\eta_{2i}) - F_1(\eta_{2i})|\big) = o_p(1), \tag{29}$$

by (A.8) and (A.3). Similarly, one has $B_{n1}^{(4)} = o_p(1)$ by (A.8) and (A.3).

Because $\bar{\eta}_{1ij}$ lies between $\tilde{\eta}_{1i} + \tilde{\eta}_{1i}v_{n1j}(\tilde{t}) + u_{n1j}(\tilde{t})$ and η_{1i}, by (27), we have that

$$|f_1(\bar{\eta}_{1ij}) - f_1(\eta_{1i})| \leq |\tilde{\eta}_{1i} - \eta_{1i} + \tilde{\eta}_{1i}v_{n1j}(\tilde{t}) + u_{n1j}(\tilde{t})|||f_1'|| \leq \|\tilde{\alpha}_1 - \alpha_1^0\|[\|a_{n1,j}\| + \|a_{n1,i}\| + (\|b_{n1,j}\| + \|b_{n1,i}\|)|\tilde{\eta}_{1i}| \tag{30}$$

It follows from (10) and part (ii) of Lemma 1, that $\sup_{1\leq i\leq n}|\tilde{\eta}_{1i}| \leq (1+\zeta_n)\sup_{1\leq i\leq n}|\eta_{1i}|]$ where $\zeta_n = o_p(1)$. Substituting these in the expression for $B_{n1}^{(3)}$ and using the fact that the process (Y_{1t}, Y_{2t}) is stationary and ergodic, we have that $B_{n1}^{(3)} = o_p(1)$. Since $B_{n1}^{(1)}, B_{n1}^{(2)}$, and $B_{n1}^{(4)}$ are also $o_p(1)$, we conclude that $B_{n1} = o_p(1)$. By similar arguments, we also have $B_{n2} = o_p(1)$. Therefore, we have $A_{n2} = o_p(1)$. By similar arguments, we can show that $A_{n3} = o_p(1)$. This completes the proof of the lemma. □

Lemma 9. *For $j \in \{4, 5, 6\}$, $|A_{nj}| = o_p(1)$.*

Proof. Let $d_{npi} = F_{pn}(\eta_{pi}) - F_p(\eta_{pi}) + c_{pi}\delta_{pi}$, where $c_{pi}\delta_{pi}$ is defined in (25). Then, $F_{pn}(\eta_{pi}) + c_{pi}\delta_{pi} = F_p(\eta_{pi}) + d_{npi}$ and, by (4), $\sup_{1\leq i\leq n}|d_{n1i}| \leq \sup_{1\leq i\leq n}\Big|F_{1n}(\eta_{1i}) - F_1(\eta_{1i})\Big| + \sup_{1\leq i\leq n}|c_1\delta_{1i}| = o_p(1)$. Because $|A_{n4}| \leq \sup_{1\leq i\leq n}|\delta_{2i}|\ n^{-1}\Sigma_{i=1}^{n}\Big|\sqrt{n}\delta_{1i}\ l_{\theta,1,2}\{\theta_0, F_1(\eta_{1i}) + d_{n1i}, F_2(\eta_{2i}) + d_{n2i}\}\Big|$, and $\sup_{1\le i\le n}|\delta_{2i}| = o_p(1)$, we only need to show that $n^{-1}\Sigma_{i=1}^n\Big|\sqrt{n}\delta_{1i}\ l_{\theta,1,2}\{\theta_0, F_1(\eta_{1i}) + d_{n1i}, F_2(\eta_{2i}) + d_{n2i}\}\Big| = O_p(1)$. By (16), one has

$$n^{-1}\Sigma_{i=1}^n\Big|\sqrt{n}\delta_{1i}\ l_{\theta,1,2}\{\theta_0, F_1(\eta_{1i}) + d_{n1i}, F_2(\eta_{2i}) + d_{n2i}\}\Big| \le \Sigma_{k=1}^3 A_{n4}^{(k)}, \tag{31}$$

where

$$\begin{aligned}
A_{n4}^{(1)} &= n^{-1}\Sigma_{i=1}^n\Big|n^{1/2}\Big\{\tilde{F}_{1n}(\tilde{\eta}_{1i}) - F_{1n}(\tilde{\eta}_{1i}) - n^{-1}\Sigma_{j=1}^n\Big[F_1\big(\tilde{\eta}_{1i} + \tilde{\eta}_{1i}v_{npj}(\tilde{t}) + u_{npj}(\tilde{t})\big) - F_1(\tilde{\eta}_{1i})\Big]\Big\} \\
&\quad \times l_{\theta,1,2}\{\theta_0, F_1(\eta_{1i}) + d_{n1i}, F_2(\eta_{2i}) + d_{n2i}\}\Big|, \\
A_{n4}^{(2)} &= n^{-1}\Sigma_{i=1}^n\Big|n^{1/2}\Big\{\big[F_{1n}(\tilde{\eta}_{1i}) - F_{1n}(\eta_{1i})\big] - \big[F_1(\tilde{\eta}_{1i}) - F_1(\eta_{1i})\big]\Big\} \\
&\quad \times l_{\theta,1,2}\{\theta_0, F_1(\eta_{1i}) + d_{n1i}, F_2(\eta_{2i}) + d_{n2i}\}\Big|, \\
A_{n4}^{(3)} &= n^{-1}\Sigma_{i=1}^n\Big|n^{1/2}\Big\{n^{-1}\Sigma_{j=1}^n\Big[F_1\big(\tilde{\eta}_{1i} + \tilde{\eta}_{1i}v_{npj}(\tilde{t}) + u_{npj}(\tilde{t})\big) - F_1(\tilde{\eta}_{1i})\Big] + \big[F_1(\tilde{\eta}_{1i}) - F_1(\eta_{1i})\big]\Big\} \\
&\quad \times l_{\theta,1,2}\{\theta_0, F_1(\eta_{1i}) + d_{n1i}, F_2(\eta_{2i}) + d_{n2i}\}\Big|.
\end{aligned}$$

Now, $A_{n4}^{(1)} = o_p(1)$ and $A_{n4}^{(2)} = o_p(1)4$ by Lemmas 3 and 5. To show that $A_{n4}^{(3)} = O_p(1)$, note that by arguments similar to those in the proof of Lemma 8, we have that $A_{n4}^{(3)} \le E_{n1} + E_{n2} + E_{n3}$, where

$$\begin{aligned}
E_{n1} &= n^{-3/2}\Sigma_{i,j=1}^{n,n}\Big|(\tilde{\alpha}_1 - \alpha_1^0)'(a_{n1,j} - a_{n1,i})f_1(\bar{\eta}_{1ij})\ l_{\theta,1,2}\{\theta_0, F_1(\eta_{1i}) + d_{n1i}, F_2(\eta_{2i}) + d_{n2i}\}\Big|, \\
E_{n2} &= n^{-3/2}\Sigma_{i,j=1}^{n,n}\Big|(\tilde{\alpha}_1 - \alpha_1^0)'(b_{n1,j} - b_{n1,i})\bar{\eta}_{1ij}f_1(\bar{\eta}_{1ij})\ l_{\theta,1,2}\{\theta_0, F_1(\eta_{1i}) + d_{n1i}, F_2(\eta_{2i}) + d_{n2i}\}\Big|, \\
E_{n3} &= n^{-3/2}\Sigma_{i=1}^n\Sigma_{j=1}^n\Big|(\tilde{\alpha}_1 - \alpha_1^0)'(b_{n1,j} - b_{n1,i})\big(\tilde{\eta}_{1i} - \bar{\eta}_{1ij}\big) \\
&\quad \times f_1(\bar{\eta}_{1ij})\ l_{\theta,1,2}\{\theta_0, F_1(\eta_{1i}) + d_{n1i}, F_2(\eta_{2i}) + d_{n2i}\}\Big|.
\end{aligned}$$

Then, one has

$$\begin{aligned}
E_{n1} &\le n^{1/2}\|\tilde{\alpha}_1 - \alpha_1^0\|\ ||f_1||\Big[\Big(n^{-1}\Sigma_{j=1}^n\|a_{n1,j}\|\Big)\ \Big(n^{-1}\Sigma_{i=1}^n\Big|l_{\theta,1}\{\theta_0, F_1(\eta_{1i}), F_2(\eta_{2i})\}\Big| + o_p(1)\Big) \\
&\quad + \Big(n^{-1}\Sigma_{i=1}^n\|a_{n1,i}\|^2\Big)^{1/2}\ \Big(n^{-1}\Sigma_{i=1}^n\Big|l_{\theta,1}\{\theta_0, F_1(\eta_{1i}), F_2(\eta_{2i})\}\Big|^2 + o_p(1)\Big)^{1/2}\Big] = O_p(1),
\end{aligned}$$

by (A.8) and Lemma 5. By replacing $||f_1||$ with $\sup_{1\le i,j\le n}|\bar{\eta}_{1ij}f_1(\bar{\eta}_{1ij})|$, one can apply similar arguments to prove that $E_{n2} = O_p(1)$. By (13), we have that $\sup_{1\le i,j\le n}|\tilde{\eta}_{1i} - \bar{\eta}_{1ij}| \le (\zeta_n + \zeta_n|\eta_i|)$, where $\zeta_n = o_p(1)$. Now, by arguments similar to those for the proof of $B_{n3} = o_p(1)$ in the proof of Lemma 8, we have $E_{n3} = o_p(1)$. Therefore, by (31), we have that $A_{n4} = O_p(1)$. Similar arguments can be applied for A_{n5} and A_{n6}. This completes the proof. $\square$

Proof of Theorem 1 (continued): It follows from the foregoing lemmas that $A_n = A_{n1} + o_p(1)$.

The asymptotic distribution of A_{n1} was established by Genest *et al.* (1995b). They showed that $A_{n1} \sim N(0, \sigma^2)$, where, $\sigma^2 = \text{var}[l_\theta\{\theta_0, F_1(\eta_1), F_1(\eta_1)\} + W_1(\eta_1) + W_2(\eta_2)]$. To complete the proof of the asymptotic normality of $n^{1/2}(\tilde{\theta} - \theta_0)$, we need to show that $C_n = o_p(1)$ and $|B_n - s| = o_p(1)$. It follows from Lemma 5 and the consistency of $\tilde{\theta}$ that $C_n = o_p(1)$. Now, to show that $|B_n - s| = o_p(1)$, note that,

$$|B_n - s| \leq \big| - n^{-1}\Sigma_{i=1}^n l_{\theta,\theta}\{\theta_0, \tilde{F}_{1n}(\tilde{\eta}_{1i}), \tilde{F}_{2n}(\tilde{\eta}_{2i})\} + n^{-1}\Sigma_{i=1}^n l_{\theta,\theta}\{\theta_0, F_1(\eta_{1i}), F_2(\eta_{2i})\}\big| \\ + \big| - n^{-1}\Sigma_{i=1}^n l_{\theta,\theta}\{\theta_0, F_1(\eta_{1i}), F_2(\eta_{2i})\} - s\big|.$$

The first term on the R.H.S. converges to zero in probability by Lemma 5 and the second term also converges to zero in probability by the Weak Law of Large Numbers because $s = -E\big[l_{\theta,\theta}\{\theta_0, F_1(\eta_1), F_2(\eta_2)\}\big]$.

This completes the proof of the asymptotic normality of the estimator. The proof of the fact that $\tilde{v}^2$ is consistent estimator of the asymptotic variance is practically the same as for the multiple regression case in Kim *et al.* (2005) and hence is omitted. This completes the proof of Theorem 1.

The proof of Theorem 2 requires only straightforward modifications to the proof of Theorem 1 to change the notation for θ being a scalar to a vector.

References

Bagdonavicius, V., Malov, S. V., and Nikulin, M. S. (2006). Semiparametric regression estimation in copula models. *Communications in Statistics-Theory and Methods*, **35**, 1449–1467.

Bandeen-Roche, K. and Liang, K.-Y. (2002). Modelling multivariate failure time associations in the presence of a competing risk. *Biometrika*, **89**(2), 299–314.

Chen, X. and Fan, Y. (2006). Estimation and model selection of semiparametric copula-based multivariate dynamic models under copula misspecification. *Journal of Econometrics*, **135**, 125–154.

Cherubini, U., Luciano, E., and Vecchiato, W. (2004). *Copula Methods in Finance*. John Wiley and Sons Ltd, Chichester, U.K.

Genest, C. and Favre, C. (2007). Everything you always wanted know about copula but were afraid to ask. *Journal of hydrologic engineering, (in press)*.

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

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

Hutchinson, T. P. and Lai, C. D. (1990). *Continuous Bivariate Distributions, Emphasizing Applications*. Adelaide: Rumsby Scientific.

Joe, H. (1997). *Multivariate Models and Dependence Concepts*. Chapman and Hall, London.

Junker, M. and May, A. (2005). Measurement of aggregate risk with copulas. *The Econometrics Journal*, **8**(3), 428–454.

Kim, G., Silvapulle, M. J., and Silvapulle, P. (2005). Semiparametric estimation of the joint distribution of two variables when each variable satisfies a regression model. *Joint Statistical Meetings, August 2005, Minneapolis*. [Available as working paper number 1/07, Department of Econometrics and Business Statistics, Monash University].

Kim, G., Silvapulle, M. J., and Silvapulle, P. (2007). Comparison of semiparametric and parametric methods for estimating copulas. *Computational Statistics and Data Analysis*, **51**, 2836–2850.

Koul, H. L. (2002). *Weighted Empirical Processes in Dynamic Nonlinear Models*. Lecture Notes in Statistics, Vol 166. Springer Verlag, New York.

Koul, H. L. and Ling, S. (2006). Fitting an error distribution in some heteroscedastic time series model. *Annals of Statistics*, **34**, 994–1012.

Lehmann, E. L. (1983). *Theory of Point Estimation.* Wiley: New York.

Nelsen, R. B. (2006). *An introduction to copulas.* Number 139 in Lecture Notes in Statistics. Springer Verlag, New York.

Oakes, D. (1994). Multivariate survival distributions. *Journal of Nonparametric Statistics*, **3**, 343–354.

Oakes, D. and Wang, A. (2003). Copula model generated by Dabrowska's association measure. *Biometrika*, **90**(2), 478–481.

Patton, A. J. (2006). Modelling asymmetric exchange rate dependence. *International Economic Review*, **47**, 527–556.

Shih, J. and Louis, T. (1995). Inferences on the association parameter in copula models for bivariate survival data. *Biometrics*, **51**, 1384–1399.

Sklar, A. (1959). Fonctions de répartition à n dimensionset leurs marges. *Publ. Inst, Statis. Univ. Paris*, **8**, 229–231.

Taniguchi, M. and Kakizawa, Y. (2000). *Asymptotic Theory of Statistical Inference for Time Series.* Springer Verlag, New York.

Wang, W. (2003). Estimating the association parameter for copula models under dependent censoring. *Journal of the Royal Statistical Society, Series B: Statistical Methodology*, **65**(1), 257–273.

Wang, W. and Ding, A. A. (2000). On assessing the association for bivariate current status data. *Biometrika*, **87**, 879–893.

Table 1: Efficiencies (%) of the Semiparametric estimator relative to the inference function for margin estimator in terms of mean square error.

θ	100 observations (N-N)	(T-T)	(T-ST)	(T-C)	500 observations (N-N)	(T-T)	(T-ST)	(T-C)
				Ali-Mikhail-Haq faimly of Copulas				
-0.50	96	130	108	107	100	288	137	129
0.40	97	163	140	147	100	341	231	255
0.71	98	73	105	116	96	228	324	420
				Frank Copula				
1.86	94	312	157	154	97	1340	362	332
5.74	98	313	133	124	97	1460	309	245
18.2	91	123	268	330	94	139	948	1290
				Gumbel Copula				
1.25	71	280	230	260	89	2610	960	510
2.00	79	131	117	143	88	330	300	410
5.00	95	188	510	630	93	160	1910	2200
				Joe Copula				
1.44	58	105	107	109	80	240	410	480
2.86	73	168	210	279	78	270	740	1050
8.77	112	230	550	650	101	220	1570	1940
				Plackett Copula				
2.50	90	420	169	166	97	2090	400	360
11.6	102	280	107	104	99	1260	139	105
115	86	117	270	310	89	180	930	1020
				Gaussian Copula				
0.31	88	102	100	105	93	174	163	156
0.71	113	204	260	320	101	630	1060	1680
0.95	89	480	1800	2620	89	1300	10200	8450
				T Copula				
0.31	96	172	99	107	98	300	111	113
0.71	117	150	150	179	105	590	360	410
0.95	102	117	1010	1020	95	400	1790	1860

Note: The error distributions are (1) N-N: normal and normal, (2) T-T: t_3 and t_3, (3) T-ST: t_3 and skew-t_3, and (4) T-C: t_3 and $\chi^2(5)$. The number of repeated samples is 250.

Table 2: Estimated means and standard deviations when the marginal distributions are t_3 and $\chi^2(5)$ but the inference function for margin method assumes that they are normal.

	100 observations				500 observations			
	IFM		Semi		IFM		Semi	
θ	mean	std	mean	std	mean	std	mean	std
				Ali-Mikhail-Haq faimly of Copulas				
-0.50	-0.53	0.36	-0.52	0.35	-0.55	0.18	-0.50	0.16
0.40	0.42	0.31	0.37	0.25	0.49	0.14	0.40	0.11
0.71	0.76	0.18	0.69	0.17	0.84	0.06	0.71	0.07
				Frank Copula				
1.86	2.13	0.71	1.88	0.61	2.26	0.39	1.89	0.31
5.74	6.02	0.85	5.75	0.81	6.19	0.44	5.75	0.40
18.2	14.2	1.61	16.8	1.90	14.4	1.01	17.6	0.88
				Gumbel Copula				
1.25	1.18	0.16	1.28	0.10	1.18	0.09	1.26	0.05
2.00	1.83	0.17	2.04	0.20	1.83	0.09	2.00	0.10
5.00	3.41	0.48	4.61	0.54	3.34	0.29	4.74	0.25
				Joe Copula				
1.44	1.29	0.13	1.50	0.18	1.27	0.07	1.46	0.09
2.86	2.32	0.31	2.88	0.37	2.30	0.18	2.84	0.18
8.77	4.64	1.09	7.54	1.15	4.37	0.73	7.92	0.55
				Plackett Copula				
2.50	2.87	0.93	2.59	0.77	2.98	0.49	2.55	0.36
11.6	11.4	3.25	11.5	3.19	11.9	1.50	11.5	1.48
115	51.0	13.6	88.2	23.5	50.0	8.92	98.6	12.1
				Gaussian Copula				
0.31	0.28	0.09	0.32	0.09	0.29	0.04	0.31	0.04
0.71	0.64	0.07	0.71	0.05	0.63	0.05	0.70	0.02
0.95	0.87	0.06	0.94	0.01	0.87	0.04	0.94	0.01
				T Copula				
0.31	0.29	0.11	0.31	0.10	0.30	0.05	0.31	0.05
0.71	0.66	0.06	0.70	0.06	0.67	0.04	0.70	0.03
0.95	0.89	0.03	0.94	0.02	0.91	0.01	0.94	0.01

Table 3: Coverage rates (in %) of an asymptotic 95% confidence interval of the copula parameter based on the semiparametric estimator

	Sample sizes							
	100 samples				500 samples			
true value	N-N	T-C	T-ST	T-T	N-N	T-C	T-ST	T-T
	AMH copula							
-0.50	95	95	95	95	96	96	96	96
0.40	94	92	92	93	93	94	93	93
0.71	91	91	91	92	92	95	95	93
	Frank copula							
1.86	98	96	96	97	94	93	94	94
5.74	97	96	97	97	94	93	93	93
18.2	94	89	90	94	94	86	87	87
	Gumbel copula							
1.25	96	96	96	97	93	93	94	94
2.00	96	97	96	95	93	92	93	92
5.00	90	80	85	90	88	79	84	86
	Joe copula							
1.44	96	95	96	96	94	94	94	94
2.86	95	94	94	94	92	93	93	91
8.77	79	78	80	73	61	62	60	49
	Plackett copula							
2.50	95	95	95	96	97	97	98	97
11.6	100	100	100	100	100	100	100	100
115	99	100	99	99	96	98	99	99
	Gaussian Copula							
0.31	100	100	100	100	100	100	100	100
0.71	100	100	100	100	100	100	100	100
0.95	100	100	100	100	100	100	100	100
	T Copula							
0.31	96	97	94	96	95	96	97	97
0.71	100	100	100	100	99	100	100	99
0.95	100	100	100	100	100	100	100	100

Note: The error distributions are (1) N-N: normal and normal, (2) T-T: t_3 and t_3, (3) T-ST: t_3 and skew-t_3, and (4) T-C: t_3 and $\chi^2(5)$. The number of repeated samples is 250.

Table 4: Estimates of the copula parameter and chi-square goodness of fit tests.

Copula	$\tilde{\theta}$	s.e.($\tilde{\theta}$)	p value(Chi-sq test)
AMH	0.94	(0.01)	0.00
Clayton	0.79	(0.05)	0.00
Frank	3.77	(0.15)	0.01
Gaussian	0.53	(0.02)	0.33
Gumbel	1.51	(0.03)	0.00
Joe	1.64	(0.05)	0.00
Plackett	5.45	(0.62)	0.03
Sym-JC	0.32	(0.03)	0.02
	0.34	(0.03)	