ISSN 1440-771X

MONASH University

Australia

Department of Econometrics
and Business Statistics

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

Semiparametric estimation of the dependence parameter of the error terms in multivariate regression

Gunky Kim, Mervyn J. Silvapulle and Paramsothy Silvapulle

February 2007

Working Paper 1/07

Semiparametric estimation of the dependence parameter of the error terms in multivariate regression

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

First version: Presented at the Joint Statistical Meetings, Minneapolis, August 2005.
Second version: August 2006.

Summary

A semiparametric method is developed for estimating the dependence parameter and the joint distribution of the error term in the multivariate linear regression model. The nonparametric part of the method treats the marginal distributions of the error term as unknown, and estimates them by suitable empirical distribution functions. Then a pseudolikelihood is maximized to estimate the dependence parameter. It is shown that this estimator is asymptotically normal, and a consistent estimator of its large sample variance is given. A simulation study shows that the proposed semiparametric estimator is better than the parametric methods available when the error distribution is unknown, which is almost always the case in practice. It turns out that there is no loss of asymptotic efficiency due to the estimation of the regression parameters. An empirical example on portfolio management is used to illustrate the method. This is an extension of earlier work by Oakes (1994) and Genest et al. (1995) for the case when the observations are independent and identically distributed, and Oakes and Ritz (2000) for the multivariate regression model.

Some key words: Copula; Pseudolikelihood; Robustness.

JEL Codes: C01, C12, C13, C14.

1. Introduction

Estimation of the joint distribution of a random vector and learning about inter-dependence among its components are important topics in statistical inference. This paper develops a method for estimating the joint distribution and the dependence parameter of the error distribution in multivariate linear regression.

It is now well-known that the joint cumulative distribution function $H(x_1, \ldots, x_k)$ of a random vector $(X_1, \ldots, X_k)$ with continuous marginals $F_i(x_i) = \text{pr}(X_i \leq x_i)$ 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 $U_i = F_i(X_i)$ is distributed uniformly on $[0, 1]$, $i = 1, \ldots, k$ (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 between the components of the random vector from the marginal distributions.

It is possible that 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 marginal parameters from the intrinsic association which is 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 intrinsic association among the marginals (Wang and Ding (2000), Oakes and Wang (2003)).

Copulas have been used in a very wide range of applied 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) provides a comprehensive and authoritative account 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. In what follows, we shall restrict our discussion to bivariate observations only, for simplicity. However, the extensions to higher dimensions would be straight forward.

The use of Copulas in risk management has been increasing substantially in the recent past where the main interest is on the whole distribution rather than just the association parameter (Cherubini et al. (2004)). As an example, let Y_1 and Y_2 denote the market values of two shares, say a bank and a mining company respectively. Let x denote a market index such as the Dow Jones Index. Suppose that $Y_1 = x_1^T\beta_1 + \epsilon_1$ and $Y_2 = x_2^T\beta_2 + \epsilon_2$, where $x_1 = x_2 = (1, x)^T$. Thus, after accounting for the overall market movements, ϵ_1 and ϵ_2 represent the risks that are not under the control of the investor. For managing the risks associated with a portfolio consisting of these two investments, for example, for assessing the need for diversification of investments, the main quantities of interest are functions of the joint distribution of (ϵ_1, ϵ_2). Examples of quantities that are of interest include, (i) $\text{pr}(Y_1 \leq a_1 \text{ and } Y_2 \leq a_2)$ and $\text{pr}(Y_1 \leq a_1 \mid Y_2 \leq a_2)$, where a_1 and a_2 are given numbers, and (ii) the *Value at Risk*, c, defined by $\text{pr}\{b_1Y_1 + (1 - b_1)Y_2 \leq c\} \leq \alpha$, where α is a given small number, for example $\alpha = 0.05$, and b_1 is the proportion of investment in the first asset.

This paper develops a new semiparametric method for estimating the dependence parameter θ and joint distribution of the error term, (ϵ_1, ϵ_2). If the assumptions for the traditional normal-theory linear model are satisfied, then it would be possible to obtain an optimal estimate of the joint distribution of the error terms by maximum likelihood. However, in many areas of applications, for example risk management, the marginal distributions of the error terms are far from being normal. In fact, returns from investments are notorious for being long-tailed and skewed. Further, the marginal distributions of (ϵ_1, ϵ_2) may also take different functional forms. For example, ϵ_1 and ϵ_2 may be distributed as gamma and normal respectively. Consequently, the well known elliptically symmetric families of distributions, such as the multivariate normal and t, are inadequate. Further, often it is of interest to apply inference methods that make as few assumptions as possible about the functional form of the distribution. In this paper, we propose a semi-parametric method that fits this requirement. An attractive feature of this method is that it does not cause any additional difficulties due to the marginal distributions being long tailed and skewed, features that are common in financial data and have attracted considerable interest in financial statistics.

The method introduced in this paper started with Oakes (1994) and Genest et al. (1995). They proposed a procedure for estimating the dependence parameter in a copula for independent and identically distributed observations when the marginal marginal distributions are treated as unknown. The method involves two stages of estimation: In the first stage, the marginal distributions are estimated by their respective empirical distribution functions, and in the second stage, the maximum likelihood is applied with the marginal distributions replaced by the corresponding empirical distributions. Genest et al. (1995) showed that this estimator is asymptotically normal for different settings; see also Wang and Ding (2000) and Shih and Louis (1995). In this paper, we extend this approach to the multivariate linear regression model when the interest is centered on the joint distribution of the multivariate error term. Oakes and Ritz (2000) also studied estimation of the copula of the error term in the same multivariate linear regression setting but they used a fully parametric approach.

Now, to introduce the method developed here, let us consider the bivariate investment example considered earlier in this section. Let $F_1(t_1) = \mathrm{pr}(\epsilon_1 \le t_1)$, $F_2(t_2) = \mathrm{pr}(\epsilon_2 \le t_2)$ and let $C(u_1, u_2; \theta)$ denote the copula of (ϵ_1, ϵ_2), θ being an unknown parameter which we shall assume to be a scalar for simplicity. Therefore, the joint distribution of (ϵ_1, ϵ_2) is given by $\mathrm{pr}(\epsilon_1 \leq t_1, \epsilon_2 \leq t_2) = C\{F_1(t_1), F_2(t_2); \theta\}$. Throughout, we shall assume that the functional form of $C(u_1, u_2; \theta)$ is known, but F_1 and F_2 are unknown. We propose an estimator of θ and show that it is consistent and asymptotically normal. Further, we also obtain a consistent estimator of its asymptotic variance so that confidence intervals may be constructed. Simulation results show that our proposed method performs better than the traditional fully parametric methods of inference when the functional forms of the marginal distributions are unknown, which is of course almost always the case.

The rest of the paper is structured 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. The main results

As indicated in the introduction, we shall consider the bivariate case for notational simplicity. The extension to the multivariate case is almost straight forward. Let the data generating process for (Y_1, Y_2) be $Y_1 = x_1^T\beta_1 + \epsilon_1$, and $Y_2 = x_2^T\beta_2 + \epsilon_2$ where x_1 and x_2 are vectors of covariates associated with Y_1 and Y_2 respectively. In what follows, these covariates are assumed to be non-stochastic. However, the results would hold, with appropriate modifications, even if they are stochastic. Suppose that there are n independent observations $(Y_{1i}, x_{1i}, Y_{2i}, x_{2i})$, $(i = 1, \ldots, n)$. Thus, we have $Y_{pi} = x_{pi}^T\beta_p + \epsilon_{pi}$, $(i = 1, \ldots, n$, $p = 1, 2)$. Let f_p and F_p denote the probability density and cumulative distribution functions of ϵ_p respectively, $p = 1, 2$. Let $C(u_1, u_2; \theta)$ and $c(u_1, u_2; \theta)$ denote the copula of (ϵ_1, ϵ_2) and the corresponding density function, respectively. Then, the loglikelihood takes the form, $\ell^*(\theta, \beta_1, \beta_2) = L^*(\theta; \beta_1, \beta_2, F_1, F_2) + B(\beta_1, \beta_2, f_1, f_2)$ where $L^*(\theta; \beta_1, \beta_2, F_1, F_2) =$

$\sum \log c\{F_1(\epsilon_{1i}), F_2(\epsilon_{2i}); \theta\}$ and $B(\beta_1, \beta_2, f_1, f_2) = \sum \log\{f_1(\epsilon_{1i}), f_2(\epsilon_{2i})\}$. The maximum likelihood estimator of $(\theta, \beta_1, \beta_2)$ is simply the point at which $\ell^*(\beta_1, \beta_2, \theta)$ reaches its maximum. If the joint distribution of (ϵ_1, ϵ_2) is correctly specified then this estimator is consistent and asymptotically normal, provided some regularity conditions are satisfied. Since the term $B(\beta_1, \beta_2, f_1, f_2)$ does not depend on θ, it may be ignored for the purposes of estimating θ by maximum likelihood.

Now, we introduce the following semiparametric estimator of the copula parameter θ, when (F_1, F_2) is unknown:

(a) Let $\tilde{\beta}_p$ be an estimator β_p such that $n^{1/2}(\hat{\beta}_p - \beta_p) = O_p(1)$, for $p = 1, 2$.
(b) Compute the residuals $\tilde{\epsilon}_{pj} = y_{pj} - x_{pj}^T \tilde{\beta}_p$, for $p = 1, 2$ and $j = 1, \ldots, n$.
(c) Estimate $F_p(t)$ by $\tilde{F}_{pn}(t)$ defined by $\tilde{F}_{pn}(t) = \{1/(n+1)\}\Sigma_{i=1}^n I(\tilde{\epsilon}_{pi} \leq t)$, where I is the indicator function; thus $\tilde{F}_{pn}$ is the empirical distribution of $\{\tilde{\epsilon}_{p1}, \ldots, \tilde{\epsilon}_{pn}\}$, except for the denominator $(n+1)$ instead of n.
(d) Estimate θ by $\tilde{\theta}$, defined by $\tilde{\theta} = \text{argmax}_\theta \; L(\theta)$ where

$$L(\theta) = \Sigma \log c\{\tilde{F}_{1n}(\tilde{\epsilon}_{1i}), \tilde{F}_{2n}(\tilde{\epsilon}_{2i}); \theta\}.$$

This four-step procedure reduces to that proposed by Oakes (1994) and Genest et al. (1995) for the case when (Y_{1i}, Y_{2i}) are independent and identically distributed for $i = 1, \ldots, n$. Since $(\tilde{F}_{1n}, \tilde{F}_{2n})$ is expected to be close to (F_1, F_2) for large n, it is reasonable to expect that the foregoing estimator is likely to be a reasonable estimator.

We will show that $\tilde{\theta}$ is consistent and asymptotically normal, and obtain a closed form expression for the asymptotic variance. By substituting sample estimates for the asymptotic variance formulae, we shall obtain a consistent estimator of the large sample variance.

While the idea that underlies our new method is intuitively simple and is a natural extension of Oakes (1994) and Genest et al. (1995), the mathematical arguments to derive its essential properties are quite involved. Therefore, in this section we shall indicate the main ideas in a simple form and relegate the rigorous details to an appendix. Even there, only the main steps are indicated. More detailed and rigorous proofs are given in a working paper of the authors at Monash University.

To indicate the general approach, let $l(\theta, u_1, u_2) = \log\{c(u_1, u_2; \theta)\}$, and let l with subscripts θ, 1, and 2 denote partial derivatives with respect to θ, u_1 and u_2 respectively. For example, $l_\theta(\theta, u_1, u_2) = (\partial/\partial\theta)l(\theta, u_1, u_2)$ and $l_{\theta,1}(\theta, u_1, u_2) = (\partial^2/\partial u_1 \partial\theta)l(\theta, u_1, u_2)$. Now, let us first expand $L(\tilde{\theta})$ in Taylor series about the true value θ_0.

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

where θ^* lies in the line segment $[\theta_0, \tilde{\theta}]$. Now, solving this for $n^{1/2}(\tilde{\theta} - \theta_0)$, we have that

$$n^{1/2}(\tilde{\theta} - \theta_0) = A_n/(B_n + C_n) \quad (1)$$

where

$$A_n = n^{-1/2}\Sigma_{i=1}^n l_\theta\{\theta_0, \tilde{F}_{1n}(\tilde{\epsilon}_{1i}), \tilde{F}_{2n}(\tilde{\epsilon}_{2i})\}, \quad (2)$$

$$B_n = -n^{-1}\Sigma_{i=1}^n l_{\theta,\theta}\{\theta_0, \tilde{F}_{1n}(\tilde{\epsilon}_{1i}), \tilde{F}_{2n}(\tilde{\epsilon}_{2i}), \quad (3)$$

$$C_n = -(\tilde{\theta} - \theta_0)[(2n)^{-1}\Sigma_{i=1}^n l_{\theta,\theta,\theta}\{\theta^*, \tilde{F}_{1n}(\tilde{\epsilon}_{1i}), \tilde{F}_{2n}(\tilde{\epsilon}_{2i})\}]. \quad (4)$$

The main reasons for the technical details leading to the asymptotic properties of $\tilde{\theta}$ turns out to be complicated are that the expressions in (2)- (4) are sums of dependent random variables and $\tilde{F}_{pn}(\tilde{\epsilon}_{pi})$ is a non-smooth function. The dependence of the random variables will be dealt with by using results for U-statistics and *multivariate rank order statistics*. To deal with the non-smoothness due to the presence of $\tilde{F}_{pn}$, results for *weighted empirical processes* in Koul (2002) will be used.

By essentially mimicking the arguments in section 6.4 of Lehmann (1983), it can be shown that the estimator $\tilde{\theta}$ is consistent. To establish the asymptotic normality of $n^{1/2}(\tilde{\theta} - \theta_0)$, we consider the behaviour of the terms in (2) - (4). We will show that $\{\tilde{F}_{1n}(\tilde{\epsilon}_{1i}), \tilde{F}_{2n}(\tilde{\epsilon}_{2i})\}$ in (3) can be approximated by $\{F_1(\epsilon_{1i}), F_2(\epsilon_{2i})\}$ so that $B_n = -n^{-1}\Sigma_{i=1}^n l_{\theta,\theta}\{\theta_0, F_1(\epsilon_{1i}), F_2(\epsilon_{2i})\} + o_p(1)$, which converges to γ in probability, where

$$\gamma = E\big[\big(l_\theta\{\theta_0, F_1(\epsilon_1), F_2(\epsilon_2)\}\big)^2\big]. \tag{5}$$

By assuming that the third order derivatives $l_{\theta,\theta,\theta}$ are bounded by integrable functions in a small neighbourhood of θ_0, we have $n^{-1}\Sigma_{i=1}^n l_{\theta,\theta,\theta}\{\theta^*, \tilde{F}_{1n}(\tilde{\epsilon}_{1i}), \tilde{F}_{2n}(\tilde{\epsilon}_{2i})\} = O_p(1)$ and hence $C_n = o_p(1)$. Therefore, $B_n + C_n$ converges to γ in probability. To obtain the asymptotic distribution of A_n in (2), we approximate $\tilde{F}_{pn}$ by F_{pn}, the empirical distribution of the unobserved error terms rather than by its true population distribution F_p, $(p = 1, 2)$. This leads to $A_n = A_{n1} + o_p(1)$, where $A_{n1} = n^{-1/2}\Sigma_{i=1}^n l_\theta\{\theta_0, F_{1n}(\epsilon_{1i}), F_{2n}(\epsilon_{2i})\}$. It may be noted that A_{n1} is a multivariate rank order statistic of the form $n^{-1}\Sigma J\{F_{1n}(\epsilon_{1i}), F_{2n}(\epsilon_{2i})\}$ for some function J. The asymptotic distribution of such general rank order statistics have been studied extensively in the literature, for example see Rüschendorf and Ruschendorf (1976) and Ruymgaart et al. (1972). The asymptotic distribution of the foregoing particular form of A_{n1} was obtained by Genest et al. (1995). Applying Proposition 2.1 therein, we have that A_{n1} converges in distribution to $N(0, \sigma^2)$, where

$$\sigma^2 = \mathrm{var}\big[l_\theta\{\theta_0, F_1(\epsilon_1), F_2(\epsilon_2)\} + W_1(\epsilon_1) + W_2(\epsilon_2)\big], \tag{6}$$

$$W_p(\epsilon_p) = \int\int I\big(F_p(\epsilon_p) \le u_p\big) l_{\theta,p}\{\theta_0, u_1, u_2\} c(u_1, u_2; \theta) du_1 du_2, \quad p = 1, 2. \tag{7}$$

Thus, we conclude that $n^{1/2}(\tilde{\theta} - \theta_0)$ converges in distribution to $N(0, \nu^2)$, where $\nu^2 = \sigma^2/\gamma^2$.

To obtain an estimate of the asymptotic variance ν^2, we estimate σ^2 and γ^2 separately. By substituting estimated quantities to the unknown quantities in (5), an estimate of γ^2 is

$$\tilde{\gamma} = -n^{-1}\Sigma_{i=1}^n l_{\theta,\theta}\{\tilde{\theta}, \tilde{F}_{1n}(\tilde{\epsilon}_{1i}), \tilde{F}_{2n}(\tilde{\epsilon}_{2i})\}. \tag{8}$$

Since $\sigma^2 = \mathrm{var}\{T(\theta_0)\}$, where $T(\theta_0) = \{l_\theta\{\theta_0, F_1(\epsilon_1), F_2(\epsilon_2)\} + W_1(\epsilon_1) + W_2(\epsilon_2)\}$, and $T(\theta_0)$

cannot be observed, we estimate σ^2 by the sample variance

$$\tilde{\sigma}^2 = \text{ Sample variance of } \tilde{T}_1(\tilde{\theta}), \ldots, \tilde{T}_n(\tilde{\theta}), \tag{9}$$

of the pseudo observations,

$$\tilde{T}_i(\theta) = l_\theta\{\theta, \tilde{F}_{1n}(\tilde{\epsilon}_{1i}), \tilde{F}_{2n}(\tilde{\epsilon}_{2i})\} + \tilde{W}_1(\tilde{\epsilon}_{1i}, \theta) + \tilde{W}_2(\tilde{\epsilon}_{2i}, \theta), \quad i = 1, \ldots, n, \tag{10}$$

$$\text{where} \quad \tilde{W}_p(t, \theta) = n^{-1}\Sigma_{j=1}^n I(t \leq \tilde{\epsilon}_{pj}) l_{\theta,p}\{\theta, \tilde{F}_{1n}(\tilde{\epsilon}_{1j}), \tilde{F}_{2n}(\tilde{\epsilon}_{2j})\}, \quad p = 1, 2. \tag{11}$$

This leads to the consistent estimator $\tilde{\nu}^2 = \tilde{\sigma}^2/\tilde{s}^2$ for ν^2. Now, let us state the main theorem. A set of regularity conditions to ensure that the theorem holds, is stated in the Appendix where the proof of theorem is also given.

Theorem 0.1. *Assume that the regularity conditions given in the Appendix hold. Then, the semiparametric estimator $\tilde{\theta}$ 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 $\nu^2 = \sigma^2/\gamma^2$. Further, a consistent estimator $\tilde{\nu}^2$ of ν^2 is given by $\tilde{\nu}^2 = \tilde{\sigma}^2/\tilde{\gamma}^2$, where $\tilde{\sigma}^2$ and $\tilde{\gamma}$ are as in (9) and (8) respectively.*

The expression for the asymptotic variance of $n^{1/2}(\tilde{\theta} - \theta_0)$ is the same as that for the case when the there is no regression structure and the observations are independent and identically distributed. Therefore, Proposition 2.2 of Genest et al. (1995) is applicable to the setting in the above theorem as well. In particular, the semiparametric estimator $\tilde{\theta}$ is fully efficient for the independent copula, otherwise there is a loss of asymptotic efficiency due to the marginal distributions being unknown.

The parameter θ can be estimated by other methods as well. The two main ones that play central roles in inference for copulas are the maximum likelihood and the *inference function for margins* (see Joe (1997)). Both of these are fully parametric. Let $F_p(t; \alpha_p)$ denote the distribution of the error term ϵ_p for every p. Now the loglikelihood takes the form

$$\ell(\theta, \beta_1, \beta_2, \alpha_1, \alpha_2) = L(\theta, \beta_1, \beta_2, \alpha_1, \alpha_2) + B(\beta_1, \beta_2, \alpha_1, \alpha_2)$$

where $B(\beta_1, \beta_2, \alpha_1, \alpha_2) = \sum \log\{f_1(\epsilon_{1i}, \alpha_1)\ f_2(\epsilon_{2i}, \alpha_2)\}$ and $L(\theta, \beta_1, \beta_2, \alpha_1, \alpha_2) = \sum \log c\{F_1(\epsilon_{1i}, \alpha_1),\ F_2(\epsilon_{2i}, \alpha_2); \theta\}$.

The maximum likelihood estimator of $(\theta, \beta_1, \beta_2, \alpha_1, \alpha_2)$ is simply the maximizer of $\ell(\theta, \beta_1, \beta_2, \alpha_1, \alpha_2)$. The method of maximum likelihood for copulas is usually not the preferred one due to difficulties such as multiple maxima for the likelihood and erratic behaviour of the estimator. The inference function method has been proposed as a close alternative - see Joe (1997) for a thorough account of this topic. In this method, the model is estimated in two stages. In the first stage, the parameter (β_p, α_p) is estimated using the data for the p^{th} margin, for every p; let this estimator be denoted by $(\hat{\beta}_p, \hat{\alpha}_p)$. Then, in the second stage, θ is estimated by maximizing the likelihood function with (β_p, α_p) replaced by its estimator ($p = 1, 2$). Thus, the inference function for margins estimator of θ is $argmax_\theta\ \ell(\theta, \hat{\beta}_1, \hat{\beta}_2, \hat{\alpha}_1, \hat{\alpha}_2)$.

One would expect that the maximum likelihood and the inference function for margins methods are likely to be non-robust against misspecification of the marginal distributions. The simulation results in section 4 illustrate that the semiparametric method is considerably better than the maximum likelihood and inference function for margins method when the form of the marginal distributions are unknown, which is almost always the case in practice.

3. Simulation study

We carried out a simulation study to compare the semiparametric method with its competitors, the maximum likelihood and inference function for margins, and also to evaluate the reliability of the large sample confidence interval for θ given in Theorem 1.

Design of the simulation study

The following five copulas were considered in the study. More details about them may be found in Joe (1997) and Nelsen (1999).

(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:*

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

These copulas cover a very wide range of distributional shapes. The maximum likelihood and inference function for margins estimators that are used in this simulation study assumed that the marginal distributions were normal. The following sets of marginal distributions were considered: (1) X_1 and X_2 are normally distributed, (2) $X_1 \sim t_r$ and $X_2 \sim t_r$, (3) $X_1 \sim t_r$ and $X_2 \sim$ skew t_r with skewness = 0.5, and (4) $X_1 \sim t_r$ and $X_2 \sim \chi^2_2$. The first one corresponds to the correct specification of the marginal distributions, while each of the others leads to a misspecification of the model. A skew 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. The values 3 and 8 were considered for the degrees of freedom r of the t_r-distribution. Since the semiparametric method estimates each marginal distribution nonparametrically, it is meant to be used when the sample size is moderate to large. In this study, we considered sample sizes ranging from 50 to 1000. This captures a broad range of realistic settings.

All the computations were programmed in MATLAB Version 7.0.4. Optimizations were performed using the procedure "fmincon.m" in the "Optimization Toolbox (3.0.2).

Results:

Only a selection of the simulation results are presented here to save space. Overall, the difference between the inference functions for margins and maximum likelihood estimators were small, with the former performing slightly better. Therefore, the results for maximum likelihood are not presented here.

——— Tables 1-2 about here ————-

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 mis-specification. Thus, as expected, the inference function for margin estimators perform slightly better than the semiparametric estimator. However, the differences are small.

Each marginal distribution is incorrectly specified as Normal:

Table 2 provides estimated bias computed as the mean of the simulated estimates of θ minus the true value of θ. The same table also provides standard deviations of the simulated estimates of θ. Table 2 shows quite clearly that (i) the maximum likelihood and inference function for margin estimators are 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 maximum likelihood and the inference function for margin estimators of θ. We recognize that the very large values for relative MSE in Table 1 are not precise, but we presented them because they convey the message that misspecification of the marginal distribution may cause the parametric estimators to be biased and the standard deviation of the estimators could become relatively unimportant compared to the bias.

Table 3 shows that an approximate 95% confidence interval based on a normal approximation for the large sample distribution of $\tilde{\theta}$, has coverage rates close to 95% for sample size ≥ 40. In some isolated cases, it dropped to a rate just below 90 %. These results show that the semiparametric method also offers a reliable and easy to compute large sample confidence interval for θ.

4. An illustrative example

To illustrate the semiparametric method, we discuss an example that is very similar to that we discussed in the Introduction. The response variables are returns on shares of ANZ Bank and of BHP-Billiton. We consider regression of these variables on the All Ordinaries Index (Australia), a market index similar to the Dow Index in the USA. The variables are defined as follows: $y_{1t} = ln(A_t/A_{t-1}) - ln(T_t/T_{t-1})$, $y_{2t} = ln(B_t/B_{t-1}) - ln(T_t/T_{t-1})$, $z_t = ln(I_t/I_{t-1}) - ln(T_t/T_{t-1})$, where A_t = ANZ price index, B_t = BHP-Billiton price index, T_t = 90-day Treasury bill rates, and I_t = All Ordinaries Index. We used monthly data for the period July 1981 to July 2001.

We consider the regression model, $Y_{pt} = x_t^T\beta_p + \epsilon_{pt},\ p = 1, 2$, where $x_t = (1, z_t)^T$. In the first stage, we estimated β_1 and β_2 by least squares. The estimated models are, $y_{1t} = -0.064 + 0.840z_t + \tilde{\epsilon}_t$ and $y_{2t} = -0.470 + 0.992z_t + \tilde{\epsilon}_t$, respectively. Then we considered Ali-Mikhail-Haq, Clayton, Gumbel, Joe, and Independent copulas for the joint distribution of the error term, (ϵ_1, ϵ_2). Closed form expressions for the copulas were given in the previous section. We assessed the goodness of fit using the chi-square statistic with a grid of 20 cells. Since the models are nonnested and method is semiparametric, the distribution theory of the chisquare statistic is not available. However, it is reasonable to compare the chi-square statistics. Based on such diagnostics, we concluded that a Gumbel copula provided the best fit for the joint distribution of the error terms, although some of the others were not too different. The estimated value of the Gumbel copula parameter is $\tilde{\theta} = 1.076$ and the standard error = 0.046. Hence, an estimate of the joint distribution of (ϵ_1, ϵ_2) is $C(\tilde{F}_1(\epsilon_1), \tilde{F}_2(\epsilon_2); 1.076)$, and an estimate of the joint distribution of (y_1, y_2) conditional on x is

$$C\{\tilde{F}_1\{y_1 - (-0.064 + 0.840z)\}, \tilde{F}_2\{y_2 - (-0.470 + 0.992z)\}; 1.076], \tag{12}$$

where $\tilde{F}_1$ and $\tilde{F}_2$ are the empirical distributions of the residuals of the error terms in the two regression models. This estimated distribution can be used for estimating various quantities associated with risk. For example, it can be used to estimate (i) the probability of the random components of returns, namely ϵ_1 and ϵ_2, falling below a_1 and a_2 respectively where a_1 and a_2 are given, (ii) the probability, $\text{pr}(Y_1 \leq a_1, Y_2 \leq a_2 \mid z)$, of the returns falling below a_1 and a_2 for a given value of the explanatory variable z, and (iii) *Value-at-Risk* c of the portfolio $w = b_1 Y_1 + (1 - b_1) Y_2$, defined by $\text{pr}\{b_1 Y_1 + (1 - b_1) Y_2 \leq c \mid z\} \leq \alpha$, where α is a given small number, for example $\alpha = 0.05$, and b_1 is the proportion of investment in the first asset.

As an example, Table 4 provides estimates of Value-at-Risk of the portfolio for several values of b_1 and z. This is very similar to those for the example on pages 68-69 in Cherubini et al. (2004). Table 4 shows that as the proportion b_1 moves closer to 50%, the Value-at-Risk decreases indicating that portfolio diversification reduces risk, and the rate at which the Value-at-Risk decreases is indicative of the effectiveness of diversification on reducing risk. Table 4 also shows that the Value-at-Risk is almost symmetric about $b_1 = 50\%$, which is consistent with our observation that the histograms, not given here, of the regression residuals for the two assets appear to have approximately equally heavy tails.

The chi-square goodness fit statistics for a 5 by 4 grid, turned out to be 5.2 and 380 for the semiparametric method and for the inference function for margins method with normal distribution for each margin, respectively. Therefore, the the former method provided a significantly better fit than the latter one.

5. Conclusion

We extended a semiparametric estimator of Oakes (1994) and Genest et al. (1995) for estimating the dependence parameter and the joint distribution of the error terms of the multivariate linear regression model. We showed that this estimator is asymptotically normal. It turns out that the form of the asymptotic variance is very similar to that obtained by Genest et al. (1995) for the case when the observations are independent and identically distributed. This helped us to use his results and construct consistent estimates for the asymptotic variance and confidence interval for the dependence parameter. Simulation results showed that our semiparametric estimator performs better than the parametric ones when true error distribution deviates from that assumed by the parametric methods, maximum likelihood and inference function for margins. Further, the semiparametric method is fully efficient for the independent copula, which extends a result of Genest et al. (1995) for case of independent and identically distributed observations. Since the form of the expression for the asymptotic variance of the semiparametric estimator is very similar to that when the observations are independent and identically distributed, we would expect that the conditions in Genest and Werker (2002) for the semiparametric estimator to be efficient are also likely to be applicable in the regression case as well.

Acknowledgment

This research was partially supported by the Australian Research Council.

Appendix: Proofs

Here we shall indicate the main steps in the proof of Theorem 1. A more detailed proof is provided in an unpublished manuscript. As in the text, the index p refers to the pth component, $p = 1, 2$; for simplicity 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), and let (U_1, U_2) denote a random variable with the same distribution as $(F_1(\epsilon_1), F_2(\epsilon_2))$ so that $(U_1, U_2) \sim C(u_1, u_2; \theta_0)$. Now, let us introduce the regularity conditions.

Condition C:

(C.1): The distribution function F_p has continuously differentiable density, denoted by f_p and it satisfies $\|f_p\|_\infty < \infty$ and $\left\|f'_p\right\|_\infty < \infty$ where f'_p is the first derivative of f_p.

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

(C.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 $|d_j| \leq \varepsilon_\theta$.

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

(C.5): The covariate x_p is non-stochastic, $n^{-1} X_p^T X_p$ converges to a positive definite matrix, and $n^{-1/2} \max_i \|x_{pi}\| \to 0$ as $n \to \infty$ for $p = 1, 2$.

(C.6): $n^{1/2}(\tilde{\beta}_p - \beta_p) = O_p(1)$, $p = 1, 2$.

Remark: The proof given below assumes that $\|x_{pi}\|$, $i = 1, 2, \ldots$, is bounded, but the results hold under the weaker assumption, $n^{-1/2} \max_i \|x_{pi}\| \to 0$, in (C.5).

Lemma 1. $\sup_i \left| [F_{pn}(\tilde{\epsilon}_{pi}) - F_{pn}(\epsilon_{pi})] - \left[F_p(\tilde{\epsilon}_{pi}) - F_p(\epsilon_{pi})\right] \right| = o_p(n^{-\frac{1}{2}})$.

Proof. Let $Z = F_p(\epsilon_p)$ and $Z_i = F_p(\epsilon_{pi})$, $i = 1, \ldots, n$. Then, $Z_1, \ldots, Z_n$ are independently and identically distributed taking values in [0,1]. Let $W(t) = n^{-1/2}\sum\{I(Z_i \le t) - t\}$. Then,

$$\sup_i \Big| F_{pn}(\tilde\epsilon_{pi}) - F_{pn}(\epsilon_{pi}) - \big[F_p(\tilde\epsilon_{pi}) - F_p(\epsilon_{pi})\big]\Big| = n^{-1/2}\sup_i \Big| W_d\{F_p(\tilde\epsilon_{pi})\} - W_d\{F_p(\epsilon_{pi})\}\Big|$$

$$\le n^{-1/2}\sup_{|t-s|<\delta} |W_d(t) - W_d(s)|, \text{ with arbitrary large probability for any } \delta > 0.$$

Now, the proof follows from

$$\lim_{\delta\to 0}\limsup_n \operatorname{pr}\Big(\sup_{|t-s|<\delta} |W_d(t) - W_d(s)| > \epsilon\Big) = 0$$

for any $\epsilon > 0$, by Theorem 2.2.1 in Koul (2002). □

Lemma 2. $\sup_i |\tilde F_{pn}(\tilde\epsilon_{pi}) - F_{pn}(\tilde\epsilon_{pi}) - \bar x_p^T(\tilde\beta_p - \beta_p) f_p(\tilde\epsilon_{pi})| = o_p(n^{-1/2})$.

Proof. Let $S_d^0(t,u) = n^{-1/2}\sum_{j=1}^n I\big\{\epsilon_{pj} \le t + n^{-1/2}x_{pj}^T u\big\}$. It follows from Theorem 2.3.1 in Koul (2002) that for any $b > 0$,

$$\sup_{-\infty<t<\infty, \|u\|<b} \Big| n^{1/2}\Big\{ S_d^0(t,u) - S_d^0(t,0) - u^T \bar x_p f_p(t)\Big\}\Big| = o_p(1). \tag{13}$$

See page 192 in Shorack and Wellner (1986) for a similar result. Let $\tilde u = n^{1/2}(\tilde\beta - \beta)$. Then, $S_d^0(t,\tilde u) = \tilde F_{pn}(t)$ and

$$\sup_t |\tilde F_{pn}(t) - F_{pn}(t) - \bar x_p^T(\tilde\beta_p - \beta_p) f_p(t)| = n^{-1/2}\sup_t \Big| n^{1/2}\Big\{ S_d^0(t,\tilde u) - S_d^0(t,0) - \bar x_p^T \tilde u f_p(t)\Big\}\Big|.$$

Now, the proof follows from (13) since $\|\tilde u\| < b$ with arbitrarily large probability for sufficiently large $b > 0$. □

Let $\vartheta_{pi} = F_{pn}(\tilde\epsilon_{pi}) - F_{pn}(\epsilon_{pi})$, $\delta_{pi} = \tilde F_{pn}(\tilde\epsilon_{pi}) - F_{pn}(\epsilon_{pi})$, $\delta_{pi}^* = \tilde F_{pn}(\tilde\epsilon_{pi}) - F_p(\epsilon_{pi})$, $\eta_{pi} = \tilde F_{pn}(\tilde\epsilon_{pi}) - F_{pn}(\tilde\epsilon_{pi})$, and $\xi_{pi} = F_p(\tilde\epsilon_{pi}) - F_p(\epsilon_{pi})$. Then, we have

Lemma 3. $\sup_i |\xi_{pi}|$, $\sup_i |\vartheta_{pi}|$, $\sup_i |\delta_{pi}|$, $\sup_i |\delta_{pi}^*|$, $\sup_i |\eta_{pi}|$ *are all of order* $O_p(n^{-1/2})$.

Proof. The proof for $\sup_i |\vartheta_{pi}|$ follows from Lemma 1. The proof for η_{pi} follows from Lemma 2. The proof for δ_{pi} follows from $\sup_i |\delta_{pi}| \le \sup_i |\eta_{pi}| + \sup_i |\vartheta_{pi}|$ and the previous parts. The proof for δ_{pi}^* follows from $\sup_i |\delta_{pi}^*| \le \sup_i |\delta_{pi}| + \sup_i |F_{pn}(\epsilon_{pi}) - F_p(\epsilon_{pi})|$, the last term being $O_p(n^{-1/2})$ since it is the empirical process for independent and identically distributed random variables. □

Lemma 4. *Let $\Psi(\theta, u_1, u_2)$ and $G(u_1, u_2)$ be the functions defined in Condition (C.3). Also let $\{d^n_{pi}\}$ be a sequence of random variables such that $\sup_i |d^n_{pi}| = O_p(n^{-1/2})$. Then, for any given θ in a small neighbourhood of θ_0, we have that*

$$n^{-1/2}\Sigma_{i=1}^n|\Psi\{\theta, F_1(\epsilon_{1i}) + d^n_{1i}, F_2(\epsilon_{2i}) + d^n_{2i}\} - \Psi\{\theta, F_1(\epsilon_{1i}), F_2(\epsilon_{2i})\}| = O_p(1),$$

$$n^{-1/2}\Sigma_{i=1}^n|\Psi\{\theta, F_{1n}(\epsilon_{1i}) + d^n_{1i}, F_{2n}(\epsilon_{2i}) + d^n_{2i}\} - \Psi\{\theta, F_{1n}(\epsilon_{1i}), F_{2n}(\epsilon_{2i})\}| = O_p(1),$$

$$n^{-1/2}\Sigma_{i=1}^n|G\{F_1(\epsilon_{1i}) + d^n_{1i}, F_2(\epsilon_{2i}) + d^n_{2i}\} - G\{F_1(\epsilon_{1i}), F_2(\epsilon_{2i})\}| = O_p(1).$$

Proof. To prove the first part, note that,

$$\begin{aligned}
& n^{-1}\Sigma_{i=1}^n|\Psi\{\theta, F_1(\epsilon_{1i}) + d^n_{1i}, F_2(\epsilon_{2i}) + d^n_{2i}\} - \Psi\{\theta, F_1(\epsilon_{1i}), F_2(\epsilon_{2i})\}| \\
& \leq n^{-1}\Sigma_{i=1}^n k\{F_1(\epsilon_{1i}), F_2(\epsilon_{2i}); \theta\}(|d^n_{1i}| + |d^n_{2i}|) \\
& \leq (\sup_i |d^n_{1i}| + \sup_i |d^n_{2i}|)\; n^{-1}\Sigma_{i=1}^n k\{F_1(\epsilon_{1i}), F_2(\epsilon_{2i}); \theta\} = O_p(n^{-\frac{1}{2}})O_p(1) = O_p(n^{-\frac{1}{2}}).
\end{aligned}$$

The other two parts follow through repeated application similar arguments and the triangle inequality. □

Lemma 5. $\quad n^{-2}\Sigma_{j=1}^n\Sigma_{i=1}^n[I\{\tilde\epsilon_{pi} \leq \tilde\epsilon_{pj}\} - I\{\epsilon_{pi} \leq \epsilon_{pj}\}]^2 = o_p(1), \quad$ *for* $p = 1, 2$.

Proof. Let $\delta_{nij} = (x_{pi} - x_{pj})^T(\tilde\beta_p - \beta_p)$ and δ be a given positive number. Then $\text{pr}\{\max_{ij} |\delta_{nij}| < \delta\} \to 1$. Now,

$$\begin{aligned}
& n^{-2}\Sigma_i\Sigma_j\ |I(\tilde\epsilon_{pj} \leq \tilde\epsilon_{pi}) - I(\epsilon_{pj} \leq \epsilon_{pi})| = n^{-2}\Sigma_i\Sigma_j\ |I(\epsilon_{pj} \leq \epsilon_{pi} + \delta_{nij}) - I(\epsilon_{pj} \leq \epsilon_{pi})|, \\
& \leq n^{-2}\Sigma_i\Sigma_j\ I(|\epsilon_{pj} - \epsilon_{pi}| \leq |\delta_{nij}|) \leq n^{-2}\Sigma_i\Sigma_j\ I(|\epsilon_{pj} - \epsilon_{pi}| \leq \delta),\ \text{with probability approaching 1}
\end{aligned}$$

Since the last expression is essentially a U-statistic, it converges in probability to $h(\delta)$ where $h(\delta) = 2E[I(|\epsilon_{p1} - \epsilon_{p2}| < \delta)]$. The proof follows since, as is easily seen, $h(\delta)$ is continuous at $\delta = 0$ and $h(0) = 0$. □

Let

$$\hat{W}_p(t,\theta) = n^{-1}\Sigma_{j=1}^{n} I\big(t \le \epsilon_{pj}\}l_{\theta,p}\{\theta, F_1(\epsilon_{1j}), F_2(\epsilon_{2j})\}, \tag{14}$$

$$\tilde{W}_p(t,\theta) = n^{-1}\Sigma_{j=1}^{n} I\{t \le \tilde{\epsilon}_{pj}\}l_{\theta,p}\{\theta, \tilde{F}_{1n}(\tilde{\epsilon}_{1j}), \tilde{F}_{2n}(\tilde{\epsilon}_{2j})\}, \tag{15}$$

and $$T_i(\theta) = l_\theta\{\theta, F_1(\epsilon_{1i}), F_2(\epsilon_{2i})\} + \hat{W}_1(\epsilon_{1i},\theta) + \hat{W}_2(\epsilon_{2i},\theta). \tag{16}$$

Lemma 6. $|n^{-1}\Sigma_{i=1}^{n}\{\tilde{T}_i(\theta_0) - T_i(\theta_0)\}| = o_p(1).$

Proof. $|n^{-1}\Sigma_{i=1}^{n}\{\tilde{T}_i(\theta_0) - T_i(\theta_0)\}| \le \Sigma_{j=1}^{3} T_j^n$, where

$$T_1^n = |n^{-1}\Sigma_{i=1}^{n}\{l_\theta\{\theta_0, \tilde{F}_{1n}(\tilde{\epsilon}_{1i}), \tilde{F}_{2n}(\tilde{\epsilon}_{2i})\} - l_\theta\{\theta_0, F_1(\epsilon_{1i}), F_2(\epsilon_{2i})\}\}|,$$

$$T_2^n = |n^{-1}\Sigma_{i=1}^{n}\{\tilde{W}_1(\tilde{\epsilon}_{1i},\theta_0) - \hat{W}_1(\epsilon_{1i},\theta_0)\}|,$$

$$T_3^n = |n^{-1}\Sigma_{i=1}^{n}\{\tilde{W}_2(\tilde{\epsilon}_{2i},\theta_0) - \hat{W}_2(\epsilon_{2i},\theta_0)\}|.$$

We will show that, $T_j^n = o_p(1)$ for $i, j = 1, 2, 3$. First, it may be seen that $T_1^n = O_p(n^{-1/2})$ by Lemma 4. To show that $T_2^n = o_p(1)$, note that

$$\begin{aligned} T_2^n &= |n^{-1}\Sigma_{i=1}^{n}\{\tilde{W}_1(\tilde{\epsilon}_{1i},\theta_0) - \hat{W}_1(\epsilon_{1i},\theta_0)\}| \\ &\le n^{-2}\Sigma_{i=1}^{n}\Sigma_{j=1}^{n}|I\{\tilde{\epsilon}_{1i} \le \tilde{\epsilon}_{1j}\}|\ |\{l_{\theta,1}\{\theta_0, \tilde{F}_{1n}(\tilde{\epsilon}_{1j}), \tilde{F}_{2n}(\tilde{\epsilon}_{2j})\} - l_{\theta,1}\{\theta_0, F_1(\epsilon_{1j}), F_2(\epsilon_{2j})\}\}| \\ &+ |n^{-2}\Sigma_{i=1}^{n}\Sigma_{j=1}^{n}\{I\{\tilde{\epsilon}_{1i} \le \tilde{\epsilon}_{1j}\} - I\{\epsilon_{1i} \le \epsilon_{1j}\}\}l_{\theta,1}\{\theta_0, F_1(\epsilon_{1j}), F_2(\epsilon_{2j})\}|. \end{aligned}$$

It may be verified that the first term is $o_p(1)$ using Lemma 4. Now, consider the second term:

$$\begin{aligned} &\Big|n^{-2}\Sigma_{i=1}^{n}\Sigma_{j=1}^{n}\Big\{I\big(\tilde{\epsilon}_{1i} \le \tilde{\epsilon}_{1j}\big) - I\big(\epsilon_{1i} \le \epsilon_{1j}\big)\Big\}l_{\theta,1}\{\theta_0, F_1(\epsilon_{1j}), F_2(\epsilon_{2j})\}\Big| \\ &\le \sup_j \Big|n^{-1}\Sigma_{i=1}^{n}\Big\{I\big(\tilde{\epsilon}_{1i} \le \tilde{\epsilon}_{1j}\big) - I\big(\epsilon_{1i} \le \epsilon_{1j}\big)\Big\}\Big|\ n^{-1}\Sigma_{j=1}^{n}\Big|l_{\theta,1}\{\theta_0, F_1(\epsilon_{1j}), F_2(\epsilon_{2j})\}\Big| \\ &= \sup_j |\delta_{1i}|\ \ n^{-1}\Sigma_{j=1}^{n}\Big|l_{\theta,1}\{\theta_0, F_1(\epsilon_{1j}), F_2(\epsilon_{2j})\}\Big| = o_p(1). \end{aligned}$$

by Lemma 3. Hence, $T_2^n = o_p(1)$. Similarly, $T_3^n = o_p(1)$. □

We need another Lemma to show that $\tilde{\sigma}^2$ is a consistent estimator of σ^2.

Lemma 7. *Let $\tilde{T}_i(\theta)$ and $T_i(\theta)$ be defined as in (10) and (16). 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: $\{\tilde{T}_i(\theta)\}^2$, $\{(\partial/\partial\theta)\tilde{T}_i(\theta)\}^2$, $\{T_i(\theta)\}^2$, $\{(\partial/\partial\theta)T_i(\theta)\}^2$.*

Proof. First, recall that $\tilde{T}_i(\theta) = \tilde{T}_{i1}(\theta)+\tilde{T}_{i2}(\theta)+\tilde{T}_{i3}(\theta)$, where $\tilde{T}_{i1}(\theta) = l_\theta\{\theta, \tilde{F}_{1n}(\tilde{\epsilon}_{1i}), \tilde{F}_{2n}(\tilde{\epsilon}_{2i})\}$, $\tilde{T}_{i2}(\theta) = \tilde{W}_1(\tilde{\epsilon}_{1i}, \theta)$, $\tilde{T}_{i3}(\theta) = \tilde{W}_2(\tilde{\epsilon}_{2i}, \theta)$. Now, by Cauchy-Schwartz inequality, to prove the first part, it suffices to establish that the lemma holds with G_{ni} replaced by $\tilde{T}_{ij}$, $j = 1, 2, 3$. Now consider $\tilde{T}_{i1}(\theta)$:

$$
\begin{aligned}
&\sup_{\theta\in\mathbf{N}(\theta_0)} n^{-1}\Sigma_{i=1}^n\{\tilde{T}_{i1}(\theta)\}^2 = \sup_{\theta\in\mathbf{N}(\theta_0)} n^{-1}\Sigma_{i=1}^n[l_\theta\{\theta, \tilde{F}_{1n}(\tilde{\epsilon}_{1i}), \tilde{F}_{2n}(\tilde{\epsilon}_{2i})\}]^2 \\
&\le n^{-1}\Sigma_{i=1}^n[G\{F_1(\epsilon_{1i}) + \delta_{1i}^*, F_2(\epsilon_{2i}) + \delta_{2i}^*\}]^2 \\
&\le n^{-1}\Sigma_{i=1}^n[G\{F_1(\epsilon_{1i}), F_2(\epsilon_{2i})\}]^2 \qquad (17) \\
&+ 2\{\sup_i |\delta_{1i}^*| + \sup_i |\delta_{2i}^*|\}n^{-1}\Sigma_{i=1}^n k^*\{F_1(\epsilon_{1i}), F_2(\epsilon_{2i})\}G\{F_1(\epsilon_{1i}), F_2(\epsilon_{2i})\} \\
&+ \{\sup_i |\delta_{1i}^*| + \sup_i |\delta_{2i}^*|\}^2 n^{-1}\Sigma_{i=1}^n(k^*\{F_1(\epsilon_{1i}), F_2(\epsilon_{2i})\})^2 = O_p(1).
\end{aligned}
$$

The claims about the other terms can also be established by applying similar arguments, although the complete proof is long. □

Lemma 8. $n^{-1}\sum_{i=1}^n\{\tilde{T}_i(\theta_0) - T_i(\theta_0)\}^2 = o_p(1)$.

Proof. The summand can be expressed as $\{T_{i1}^n + T_{i2}^n + T_{i3}^n\}^2$, and hence it suffices to establish that $n^{-1}\sum_{i=1}^n\{T_{ij}^n(\theta_0)\}^2 = o_p(1)$, for $j = 1, 2, 3$, where $T_{i1}^n = \{l_\theta\{\theta_0, \tilde{F}_{1n}(\tilde{\epsilon}_{1i}), \tilde{F}_{2n}(\tilde{\epsilon}_{2i})\} - l_\theta\{\theta_0, F_1(\epsilon_{1i}), F_2(\epsilon_{2i})\}\}$, $T_{i2}^n = \{(\tilde{W}_1(\tilde{\epsilon}_{1i}, \theta_0) - \hat{W}_1(\epsilon_{1i}, \theta_0)\}$, $T_{i3}^n = \{\tilde{W}_2(\tilde{\epsilon}_{2i}, \theta_0) - \hat{W}_2(\epsilon_{2i}, \theta_0)\}$. We shall indicate the proof for one term; the rest of the claims can be established by similar arguments and appealing to the earlier lemmas. To show that $n^{-1}\sum_{i=1}^n\{T_{i2}^n\}^2 = o_p(1)$, let

$$
\begin{aligned}
T_{i21}^n &= n^{-1}\Sigma_{j=1}^n I(\tilde{\epsilon}_{1i} \le \tilde{\epsilon}_{1j})\ [l_{\theta,1}\{\theta_0, \tilde{F}_{1n}(\tilde{\epsilon}_{1j}), \tilde{F}_{2n}(\tilde{\epsilon}_{2j})\} - l_{\theta,1}\{\theta_0, F_1(\epsilon_{1j}), F_2(\epsilon_{2j})\}], \\
T_{i22}^n &= n^{-1}\Sigma_{j=1}^n[I(\tilde{\epsilon}_{1i} \le \tilde{\epsilon}_{1j}) - I(\epsilon_{1i} \le \epsilon_{1j})]\ l_{\theta,1}\{\theta_0, F_1(\epsilon_{1j}), F_2(\epsilon_{2j})\}.
\end{aligned}
$$

By Cauchy-Schwartz inequality, it suffices to establish that $n^{-1}\Sigma_{i=1}^n\{T_{i2j}^n\}^2 = o_p(1)$, for $j = 1, 2$. The proof involves breaking the terms into separate parts and applying Cauchy-Schwartz inequality and the previous lemmas. For example, $n^{-1}\Sigma_{i=1}^n(T_{i22}^n)^2$ is less than or equal to

$$
\{n^{-2}\Sigma_{j=1}^n\Sigma_{i=1}^n[I\{\tilde{\epsilon}_{1i} \le \tilde{\epsilon}_{1j}\} - I\{\epsilon_{1i} \le \epsilon_{1j}\}]^2\}\ \ n^{-1}\Sigma_{j=1}^n\{l_{\theta,1}\{\theta_0, F_1(\epsilon_{1j}), F_2(\epsilon_{2j})\}^2, \qquad (18)
$$

which is of order $o_p(1)$. Similarly, the other terms can also be shown to be of order $o_p(1)$, which completes the proof. $\square$

Now, the proof of the consistency of the estimator follows essentially the same arguments as in section 6.4 of Lehmann (1983), in particular Theorem 4.1. The intermediate arguments required for this are contained in the Lemmas established thus far. The main approach is that $(\tilde{F}_{pn}, \tilde{\epsilon}_{pi})$ can be replaced by (F_{pn}, ϵ_{pi}) for $p = 1, 2$ and $i = 1, \ldots, n$ in the derivatives of $n^{-1}\Sigma_{i=1}^{n} l_\theta\{\theta_0, F_{1n}(\epsilon_{1i}), F_{2n}(\epsilon_{2i})\}$, because the remainder terms can be shown to be negligible.

Proof of the asymptotic normality of $n^{1/2}(\tilde{\theta} - \theta_0)$:

We shall prove that the numerator in (1) converges to a normal distribution and that the denominator converges to the constant γ in probability. The main approach to obtaining the asymptotic distribution of the numerator is to avoid expanding it about the true value of the unknown parameters but to ensure that the first term is the rank order statistic in terms of the errors $\{\epsilon_{pi}\}$. Thus, we express the numerator A_n in (1) as $A_n = \sum_{k=1}^{6} A_{nk}$, where

$$A_{n1} = n^{-1/2}\Sigma_{i=1}^{n} l_\theta\{\theta_0, F_{1n}(\epsilon_{1i}), F_{2n}(\epsilon_{2i})\} \quad (19)$$

$$A_{n2} = n^{-1/2}\Sigma_{i=1}^{n} \delta_{1i}\ l_{\theta,1}\{\theta_0, F_{1n}(\epsilon_{1i}), F_{2n}(\epsilon_{2i})\} \quad (20)$$

$$A_{n3} = n^{-1/2}\Sigma_{i=1}^{n} \delta_{2i}\ l_{\theta,2}\{\theta_0, F_{1n}(\epsilon_{1i}), F_{2n}(\epsilon_{2i})\} \quad (21)$$

$$A_{n4} = n^{-1/2}\Sigma_{i=1}^{n} \delta_{1i}\ \delta_{2i}\ l_{\theta,1,2}\{\theta_0, F_{1n}(\epsilon_{1i}) + c_1\delta_{1i}, F_{2n}(\epsilon_{2i}) + c_2\delta_{2i}\} \quad (22)$$

$$A_{n5} = n^{-1/2}\Sigma_{i=1}^{n} 2^{-1}(\delta_{1i})^2\ l_{\theta,1,1}\{\theta_0, F_{1n}(\epsilon_{1i}) + c_1\delta_{1i}, F_{2n}(\epsilon_{2i}) + c_2\delta_{2i}\} \quad (23)$$

$$A_{n6} = n^{-1/2}\Sigma_{i=1}^{n} 2^{-1}(\delta_{2i})^2\ l_{\theta,2,2}\{\theta_0, F_{1n}(\epsilon_{1i}) + c_1\delta_{1i}, F_{2n}(\epsilon_{2i}) + c_2\delta_{2i}\}, \quad (24)$$

for some $0 < c_1, c_2 < 1$. The next two lemmas show that $A_{nj} = o_p(1)$, for $j = 2, \ldots, 6$.

Lemma 9. $|A_{nj}| = o_p(1)$*, for* $j = 2, 3$*.*

Proof. Let $A^*_{n2i} = (\bar{x}_1 - x_{1i})^T(\tilde{\beta}_1 - \beta_1) f_1(\epsilon_{1i}) l_{\theta,1}\{\theta_0, F_1(\epsilon_{1i}), F_2(\epsilon_{2i})\}$. We will show that

$n^{-1/2}\Sigma_{i=1}^{n}A_{n2i}^{*} = o_p(1)$ and $A_{n2} - n^{-1/2}\Sigma_{i=1}^{n}A_{n2i}^{*} = o_p(1)$. First note that

$$|n^{-1/2}\Sigma_{i=1}^{n}A_{n2i}^{*}| \le |n^{-1}\Sigma_{i=1}^{n}\big(\bar{x}_1 - x_{1i}\big)^{T} f_1(\epsilon_{1i}) l_{\theta,1}\{\theta_0, F_1(\epsilon_{1i}), F_2(\epsilon_{2i})\}| \; \|\sqrt{n}(\tilde{\beta}_1 - \beta_1)\| = o_p(1). \tag{25}$$

Now, let us write $A_{n2} - n^{-1/2}\Sigma_{i=1}^{n}A_{n2i}^{*} = B_1 + n^{1/2}(\tilde{\beta} - \beta)^T B_2$ where

$$B_1 = n^{-1/2}\Sigma_{i=1}^{n}[\{\delta_{1i} - (\bar{x}_1 - x_{1i})^T(\tilde{\beta}_1 - \beta_1) f_1(\epsilon_{1i})\} l_{\theta,1}\{\theta_0, F_{1n}(\epsilon_{1i}), F_{2n}(\epsilon_{2i})\}]$$

$$\text{and } B_2 = n^{-1}\Sigma_{i=1}^{n}(\bar{x}_1 - x_{1i})^T f_1(\epsilon_{1i})[l_{\theta,1}\{\theta_0, F_{1n}(\epsilon_{1i}), F_{2n}(\epsilon_{2i})\} - l_{\theta,1}\{\theta_0, F_1(\epsilon_{1i}), F_2(\epsilon_{2i})\}].$$

Let us write $\delta_{1i} = \{F_{pn}(\tilde{\epsilon}_{pi}) - F_{pn}(\epsilon_{pi})\} + \tilde{F}_{pn}(\tilde{\epsilon}_{pi}) - F_{pn}(\tilde{\epsilon}_{pi})$. Now, using Lemmas 1 and 2 to approximate δ_{1i}, and separating terms to apply the triangle inequality several times, we have $B_1 = o_p(1)$. The presence of the term $(\bar{x}_1 - x_{1i})$ in the summand of B_2 ensures that $B_2 = o_p(1)$. The proof follows by combining these results. □

Lemma 10. *For $j \in \{4, 5, 6\}$, $|A_{nj}| = o_p(1)$.*

Proof. Let, for $p = \{1, 2\}$, let $d_{npi} = F_{pn}(\epsilon_{pi}) - F_p(\epsilon_{pi}) + c_p\delta_{pi}$, where $c_p\delta_{pi}$ is defined in (19). Then, $\sup_i \Big|d_{npi}\Big| \le \sup_i \Big|F_{pn}(\epsilon_{pi}) - F_p(\epsilon_{pi})\Big| + \sup_i |c_p\delta_{pi}| = O_p(n^{-1/2})$. Therefore, we have

$$\begin{aligned}
|A_{n4}| &= \big|n^{1/2}[n^{-1}\Sigma_{i=1}^{n}\delta_{1i}\ \delta_{2i} l_{\theta,1,2}\{\theta_0, F_{1n}(\epsilon_{1i}) + c_1\delta_{1i}, F_{2n}(\epsilon_{2i}) + c_2\delta_{2i}\}]\big| \\
&\le n^{1/2}\sup_i |\delta_{1i}|\ \sup_i |\delta_{2i}|\ n^{-1}\Sigma_{i=1}^{n}\Big|l_{\theta,1,2}\{\theta_0, F_1(\epsilon_{1i}) + d_{n1i}, F_2(\epsilon_{2i}) + d_{n2i}\}\Big| \\
&= n^{1/2}\ \sup_i |\delta_{1i}| \sup_i |\delta_{2i}|\ \Big(n^{-1}\Sigma_{i=1}^{n}\Big|l_{\theta,1,2}\{\theta_0, F_1(\epsilon_{1i}), F_2(\epsilon_{2i})\}\Big| + O_p(n^{-1/2})\Big) \\
&\le n^{1/2}\ \sup_i |\delta_{1i}|\ \ \sup_i |\delta_{2i}|\ \Big(n^{-1}\Sigma_{i=1}^{n}G\{F_1(\epsilon_{1i}), F_2(\epsilon_{2i})\} + O_p(n^{-1/2})\Big) = o_p(1).
\end{aligned}$$

By similar arguments, we also have A_{n5} and A_{n6} are also of order $o_p(1)$. □

It follows from the foregoing two Lemmas that $A_n = A_{n1} + o_p(1)$ and hence A_n and A_{n1} have the same asymptotic distributions. The asymptotic distribution of the rank order statistic A_{n1} was obtained in Genest et al. (1995). It follows from the results therein that A_{n1} converges in distribution to $N(0, \sigma^2)$ where

$$\sigma^2 = \text{var}[l_\theta\{\theta_0, F_1(\epsilon_1), F_1(\epsilon_1)\} + W_1(\epsilon_1) + W_2(\epsilon_2)],$$

$$\text{and} \quad W_p(\epsilon_p) = \int\int I\big(F_p(\epsilon_p) \le u_p\big) l_{\theta,p}\{\theta_0, u_1, u_2\} c(u_1, u_2; \theta) du_1 du_2.$$

To complete the proof, we shall show that $C_n = o_p(1)$ and $|B_n - \gamma| = o_p(1)$.

$$|C_n| = \left|(2n)^{-1}\Sigma_{i=1}^n(\tilde{\theta} - \theta_0)l_{\theta,\theta,\theta}\{\theta^*, \tilde{F}_{1n}(\tilde{\epsilon}_{1i}), \tilde{F}_{2n}(\tilde{\epsilon}_{2i})\}\right|$$

$$\leq (1/2)\|\tilde{\theta} - \theta_0\|n^{-1}\Sigma_{i=1}^n G\{\tilde{F}_{1n}(\tilde{\epsilon}_{1i}), \tilde{F}_{2n}(\tilde{\epsilon}_{2i})\}, \text{ with probability close to 1, for large } n$$

$$= (1/2)\|\tilde{\theta} - \theta_0\|[n^{-1}\Sigma_{i=1}^n G\{F_1(\epsilon_{1i}), F_2(\epsilon_{2i})\} + O_p(n^{-1/2})] = o_p(1).$$

Now, to prove the convergence of B_n, note that

$$|B_n - \gamma| \leq \left| - n^{-1}\Sigma_{i=1}^n l_{\theta,\theta}\{\theta_0, \tilde{F}_{1n}(\tilde{\epsilon}_{1i}), \tilde{F}_{2n}(\tilde{\epsilon}_{2i})\} + n^{-1}\Sigma_{i=1}^n l_{\theta,\theta}\{\theta_0, F_1(\epsilon_{1i}), F_2(\epsilon_{2i})\}\right|$$

$$+ \left| - n^{-1}\Sigma_{i=1}^n l_{\theta,\theta}\{\theta_0, F_1(\epsilon_{1i}), F_2(\epsilon_{2i})\} - \gamma\right|.$$

The first term on the right hand side converges to zero in probability by Lemma 4 and the second term also converges to zero in probability by the Weak Law of Large Numbers. This completes the proof of the asymptotic normality of $n^{1/2}(\tilde{\theta} - \theta_0)$.

The proof of the consistency of $\tilde{\nu}$ is established by showing that $\tilde{\sigma} = \sigma + o_p(1)$ and $\tilde{\gamma} = \gamma + o_p(1)$. These proofs use the lemmas established thus far as the building blocks. The proof of $\tilde{\gamma} = \gamma + o_p(1)$ follows by applying a Law of Large Numbers for independent random variables, U-statistics and rank order statistics. The proof of $\tilde{\sigma} = \sigma + o_p(1)$ is long but follows arguments similar those used in the previous parts. All of these are given in Kim et al. (2005).

References

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.

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

Genest, C. and Werker, B. J. M. (2002). Conditions on the asymptotic semiparametric efficiency of an omnibus estimator of dependence parameters in copula models. In C. M. Cuadras, J. Fortiana, and J. A. Rodríguez-Lallela, editors, *Distributions with Given Marginals and Statistical Modelling*, pages 103–112. Kluwer, Dordrecht, The Netherlands.

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

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. *Presented at the Joint Statistical Meetings, Minneapolis, August 2005.*

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

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

Nelsen, R. B. (1999). *An Introduction to Copulas*. Springer-Verlag Inc.

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

Oakes, D. and Ritz, J. (2000). Regression in a bivariate copula model. *Biometrika*, **87**, 345–352.

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

Rüschendorf, L. and Ruschendorf, L. (1976). Asymptotic distributions of multivariate rank order statistics (Corr: V10 p1311). *The Annals of Statistics*, **4**, 912–923.

Ruymgaart, F. H., Shorack, G. R., and van Zwet, W. R. (1972). Asymptotic normality of nonparametric tests for independence. *The Annals of Mathematical Statistics*, **43**, 1122–1135.

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

Shorack, G. R. and Wellner, J. A. (1986). *Empirical Processes with Applications to Statistics*. Wiley: New York.

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

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.

	100 observations				500 observations			
θ	(N-N)	(T-T)	(T-ST)	(T-C)	(N-N)	(T-T)	(T-ST)	(T-C)
				Ali-Mikhail-Haq faimly of Copulas				
-0.7	100	107	112	118	100	221	152	112
-0.3	97	176	168	136	99	377	273	143
0.1	97	181	199	194	99	334	551	257
0.5	93	124	182	195	100	468	1100	1260
				Frank Copula				
-2.0	93	307	357	190	98	2010	1610	565
-0.5	96	322	382	235	101	690	990	318
3.5	94	353	291	182	96	2180	1850	501
5.0	92	277	285	147	95	1970	1740	408
				Gumbel Copula				
1.5	88	276	300	192	93	18400	831	548
3.0	98	211	318	466	98	410	824	2260
4.5	93	203	649	1020	101	378	2330	3800
6.0	94	227	883	990	95	385	4210	3900
				Joe Copula				
1.5	80	178	224	150	82	320	802	870
3.0	93	254	344	332	90	326	1020	1680
4.5	85	170	581	923	85	288	1480	3270
6.0	87	152	965	935	93	359	2350	4570
				Plackett Copula				
0.5	97	168	213	125	98	555	754	330
2.0	93	636	512	229	97	2730	2780	800
3.5	91	339	432	201	99	3230	2650	618
5.0	90	302	401	165	97	2180	1940	470

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 χ^2_2. The number of repeated samples is 250.

Table 2: Estimated means and standard deviations when the marginal distributions are t_3 and $\chi^2(2)$ 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							
-0.7	-0.75	0.31	-0.74	0.28	-0.74	0.17	-0.69	0.17
-0.3	-0.33	0.40	-0.31	0.34	-0.35	0.18	-0.31	0.16
0.1	0.14	0.47	0.08	0.33	0.13	0.20	0.09	0.12
0.5	0.63	0.29	0.47	0.22	0.77	0.14	0.51	0.09
	Frank Copula							
-2.0	-2.41	0.89	-2.00	0.71	-2.53	0.41	-1.99	0.28
-0.5	-0.81	0.72	-0.61	0.50	-0.70	0.40	-0.49	0.25
3.5	3.96	0.88	3.52	0.73	4.16	0.45	3.52	0.33
5.0	5.31	0.85	4.95	0.75	5.58	0.44	4.97	0.36
	Gumbel Copula							
1.5	1.36	0.16	1.53	0.15	1.36	0.08	1.51	0.07
3.0	2.36	0.34	3.00	0.33	2.36	0.17	2.98	0.14
4.5	2.93	0.46	4.29	0.47	2.91	0.32	4.37	0.23
6.0	3.28	0.57	5.42	0.67	3.22	0.57	5.70	0.34
	Joe Copula							
1.5	1.29	0.15	1.58	0.20	1.25	0.07	1.51	0.09
3.0	2.30	0.38	3.06	0.43	2.23	0.20	3.01	0.19
4.5	2.92	0.53	4.37	0.53	2.95	0.38	4.49	0.28
6.0	3.54	0.78	5.81	0.82	3.37	0.52	5.89	0.38
	Plackett Copula							
0.5	0.45	0.18	0.51	0.17	0.41	0.07	0.50	0.07
2.0	2.44	0.83	2.10	0.61	2.56	0.54	2.03	0.27
3.5	4.17	1.27	3.57	1.01	4.38	0.67	3.52	0.44
5.0	5.82	1.78	5.18	1.51	6.06	0.82	4.98	0.62

Table 3: The estimated coverage rates (in %) of a 95% large sample confidence interval of the copula parameter.

	Sample sizes											
	40				100				500			
θ	N-N	T-C	T-ST	T-T	N-N	T-C	T-ST	T-T	N-N	T-C	T-ST	T-T
					AMH copula							
-0.70	94	91	94	94	96	96	95	95	93	96	93	96
-0.30	91	90	93	94	94	96	96	93	94	95	91	95
0.10	93	91	89	93	91	93	92	92	94	91	97	95
0.50	88	89	87	90	92	93	95	94	96	94	95	92
					Frank copula							
-0.50	98	96	96	96	96	97	96	96	97	98	98	98
-2.00	95	95	95	97	94	95	96	94	93	94	95	96
3.50	97	94	98	95	97	95	95	96	95	97	95	98
5.00	98	95	94	96	95	95	96	97	97	92	95	93
					Gumbel copula							
1.50	95	97	96	97	96	96	95	97	97	98	95	95
3.00	95	97	97	97	96	95	96	95	94	95	94	96
4.50	94	93	96	94	96	90	94	95	94	91	94	94
6.00	96	79	93	93	92	82	93	96	95	86	96	93
					Joe copula							
1.50	94	96	96	95	99	96	95	94	96	97	96	95
3.00	95	97	99	96	95	98	94	96	94	95	96	95
4.50	95	97	96	92	94	95	95	95	94	93	95	94
6.00	96	94	91	92	93	92	94	92	95	97	92	93
					Plackett copula							
0.50	89	90	91	91	94	94	95	94	96	97	95	96
2.00	89	92	91	94	96	92	94	94	96	98	97	96
3.50	94	92	92	93	95	97	97	97	97	99	99	99
5.00	96	96	96	93	97	96	98	96	99	99	99	99

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 χ^2_2. The number of repeated samples is 250.

Table 4: Value at Risk corresponding to $\alpha = 5\%$.

z	Percentage invested on Bank shares				
	10	25	50	75	90
-8	-16.2	-14.9	-14.0	-14.7	-15.6
-6	-14.2	-13.0	-12.2	-13.0	-13.9
-4	-12.3	-11.1	-10.4	-11.2	-12.2
-2	-10.3	-9.2	-8.5	-9.5	-10.5
0	-8.4	-7.3	-6.7	-7.7	-8.8
2	-6.41	-5.39	-4.88	-5.95	-7.07
4	-4.45	-3.49	-3.05	-4.20	-5.35
6	-2.50	-1.58	-1.21	-2.44	-3.64