Statistica Sinica **19** (2009), 53-70

STATISTICAL INFERENCE FOR MULTIVARIATE RESIDUAL COPULA OF GARCH MODELS

Ngai-Hang Chan, Jian Chen, Xiaohong Chen, Yanqin Fan and Liang Peng

Chinese University of Hong Kong, Sage SB Inc., Yale University Vanderbilt University and Georgia Institute of Technology

Abstract: Recently a flexible class of semiparametric copula-based multivariate GARCH models has been proposed to quantify multivariate risks, in which univariate GARCH models are used to capture the dynamics of individual financial series, and parametric copulas are used to model the contemporaneous dependence among GARCH residuals with nonparametric marginals. In this paper we address two questions regarding statistical inference for this class of models. (1) Under what mild sufficient conditions is the asymptotic distribution of the pseudo maximum likelihood estimator (MLE) of the residual copula parameter of Chen and Fan (2006a) justified? (2) How do we test the correct specification of a parametric copula for the GARCH residuals? In order to answer both questions rigorously, we establish a new weighted approximation for the empirical distributions of the GARCH residuals, which is of interest in its own right. Simulation studies and data examples are provided to examine the finite sample performance of the pseudo MLE of the residual copula parameter and the proposed goodness-of-fit test.

Key words and phrases: Copula, GARCH, goodness-of-fit test, pseudo maximum likelihood estimation, residual empirical distribution.

1. Introduction

On June 26, 2004, governors of the G-10 central banks endorsed the publication of the revised capital accord, known as Basel II, in which the Basel Committee proposed to adopt a more holistic approach that focuses on the interaction between the different risk categories in risk management; see McNeil, Frey and Embrechts (2005) for a succinct account of the developments of Basel II. In order to comply with Basel II, banks face the critical issue of adequately modeling dependence between different risk factors. Since copulas capture dependence structures among individual risk factors that are invariant to any monotonic transformation of the individual risks, they have become standard tools in risk management. On the insurance front, the International Actuarial Association recommends using copulas for modeling dependence structure of insurance portfolios in Solvency II. Since then, major software providers have built various copula models to serve the industrial needs. For details on copulas, we refer to Joe (1997) and Nelsen (2005). Because individual risk series in finance and insurance are typically serially dependent, Chen and Fan (2006a) introduced a class of Semiparametric COpula-based Multivariate DYnamic (SCOMDY) models, in which the conditional mean and conditional variance of individual risk series are parametrically specified, but the joint distribution of the (standardized) innovations is semiparametrically specified as a parametric copula evaluated at the nonparametric marginals. This class of models is very flexible in capturing a wide range of temporal and contemporaneous dependence structures of multivariate (nonlinear) time series.

An important class of the SCOMDY models is the so-called semiparametric copula-based multivariate GARCH models, where a scalar GARCH model is used to capture volatility of individual risk series and a parametric copula is used to model the contemporaneous dependence between different risks. We formally introduce this class of SCOMDY models. Suppose the observations $\{Y_t = (Y_{1,t}, \ldots, Y_{r,t})^T\}_{t=1}^n$ satisfy

$$Y_{j,t} = \sqrt{h_{j,t}}\epsilon_{j,t}, \quad h_{j,t} = c_j + \sum_{i=1}^{p_j} \alpha_{j,i} Y_{j,t-i}^2 + \sum_{i=1}^{q_j} \beta_{j,i} h_{j,t-i}, \quad j = 1, \ldots, r, \quad (1.1)$$

where $\{\epsilon_t = (\epsilon_{1,t}, \ldots, \epsilon_{r,t})^T\}_{t=1}^n$ is a sequence of i.i.d. random vectors with $E[\epsilon_{j,t}] = 0, E[(\epsilon_{j,t})^2] = 1$, and the joint distribution function F_ϵ of ϵ_t is assumed to take the semiparametric form:

$$F_\epsilon(\epsilon_1, \ldots, \epsilon_r) = C(F_{\epsilon,1}(\epsilon_1), \ldots, F_{\epsilon,r}(\epsilon_r); \theta_0). \quad (1.2)$$

Here $C(x_1, \ldots, x_r; \theta)$ is a parametrized copula function up to unknown $\theta \in \Theta \subset R^m$, and for $j = 1, \ldots, r$, $F_{\epsilon,j}$ is the marginal distribution function of $\epsilon_{j,t}$, assumed to be continuous but otherwise unspecified. By Sklar's Theorem (see Nelsen (2005)), any multivariate distribution with continuous marginals can be uniquely represented by its copula function evaluated at its marginals. Let C_ϵ denote the unique copula corresponding to the true joint distribution F_ϵ of the GARCH residual vector ϵ_t. We call C_ϵ the residual copula. It is defined as

$$C_\epsilon(x_1, \ldots, x_r) = F_\epsilon(F_{\epsilon,1}^-(x_1), \ldots, F_{\epsilon,r}^-(x_r)),$$

where $F_{\epsilon,j}^-(\cdot)$ is the generalized inverse of $F_{\epsilon,j}(\cdot)$, $j = 1, \ldots, r$. Model (1.2) effectively assumes that the true residual copula belongs to a parametric family: $C_\epsilon(x_1, \ldots, x_r) = C(x_1, \ldots, x_r; \theta_0)$ for some unknown $\theta_0 \in \Theta \subset R^m$.

By fitting the semiparametric distribution (1.2) to the GARCH residuals, dimensionality is reduced from r to m, which leads to more efficient estimation of the copula parameter. As a result, measures of portfolio risks such as conditional VaR can be more efficiently estimated; see Hull and White (1998) and Breymann, Dias and Embrechts (2003) for applications to exchange rate data, and Giacomini, Härdle, Ignatieva and Spokoiny (2008) to stock data. This simple multivariate GARCH model bypasses the overparametrization issue that is commonly encountered in generalizing univariate GARCH to multivariate GARCH models. For a survey on multivariate GARCH models, see Bauwens, Laurent and Rombouts (2006). The class of semiparametric copula-based multivariate GARCH models (1.1)–(1.2) provides a plausible and efficient means to manage multiple financial risk factors. However, before they can be readily used by banks and insurance companies, valid statistical inference methodologies must be developed. In particular, estimation of the residual copula parameter and tests for the correct specification of the parametric residual copula are of importance and are addressed in this paper.

Estimation and inference for copulas that directly couple multivariate observed variables have been pursued extensively. For example, in the context of nonparametric copulas, Fermanian, Radulovic and Wegkamp (2004) considered empirical copula estimation, while Fermanian and Scaillet (2003) and Chen and Huang (2007) proposed kernel smoothing. For parametric copulas coupled with nonparametric marginals, Genest, Ghoudi and Rivest (1995) investigated pseudo maximum likelihood estimation (MLE), while Chen, Fan and Tsyrennikov (2006) considered sieve MLE. Chen and Fan (2006b) studied the pseudo MLE and its properties in estimating copulas that generate nonlinear Markov models. For i.i.d. data, Klugman and Parsa (1999), Fermanian (2005) Scaillet (2007) and Genest, Quessy and Rmillard (2006) examined goodness-of-fit tests of parametric copulas. Chen and Fan (2005) developed model selection tests for multiple parametric copula comparison.

The main technical difficulty in establishing the asymptotic distribution of the pseudo MLE of the copula parameter is that the score function and its derivatives in copula-based models can blow up to infinity near the boundaries. Chen and Fan (2005, 2006b) overcome this difficulty by making use of the weak convergence of the empirical distribution function in a weighted metric. Currently, there are sporadic results on the convergence of empirical distributions using residuals of non-linear time series; see, for example, Berkes and Horváth (2003), Horváth, Kokoszka and Teyssiére (2001) and Koul and Ling (2006). However, to the best of our knowledge, a weighted approximation result is not available for empirical distributions of residuals obtained from an initial step estimation of time series models. Although Chen and Fan (2006a) developed copula model selection tests for SCOMDY models, their tests rely on the asymptotic property of the pseudo MLE of the residual copula parameter θ. Let $\widehat{\theta}$ denote this estimator (see Section 2 for its definition). Crucial to the validity of their model selection tests is the surprising result that the asymptotic distribution of $\widehat{\theta}$ is not affected by the initial step estimation of the GARCH parameters. Chen and Fan (2006a) established this result by means of a heuristic argument with stringent conditions, and by assuming the validity of a weighted approximation for the empirical distributions of GARCH residuals.

In this paper, we first establish a weighted approximation for the empirical distributions of residuals of univariate GARCH models, which is important in its own right. This weighted approximation allows us to provide a rigorous justification of the limiting distribution result for the pseudo MLE $\widehat{\theta}$ under mild sufficient conditions; see Section 2. In addition, we develop a consistent test for the correct specification of the residual copula $C_\epsilon(x_1, \ldots, x_r)$ by a particular parametric copula class $\mathcal{C} = \{C(x_1, \ldots, x_r; \theta) : \theta \in \Theta\}$. This extends existing goodness-of-fit tests for i.i.d. data to GARCH residuals. In Section 3, we provide some simulation studies and data examples to demonstrate finite sample properties of the pseudo MLE for θ, and the goodness-of-fit test for the parametric copula. All proofs can be found in the on-line supplement of this paper.

2. Estimation and Testing

2.1. Estimation of GARCH models

For each $j = 1, \ldots, r$, let $\gamma_j = (c_j, \alpha_{j,1}, \ldots, \alpha_{j,p_j}, \beta_{j,1}, \ldots, \beta_{j,q_j})^T$ denote the true GARCH parameters associated with the model (1.1). Let $\hat{\gamma}_j = (\hat{c}_j, \hat{\alpha}_{j,1}, \ldots, \hat{\alpha}_{j,p_j}, \hat{\beta}_{j,1}, \ldots, \hat{\beta}_{j,q_j})^T$ denote the quasi MLE of γ_j based on the sample $\{Y_{j,t}\}_{t=1}^n$, which is the MLE if $\epsilon_{j,t}$ is standard normal.

Similar to Berkes and Horváth (2003), for $q_j \geq p_j$, define

$$\begin{cases} d_{j,0}(\gamma_j) = \frac{c_j}{(1-\beta_{j,1}-\cdots-\beta_{j,q_j})} \\ d_{j,1}(\gamma_j) = \alpha_{j,1} \\ d_{j,2}(\gamma_j) = \alpha_{j,2} + \beta_{j,1} d_{j,1}(\gamma_j) \\ \quad \vdots \\ d_{j,p_j}(\gamma_j) = \alpha_{j,p_j} + \beta_{j,1} d_{j,p_j-1}(\gamma_j) + \cdots + \beta_{j,p_j-1} d_{j,1}(\gamma_j) \\ d_{j,p_j+1}(\gamma_j) = \beta_{j,1} d_{j,p_j}(\gamma_j) + \cdots + \beta_{j,p_j} d_{j,1}(\gamma_j) \\ \quad \vdots \\ d_{j,q_j}(\gamma_j) = \beta_{j,1} d_{j,q_j-1}(\gamma_j) + \cdots + \beta_{j,q_j-1} d_{j,1}(\gamma_j); \end{cases}$$

for $q_j < p_j$, define

$$\begin{cases} d_{j,0}(\gamma_j) = \frac{c_j}{(1-\beta_{j,1}-\cdots-\beta_{j,q_j})} \\ d_{j,1}(\gamma_j) = \alpha_{j,1} \\ d_{j,2}(\gamma_j) = \alpha_{j,2} + \beta_{j,1}d_{j,1}(\gamma_j) \\ \quad\vdots \\ d_{j,q_j+1}(\gamma_j) = \alpha_{j,q_j+1} + \beta_{j,1}d_{j,q_j}(\gamma_j) + \cdots + \beta_{j,q_j}d_{j,1}(\gamma_j) \\ \quad\vdots \\ d_{j,p_j}(\gamma_j) = \alpha_{j,p_j} + \beta_{j,1}d_{j,p_j-1}(\gamma_j) + \cdots + \beta_{j,q_j}d_{j,p_j-q_j}(\gamma); \end{cases}$$

for $i > \max(p_j, q_j)$, define

$$d_{j,i}(\gamma_j) = \beta_{j,1}d_{j,i-1}(\gamma_j) + \beta_{j,2}d_{j,i-2}(\gamma_j) + \cdots + \beta_{j,q_j}d_{j,i-q_j}(\gamma_j).$$

Set $w_{j,k}(\gamma_j) = d_{j,0}(\gamma_j) + \sum_{i=1}^{\infty} d_{j,i}(\gamma_j)Y_{j,k-i}^2$ and

$$\Gamma_j = \{u = (u_1, \ldots, u_{p_j+q_j+1})^T : u > 0, u_{p_j+2} + \cdots + u_{p_j+q_j+1} \le \Delta_0^* < 1, \\ 0 < \Delta_1^* \le \min(u_1, \ldots, u_{p_j+q_j+1}) \le \max(u_1, \ldots, u_{p_j+q_j+1}) \le \Delta_2^*, q_j\Delta_1^* < \Delta_0^*\}.$$

Remark 1. When $E\epsilon_{j,1}^4 < \infty$ and $\gamma_j \in \Gamma_j$, it follows from Berkes and Horváth (2003, equations 1.8 and 3.4) that

$$\begin{cases} h_{j,k} = w_{j,k}(\gamma_j) \\ \hat{\gamma}_j - \gamma_j = \frac{1}{n}\sum_{t=1}^{n}(\epsilon_{j,t}^2 - 1)A_j^{-1}\frac{w'_{j,t}(\gamma_j)}{w_{j,t}(\gamma_j)} + o_p(n^{-\frac{1}{2}}), \end{cases} \tag{2.1}$$

where $A_j = E\{\frac{w'_{j,1}(\gamma_j)}{w_{j,1}(\gamma_j)}(\frac{w'_{j,1}(\gamma_j)}{w_{j,1}(\gamma_j)})^T\}$ and

$$w'_{j,t}(\gamma_j) = \Big(\frac{\partial}{\partial c_j}w_{j,t}(\gamma_j), \frac{\partial}{\partial \alpha_{j,1}}w_{j,t}(\gamma_j), \ldots, \frac{\partial}{\partial \beta_{j,q_j}}w_{j,t}(\gamma_j)\Big)^T.$$

Put $\hat{w}_{j,1}(\gamma_j) = 1$ and $\hat{w}_{j,k}(\gamma_j) = d_{j,0}(\gamma_j) + \sum_{i=1}^{k-1} d_{j,i}(\gamma_j)Y_{j,k-i}^2$ for $2 \le k \le n$. Then we can estimate ϵ_t by

$$\hat{\epsilon}_t = (\hat{\epsilon}_{1,t}, \ldots, \hat{\epsilon}_{r,t})^T = \Big(\frac{Y_{1,t}}{\sqrt{\hat{w}_{1,t}(\hat{\gamma}_j)}}, \ldots, \frac{Y_{r,t}}{\sqrt{\hat{w}_{r,t}(\hat{\gamma}_j)}}\Big)^T. \tag{2.2}$$

2.2. Estimation of residual copula parameters

We can estimate the true marginal distribution of $\epsilon_{j,t}$, $F_{\epsilon,j}(x)$, by

$$\hat{F}_{\epsilon,j}(x) = \frac{1}{n-\nu+1}\sum_{t=\nu}^{n} I(\hat{\epsilon}_{j,t} \le x),$$

where $\nu = \nu(n)$ is an integer. We then estimate the residual copula parameter θ by $\hat{\theta}$, the pseudo MLE based on the pseudo sample

$$\left\{(\hat{F}_{\epsilon,1}(\hat{\epsilon}_{1,t}), \ldots, \hat{F}_{\epsilon,r}(\hat{\epsilon}_{r,t}))^T\right\}_{t=\nu}^{n}, \tag{2.3}$$

i.e.,

$$\begin{aligned}\hat{\theta} &= \arg\max_{\theta} \frac{1}{n-\nu+1}\sum_{t=\nu}^{n} \log c(\hat{F}_{\epsilon,1}(\hat{\epsilon}_{1,t}), \ldots, \hat{F}_{\epsilon,r}(\hat{\epsilon}_{r,t}); \theta) \\ &:= \arg\max_{\theta} l_n(\theta),\end{aligned}$$

where $c(x_1, \ldots, x_r; \theta) = \partial^r C(x_1, \ldots, x_r; \theta)/\partial x_1 \cdots \partial x_r$ is the copula density function. This estimation approach was employed by Genest, Ghoudi and Rivest (1995) for independent data, and by Chen and Fan (2006a) for dependent data.

2.3. Weighted approximation for residual empirical distributions

Let U be a Gaussian process with

$$EU(x) = 0, \quad E\{U(x)U(y)\} = \prod_{i=1}^{r}\{x_i \wedge y_i\} - \prod_{i=1}^{r}\{x_i y_i\},$$

where $x = (x_1, \ldots, x_r)^T$ and $y = (y_1, \ldots, y_r)^T$. The following conditions are imposed for the study of the empirical process and the weighted empirical process of the estimated residuals of GARCH models.

A1. For $j = 1, \ldots, r$, $\gamma_j \in \Gamma_j$, $E\epsilon_{j,1}^4 < \infty$, and there exists $\mu > 0$ such that $\lim_{t\to 0} t^{-\mu} P(\epsilon_{j,1}^2 \le t) = 0$.

A2. For $j = 1, \ldots, r$, the support of $\epsilon_{j,t}$ is $(-\infty, \infty)$, $F_{\epsilon,j}$ has continuous density $F'_{\epsilon,j}$, and there exist $\beta_3 \in (0, 1/4)$ and $\Delta_3 > 0$ such that

$$\sup_{s} \sup_{|x-1|\le\Delta_3} \frac{sF'_{\epsilon,j}(sx)}{\{F_{\epsilon,j}(s)(1 - F_{\epsilon,j}(s))\}^{\beta_3}} < \infty$$

for $j = 1, \ldots, r$.

Theorem 2.1. *Suppose* A1 *and* A2 *hold and* $\nu/\log n \to \infty$, $\nu/n \to 0$ *as* $n \to \infty$. *Then*

$$\begin{aligned}&\sup_{x} \frac{|\sqrt{n-\nu+1}\{\hat{F}_{\epsilon,j}(x) - F_{\epsilon,j}(x)\} - U((1, \ldots, 1, F_{\epsilon,j}(x), 1, \ldots, 1)^T) - \frac{1}{2}xF'_{\epsilon,j}(x)\tau_j|}{\{F_{\epsilon,j}(x)(1 - F_{\epsilon,j}(x))\}^{\beta_3}} \\ &\quad = o_p(1),\end{aligned} \tag{2.4}$$

where $(U(T_r(x_1,\ldots,x_r)),\tau_1,\ldots,\tau_r)^T$ *is a vector valued Gaussian process with zero mean, and covariance structure*

$$E(\tau_j\tau_i)=E\{(\epsilon_{j,1}^2-1)(\epsilon_{i,1}^2-1)\}E\{(\frac{w'_{j,1}(\gamma_j)}{w_{j,1}(\gamma_j)})^T A_j^{-1} \times E(\frac{w'_{j,1}(\gamma_j)}{w_{j,1}(\gamma_j)})(\frac{w'_{i,1}(\gamma_i)}{w_{i,1}(\gamma_i)})^T A_i^{-1}E(\frac{w'_{i,1}(\gamma_i)}{w_{i,1}(\gamma_i)})\}, \tag{2.5}$$

$$E\{U(T_r(x_1,\ldots,x_r))\tau_j\} = E\{(\epsilon_{j,1}^2-1)I(F_1(\epsilon_{1,1})\le x_1,\ldots,F_r(\epsilon_{r,1})\le x_r)\} \times E\{(\frac{w'_{j,1}(\gamma_j)}{w_{j,1}(\gamma_j)})^T\}A_j^{-1}E\{\frac{w'_{j,t}(\gamma_j)}{w_{j,t}(\gamma_j)}\}, \tag{2.6}$$

$T_r(x_1,\ldots,x_r)$ *to be defined.*

Remark 2. Theorem 2.1 establishes a weighted approximation for the empirical distributions of the residuals $(\epsilon_{j,t})$ in GARCH models. Later on, we use it to derive the asymptotic distributions of the pseudo MLE of the residual copula parameter θ, and of the goodness-of-fit test statistic for testing the parametric specification of the residual copula. For (unweighted) approximation to the empirical process of squared residuals $(\epsilon_{j,t}^2)$ in ARCH and GARCH models, we refer to Horváth, Kokoszka and Teyssiére (2001) and Berkes and Horváth (2003).

2.4. Asymptotic properties of the pseudo MLE of θ

Let θ_0 denote the true value of θ and assume the following conditions for consistency of $\hat\theta$.

- C1. $\log c(x_1,\ldots,x_r;\theta)$ is a continuous function of θ for each $(x_1,\ldots,x_r)^T\in[0,1]^r$.
- C2. Θ is a compact subset of R^m.
- C3. $E\sup_{\theta\in\Theta}|\log c(F_{\epsilon,1}(\epsilon_{1,1}),\ldots,F_{\epsilon,r}(\epsilon_{r,1});\theta)|<\infty$.
- C4. For any $\Delta_0\in(0,1/2)$ and $\Delta_1\in(1/2,1)$, there exist $\beta_0\in(0,1)$, $M_0>0$, $\beta_1>0$ and $M_1>0$ such that

$$\sup_{\theta\in\Theta}|\log c(x_1,\ldots,x_r;\theta)|\le M_0\{\wedge_{i=1}^r x_i\}^{-\beta_0}$$

for $\wedge_{i=1}^r x_i\le\Delta_0$,

$$\sup_{\theta\in\Theta}|\log c(x_1,\ldots,x_r;\theta)|\le M_0\{1-\vee_{i=1}^r x_i\}^{-\beta_0}$$

for $\vee_{i=1}^r x_i \ge \Delta_1$, and

$$\sup_{\theta\in\Theta} |\log c(x_1,\ldots,x_r;\theta) - \log c(y_1,\ldots,y_r;\theta)| < M_1 \sum_{i=1}^r |x_i - y_i|^{\beta_1}$$

for $\Delta_0/2 < \wedge_{i=1}^r x_i \le \vee_{i=1}^r x_i < \Delta_1 + (1-\Delta_1)/2$ and $\Delta_0/2 < \wedge_{i=1}^r y_i \le \vee_{i=1}^r y_i < \Delta_1 + (1-\Delta_1)/2$.

Note that Conditions C1–C3 are standard conditions for consistency of MLE based on i.i.d. data. Condition C4 is similar to that imposed by Genest, Ghoudi and Rivest (1995) and Chen and Fan (2005) for i.i.d. data; it controls the speed of divergence of the logarithm of the copula density at the boundaries, and is satisfied by all the commonly used copula densities.

Theorem 2.2. *Suppose that A1 and C1–C4 hold, and $\nu(n)/n \to 0$ as $n \to \infty$. Then $\hat\theta \xrightarrow{p} \theta_0$ as $n \to \infty$.*

Remark 3. Since the approximation between the estimated residuals $\hat\epsilon_{j,t}$ and the residuals $\epsilon_{j,t}$ is poor for small t, we employ ν in the above Theorem to obtain a good approximation rate. The same idea was employed in Hall and Yao (2003) for deriving the limiting distribution of Quasi-MLE for GARCH models.

Before we state the asymptotic normality result, we introduce some additional notations. Put

$$\dot c(x_1,\ldots,x_r;\theta) = (\frac{\partial}{\partial\theta_1} c(x_1,\ldots,x_r;\theta),\ldots,\frac{\partial}{\partial\theta_m} c(x_1,\ldots,x_r;\theta))^T,$$
$$\delta(x_1,\ldots,x_r;\theta) = \frac{\dot c(x_1,\ldots,x_r;\theta)}{c(x_1,\ldots,x_r;\theta)},$$

and for $i = 1,\ldots,r$, $\delta_i(x_1,\ldots,x_r;\theta) = \partial\delta(x_1,\ldots,x_r;\theta)/\partial x_i$. Define $C_1(x_1) = x_1$,

$$C_i(x_i|x_1,\ldots,x_{i-1}) = P(F_{\epsilon,i}(\epsilon_{i,1}) \le x_i | F_{\epsilon,1}(\epsilon_{1,1}) = x_1,\ldots,F_{\epsilon,i-1}(\epsilon_{i-1,1}) = x_{i-1})$$

for $i = 2,\ldots,r$, $T_i(x_1,\ldots,x_i) = (C_1(x_1), C_2(x_2|x_1),\ldots,C_i(x_i|x_1,\ldots,x_{i-1}))$ for $i = 1,\ldots,r$, and

$$\Sigma(\theta) = \Big(E\Big\{\frac{\partial^2}{\partial\theta_i\partial\theta_j}\log c(F_{\epsilon,1}(\epsilon_{1,1}),\ldots,F_{\epsilon,r}(\epsilon_{r,1});\theta)\Big\}\Big)_{1\le i,j\le m}.$$

We impose the following additional regularity conditions for asymptotic normality.

N1. For $j = 2,\ldots,r$, the function $C_j(x_j|T_{j-1}^-(x_1,\ldots,x_{j-1}))$ is differentiable with respect to $x_1,\ldots,x_{j-1}$ over the interior of $[0,1]^{j-1}$, and

$$\sum_{i=1}^{j-1}\int_{[0,1]^{j-1}} |\frac{\partial}{\partial x_i} C_j(x_j|T_{j-1}^-(x_1,\ldots,x_{j-1}))|\,dx_1\cdots dx_{j-1} \le M_2 \in (0,\infty).$$

N2. There exists $\beta_2 \in (0, 1/2)$ such that

$$\sup_{0 \le x_1, \ldots, x_r \le 1} \prod_{i=1}^{r} (x_i)^{\beta_2} (1 - x_i)^{\beta_2} |\delta(T_r^-(x_1, \ldots, x_r); \theta_0)| < \infty,$$

$$\int \prod_{i=1}^{r} (x_i)^{\beta_2} (1 - x_i)^{\beta_2} |\, d\delta(T_r^-(x_1, \ldots, x_r; \theta_0))| < \infty.$$

N3. For any $\Delta_4 \in (0, 1/2)$ and $\Delta_5 \in (1/2, 1)$, there exist $\beta_4 \in (0, \beta_2)$, $M_4 > 0$, $\beta_5 > 0$ and $M_5 > 0$ such that

$$|\delta_j(x_1, \ldots, x_r; \theta_0)| \le M_4 x_j \{\wedge_{i=1}^r x_i\}^{-\beta_4}$$

for $\wedge_{i=1}^r x_i \le \Delta_4$,

$$|\delta_j(x_1, \ldots, x_r; \theta_0)| \le M_4 (1 - x_j) \{1 - \vee_{i=1}^r x_i\}^{-\beta_4}$$

for $\vee_{i=1}^r x_i \ge \Delta_5$, and

$$|\delta(x_1, \ldots, x_r; \theta_0) - \delta(y_1, \ldots, y_r; \theta_0)| < M_5 \sum_{i=1}^{r} |x_i - y_i|^{\beta_5}$$

for $\Delta_4/2 < \wedge_{i=1}^r x_i \le \vee_{i=1}^r x_i < \Delta_5 + (1 - \Delta_5)/2$ and $\Delta_4/2 < \wedge_{i=1}^r y_i \le \vee_{i=1}^r y_i < \Delta_5 + (1 - \Delta_5)/2$.

N4. For $1 \le i, j \le m$, $\partial^2 \log c(x_1, \ldots, x_r; \theta)/\partial\theta_i \partial\theta_j$ is a continuous function of θ in an open neighborhood of θ_0 for each $(x_1, \ldots, x_r)^T \in [0, 1]^r$.

N5. There exists an open neighborhood Θ_0 of θ_0 such that, for $1 \le i, j \le m$,

$$E \sup_{\theta \in \Theta_0} |\frac{\partial^2}{\partial\theta_i \partial\theta_j} \log c(F_{\epsilon,1}(\epsilon_{1,1}), \ldots, F_{\epsilon,r}(\epsilon_{r,1}); \theta)| < \infty.$$

N6. For any $\Delta_6 \in (0, 1/2)$ and $\Delta_7 \in (1/2, 1)$, there exist $\beta_6 \in (0, 1)$, $M_6 > 0$, $\beta_7 > 0$ and $M_7 > 0$ such that

$$\sup_{\theta \in \Theta_0} |\frac{\partial^2}{\partial\theta_i \partial\theta_j} \log c(x_1, \ldots, x_r; \theta)| \le M_6 \{\wedge_{i=1}^r x_i\}^{-\beta_6}$$

for $\wedge_{i=1}^r x_i \le \Delta_6$,

$$\sup_{\theta \in \Theta_0} |\frac{\partial^2}{\partial\theta_i \partial\theta_j} \log c(x_1, \ldots, x_r; \theta)| \le M_6 \{1 - \vee_{i=1}^r x_i\}^{-\beta_6}$$

for $\vee_{i=1}^r x_i \ge \Delta_7$, and

$$\sup_{\theta\in\Theta_0} |\frac{\partial^2}{\partial\theta_i\partial\theta_j}\log c(x_1,\ldots,x_r;\theta) - \frac{\partial^2}{\partial\theta_i\partial\theta_j}\log c(y_1,\ldots,y_r;\theta)| < M_7 \sum_{i=1}^r |x_i - y_i|^{\beta_7}$$

for $\Delta_6/2 < \wedge_{i=1}^r x_i \le \vee_{i=1}^r x_i < \Delta_7 + (1-\Delta_7)/2$ and $\Delta_6/2 < \wedge_{i=1}^r y_i \le \vee_{i=1}^r y_i < \Delta_7 + (1-\Delta_7)/2$.

Note that conditions N4–N5 are standard for proving asymptotic normality of MLE based on i.i.d. data. Conditions N1-N2 are imposed by Csörgő and Révész (1975) for multivariate empirical processes. Conditions N3 and N6 are similar to the ones imposed by Genest, Ghoudi and Rivest (1995) and Chen and Fan (2005) for asymptotic normality based on i.i.d. data; they are employed to control the speed of divergence of partial derivatives of the logarithm of the copula density at the boundaries, and are again satisfied by all the commonly used copula densities.

Theorem 2.3. *Suppose A1–A2, C1–C4 and N1–N6 hold, and that* $\nu(n)/\log n \to \infty, \quad \nu(n)/n \to 0 \quad as \quad n \to \infty$. *Then* $\sqrt{n}(\hat\theta - \theta_0)$ *converges in distribution to* Z *given by*

$$-\Sigma^{-1}(\theta_0)\Big\{\int \delta(T_r^-(x_1,\ldots,x_r);\theta_0)\,dU(x_1,\ldots,x_r) + \sum_{i=1}^r \int \delta_i(x_1,\ldots,x_r;\theta_0)U((1,\ldots,1,x_i,1,\ldots,1)^T)c(x_1,\ldots,x_r;\theta_0)\,dx_1\cdots dx_r.$$

Remark 4. Theorem 2.3 states that under mild sufficient conditions, the limit distribution of $\hat\theta$ is independent of the GARCH filtering. In Chen and Fan (2006a), the normal limiting distribution was obtained by means of heuristic arguments under stringent conditions, and by assuming the existence of the weighted approximation for the empirical distributions of GARCH residuals. Since the variance given in Chen and Fan (2006a) is expressed in terms of a conditional distribution, it is hard to see whether the limit in the above theorem is the same as that in Chen and Fan (2006a). However, both limits are normal and independent of the parameter estimation in the GARCH models. Due to this independence of GARCH models, we can employ the parametric bootstrap method to estimate the variance of $\hat\theta$ and construct confidence intervals for θ_0; see the simulation study below.

2.5. A goodness-of-fit test of residual copulas

The results established in the preceding subsection assume the correct specification of the residual copula by the parametric copula class $\mathcal{C} = \{C(x_1,\ldots,x_r;\theta) :$

$\theta \in \Theta\}$. In this subsection, we propose a consistent test for this assumption. Let

$$H_0 : P\left(C_\epsilon(\epsilon_1, \ldots, \epsilon_r) = C(\epsilon_1, \ldots, \epsilon_r; \theta_0)\right) = 1 \text{ for some } \theta_0 \in \Theta,$$
$$H_1 : P\left(C_\epsilon(\epsilon_1, \ldots, \epsilon_r) = C(\epsilon_1, \ldots, \epsilon_r; \theta)\right) < 1 \text{ for all } \theta \in \Theta.$$

Take the empirical estimator of $C_\epsilon(x_1, \ldots, x_r)$ to be

$$\hat{C}_\epsilon(x_1, \ldots, x_r) = \frac{1}{n}\sum_{t=1}^{n} I(\hat{F}_{\epsilon,1}(\hat{\epsilon}_{1,t}) \le x_1, \ldots, \hat{F}_{\epsilon,r}(\hat{\epsilon}_{r,t}) \le x_r).$$

Then our test statistic is

$$T_n = \int \{\hat{C}_\epsilon(x_1, \ldots, x_r) - C(x_1, \ldots, x_r; \hat{\theta})\}^2 c(x_1, \ldots, x_r; \hat{\theta})\, dx_1 \cdots dx_r,$$

where $\hat{\theta}$ is the pseudo MLE of θ_0 under H_0. Let

$$\dot{C}(x_1, \ldots, x_r; \theta) = \Big(\frac{\partial}{\partial \theta_1} C(x_1, \ldots, x_r; \theta), \ldots, \frac{\partial}{\partial \theta_m} C(x_1, \ldots, x_r; \theta)\Big)^T.$$

Theorem 2.4. *Assume the conditions of Theorem* 2.3 *hold. Further, suppose*

$$\begin{cases} \max\limits_{1\le i\le r} \sup\limits_{0\le x\le 1} F'_{\epsilon,i}(F^{-}_{\epsilon,i}(x))F^{-}_{\epsilon,i}(x) < \infty, \\ \sup\limits_{\theta\in\Theta_0} \sup\limits_{0\le x_1,\ldots,x_r\le 1} |\dot{C}(x_1, \ldots, x_r; \theta)| < \infty, \\ \sup\limits_{0\le x_1,\ldots,x_r\le 1} \sum\limits_{i=1}^{r} \frac{\partial}{\partial x_i} C(x_1, \ldots, x_r; \theta_0) < \infty. \end{cases} \tag{2.7}$$

Then, under H_0,

$$\begin{aligned} nT_n \xrightarrow{d} \int \Big\{ & U(T_r(x_1, \ldots, x_r)) + \sum_{j=1}^{r} \frac{\partial}{\partial x_j} C(x_1, \ldots, x_r; \theta_0) \\ & \times U((1, \ldots, 1, x_j, 1, \ldots, 1)^T) - Z^T \dot{C}(x_1, \ldots, x_r; \theta_0)\Big\}^2 \\ & \times c(x_1, \ldots, x_r; \theta_0)\, dx_1 \cdots dx_r, \end{aligned}$$

where Z *is given in Theorem* 2.3.

Remark 5. As in estimation (see Remark 4), the asymptotic distribution of the test statistic under H_0 is independent of GARCH filtering. This motivates us to employ a parametric bootstrap method to obtain the critical point of the test instead of simulating one from the limiting distribution; see the data examples below.

3. Simulations and Data Examples

We generate 1,000 random samples of size $n = 500$ from model (1.1), with residual copula specified as the mixture

$$C(u_1, \ldots, u_r; \theta_1, \theta_2, \lambda) = \lambda C_1(u_1, \ldots, u_r; \theta_1) + (1 - \lambda) C_2(u_1, \ldots, u_r; \theta_2),$$

where

$$C_1(u_1, \ldots, u_r; \theta_1) = \{\sum_{i=1}^{r} u_i^{-\theta_1} - r + 1\}^{-\frac{1}{\theta_1}}, \theta_1 > 0\,,$$

$$C_2(u_1, \ldots, u_r; \theta_2) = \exp\{-[\sum_{i=1}^{r}(-\log(u_i))^{\theta_2}]^{\frac{1}{\theta_2}}\}, \theta_2 \geq 1.$$

In order to demonstrate that the pseudo MLE is independent of the parameter estimation in the GARCH models, we take $\theta_1 = 3.0$, $\theta_2 = 2.0$, $\lambda = 0.3$ or 0.7, the marginal distributions of ϵ_t as $N(0, 1)$, $r = 3$, and the GARCH model with either $c_1 = c_2 = c_3 = 1$, $\alpha_{1,1} = 0.2$, $\alpha_{2,1} = 0.3$, $\alpha_{3,1} = 0.4$, $\beta_{1,1} = \beta_{2,1} = \beta_{3,1} = 0.2$ (case (i)), or $c_1 = c_2 = c_3 = 0.2$, $\alpha_{1,1} = \alpha_{2,1} = \alpha_{3,1} = 0.2$, $\beta_{1,1} = 0.6$, $\beta_{2,1} = 0.5$, $\beta_{3,1} = 0.4$ (case (ii)). The average and corresponding standard deviation of the proposed pseudo maximum likelihood estimator are reported in Table 1, along with the true values of the model parameters. The proposed method works well. By looking at the two cases with $\lambda = 0.3$ in Table 1, the estimators and their standard deviations for the parameters in GARCH models are different for these two cases, but those for the parameters in the residual copula almost remain the same. Similar observations hold when $\lambda = 0.7$. This is in line with the limiting distribution for the proposed pseudo maximum likelihood estimator $\hat{\theta}$ in Theorem 2.3 being independent of the GARCH filtering, as indicated in Remark 4. We also observe that the standard deviation of $\hat{\theta}_1$ for the case $\lambda = 0.3$ is larger than that for the case $\lambda = 0.7$, and that the standard deviation of $\hat{\theta}_2$ for the case $\lambda = 0.3$ is smaller than that for the case $\lambda = 0.7$. This observation is in line with the role of the parameter λ in the mixture copula.

For constructing confidence intervals, we draw 400 random samples of size $n = 500$ from $C(x_1, x_2, x_3; \hat{\theta}_1, \hat{\theta}_2, \hat{\lambda})$. For each bootstrap sample, we compute the bootstrap version of the MLE, say $\hat{\theta}_1^*, \hat{\theta}_2^*, \hat{\lambda}^*$. Use these 400 bootstrap MLE's to estimate the variances of $\hat{\theta}_1, \hat{\theta}_2, \hat{\lambda}$ so that confidence intervals can be obtained. Based on 1,000 random samples, the coverage probabilities for $\theta_1, \theta_2, \lambda$ with level 0.9 are 0.903, 0.928, 0.899 for case (i) with $\lambda = 0.3$; 0.900, 0.926, 0.893 for case (i) with $\lambda = 0.7$; 0.908, 0.909, 0.890 for case (ii) with $\lambda = 0.3$; 0.898, 0.920, 0.903 for case (ii) with $\lambda = 0.7$. These numbers show that the proposed parametric bootstrap method works well and further confirms the property of independence of GARCH filtering given in the Theorem 2.3.

Table 1. Estimation results for mixture copula $C(u_1, u_2, u_3; \theta_1, \theta_2, \lambda)$. Standard deviations are given in parenthesis.

	Case (i) $\lambda = 0.3$	Case (i) $\lambda = 0.7$	Case (ii) $\lambda = 0.3$	Case (ii) $\lambda = 0.7$
c_1	1.033 (0.292)	1.111 (0.306)	0.241 (0.117)	0.260 (0.122)
$\alpha_{1,1}$	0.206 (0.073)	0.215 (0.071)	0.209 (0.065)	0.219 (0.062)
$\beta_{1,1}$	0.207 (0.181)	0.197 (0.177)	0.566 (0.149)	0.563 (0.137)
c_2	1.062 (0.269)	1.129 (0.278)	0.237 (0.109)	0.254 (0.110)
$\alpha_{2,1}$	0.309 (0.083)	0.327 (0.081)	0.208 (0.068)	0.217 (0.067)
$\beta_{2,1}$	0.195 (0.146)	0.188 (0.135)	0.457 (0.173)	0.456 (0.174)
c_3	1.071 (0.245)	1.124 (0.247)	0.219 (0.086)	0.236 (0.091)
$\alpha_{3,1}$	0.408 (0.089)	0.437 (0.085)	0.206 (0.070)	0.216 (0.067)
$\beta_{3,1}$	0.192 (0.114)	0.192 (0.108)	0.381 (0.191)	0.375 (0.185)
θ_1	2.956 (0.504)	2.959 (0.242)	2.984 (0.537)	2.981 (0.250)
θ_2	2.064 (0.091)	2.166 (0.163)	2.061 (0.102)	2.171 (0.195)
λ	0.298 (0.064)	0.693 (0.061)	0.295 (0.065)	0.690 (0.060)

Next we apply the proposed estimate and test to two data sets. The first one consists of 2,275 daily log-returns of the S&P 500 index, Cisco System and Intel Corporation, from January 2, 1991 to December 31, 1999; see Figure 1 and Fan, Wang and Yao (2008). The second data set contains 2,635 daily log-returns of stock prices of Nortel, Lucent and Cisco, from April 4, 1996 to September 22, 2006; see Figure 2. Here, we fit the mixture copula $C(x_1, x_2, x_3; \theta_1, \theta_2, \lambda)$ to the residuals from filtering a GARCH(1,1) for each series; see Tables 2−5 for parameter estimates. As mentioned in Remark 3, we employ the parametric bootstrap method to obtain p-values of the test.

We draw 200 random samples with size 2,275 for the first data set and size 2,635 for the second data set from $C(x_1, x_2, x_3; \hat{\theta}_1, \hat{\theta}_2, \hat{\lambda})$. Based on each sample, we compute the bootstrap version of T_n, say T_n^*. Hence we have $T_n^*(1), \ldots,$ $T_n^*(200)$, and the p-value is calculated as $(1/200) \sum_{i=1}^{200} I(T_n^*(i) \geq T_n)$, see Tables

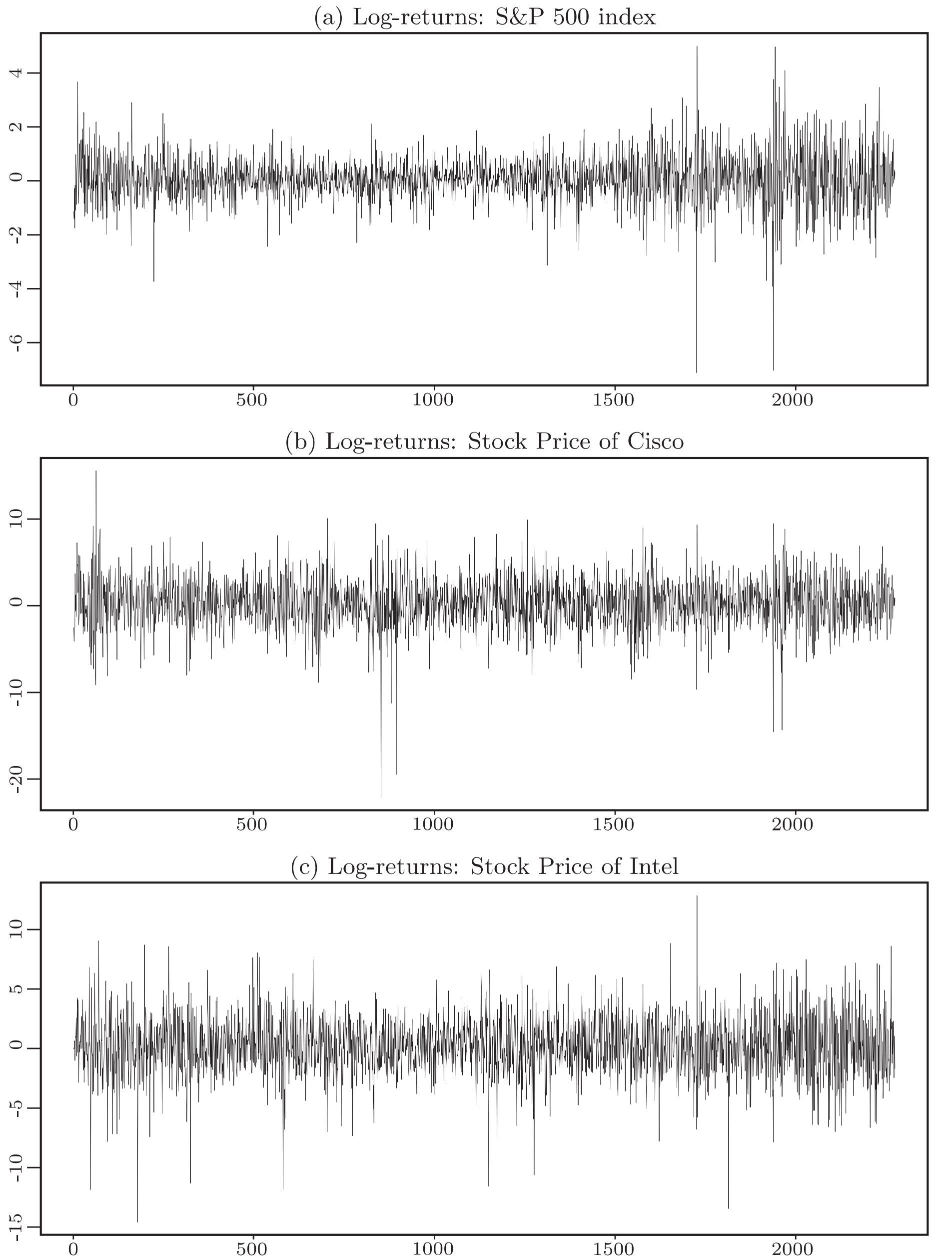

Figure 1. Daily Log-returns of S&P 500 index (a), Stock Price of Cisco Systems (b) and Stock Price of Intel Corporation (c) from January 2, 1991 to December 31, 1999.

3 and 5. The p-values in Tables 3 and 5 clearly reject the mixture copula for both data sets. Since this mixture copula is mainly designed to catch both tails

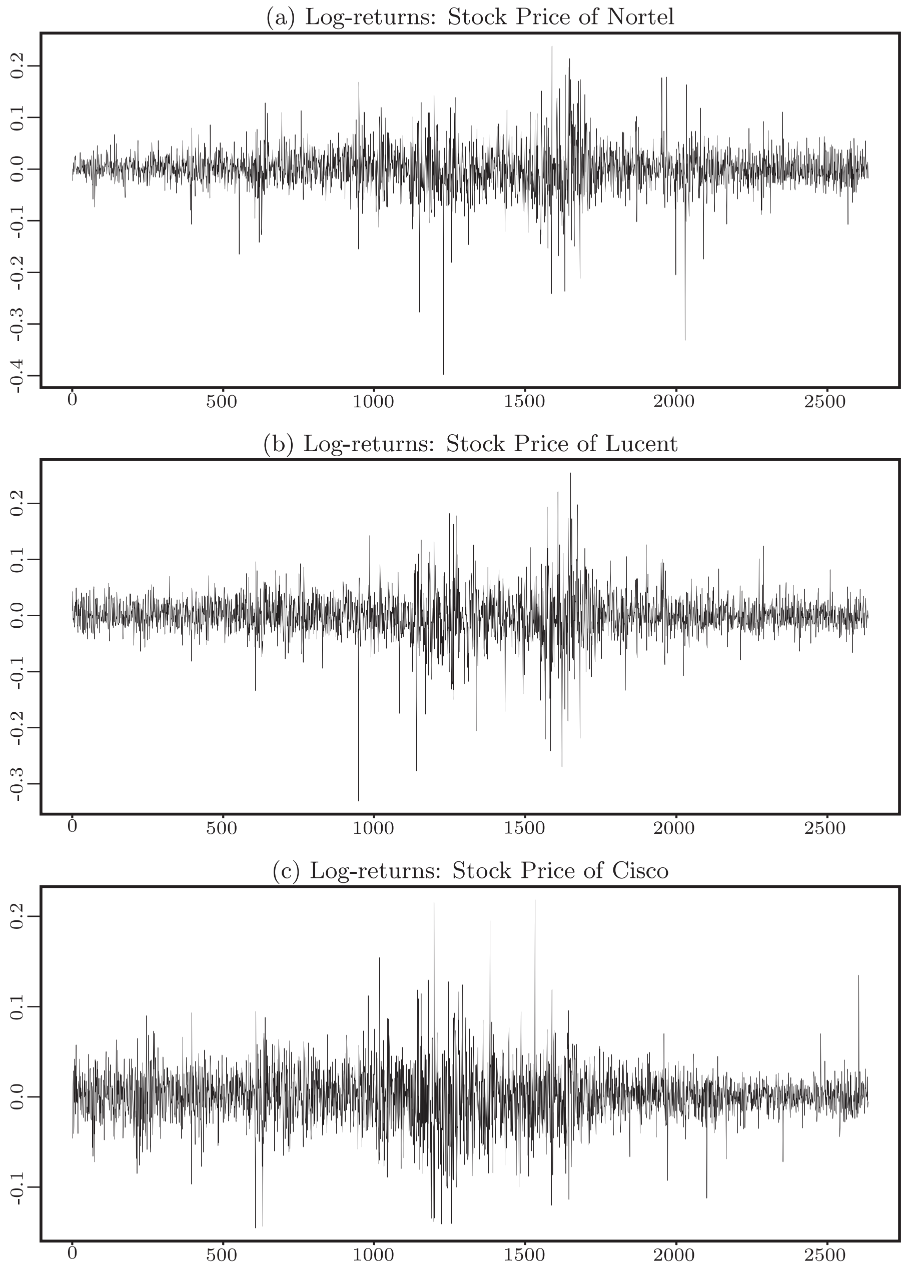

Figure 2. Daily Log-returns of stock prices of Nortel (a), Lucent (b) and Cisco (c) from April 4, 1996 to September 22, 2006.

of a data set, it is quite challenging to fit a parametric residual copula to catch both the tail and middle parts of a data set; seeking more flexible parametric models is of importance.

Table 2. Parameter estimates for GARCH(1,1) of the daily log-returns of S&P 500 index, stock prices of Cisco and Intel.

	S&P 500 index	Cisco systems	Intel Corporation
c	0.0096	0.3272	0.1336
α	0.0636	0.0737	0.0186
β	0.9247	0.8879	0.9597

Table 3. Copula parameter estimates and test statistic of the daily log-returns of S&P 500 index, stock prices of Cisco and Intel.

	$C_2(;\theta_1,\theta_2,\lambda)$
Parameter estimation	(1.0549, 1.4618, 0.4716)
Test statistic nT_n	0.1806
P-value	0.000

Table 4. Parameter estimates for GARCH(1,1) of the daily log-returns of stock prices of Nortel, Lucent and Cisco.

	Nortel	Lucent	Cisco
c	7.0×10^{-6}	1.0×10^{-5}	8.0×10^{-6}
α	0.0360	0.0436	0.0627
β	0.9609	0.9504	0.9301

Table 5. Copula parameter estimates and test statistic of the daily log-returns of stock prices of Nortel, Lucent and Cisco.

	$C_2(;\theta_1,\theta_2,\lambda)$
Parameter estimation	(1.05839, 1.3456, 0.4284)
Test statistic nT_n	0.2301
P-value	0.000

Acknowledgements

We thank the Co-Editors, an associate editor, and a reviewer for helpful comments. Chan and Peng's research was supported in part by HKSAR-RGC Earmarked Grants 400305, 400306 and NSF Grant SES-0631608. Chen and Fan's research was supported in part by NSF Grants SES-0318091 and SES-0613613, and SES-0317984 respectively.

References

Amemiya, T. (1985). *Advanced Econometrics.* Harvard University Press, Boston.

Bauwens, J., Laurent, S. and Rombouts, J. V. K. (2006). Multivariate GARCH models: a survey. *J. Appl. Econometrics* **21**, 79-110.

Berkes, I. and Horváth, L. (2003). Limit results for the empirical process of squared residuals in GARCH models. *Stochastic Process. Appl.* **105**, 271-298.

Breymann, W., Dias, A. and Embrechts, P. (2003). Dependence structures for multivariate high-frequency data in finance. *Quantitative Finance* **3**, 1-16.

Chen, S. and Huang, T. (2007). Nonparametric estimation of copula functions for dependence modeling. *Canad. J. Statist.* **35**, 265-282.

Chen, X. and Fan, Y. (2005). Pseudo-likelihood ratio tests for model selection in semiparametric multivariate copula models. *Canad. J. Statist.* **33**, 389-414.

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

Chen, X. and Fan, Y. (2006b). Semiparametric estimation of copula-based time series models. *J. Econometrics* **130**, 307-335.

Chen, X., Fan, Y. and Tsyrennikov, V. (2006). Efficient estimation of semiparametric multivariate copula models. *J. Amer. Statist. Assoc.* **101**, 1228-1240.

Claeskens, G. and Van Keilegom, I. (2003). Bootstrap confidence bands for regression curves and their derivatives. *Ann. Statist.* **31**, 1852-1884.

Csörgő, M. and Révész, P. (1975). A strong approximation of the multivariate empirical process. *Studia Scientiarum Mathematicarum Hungarica* **10**, 427-434.

Fan, J., Wang, M. and Yao, Q. (2008). Modeling multivariate volatilities via conditionally uncorrelated components. *J. Roy. Statist. Soc. Ser. B* **70**, 679-702.

Fermanian, J. D. (2005). Goodness-of-fit tests for copulas. *J. Multivariate Anal.* **95**, 119-152.

Fermanian, J. D., Radulovic, D. and Wegkamp, M. (2004). Weak convergence of empirical copula processes. *Bernoulli* **10**, 847-860.

Fermanian, J. D. and Scaillet, O. (2003). Nonparametric estimation of copulas for time series. *J. Risk* **5**, 25-54.

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

Genest, C., Quessy, J.-F. and Rmillard, B. (2006). Goodness-of-fit procedures for copula models based on the probability integral transformation. *Scand. J. Statist.* **33**, 337-366.

Giacomini, E., Härdle, W., Ignatieva, E. and Spokoiny, V. (2008). Inhomogeneous dependency modelling with time varying copulae. *J. Business and Economic Statist.* **26**, in press.

Hall, P. and Yao, Q. (2003). Inference in ARCH and GARCH models. *Econometrica* **71**, 285-317.

Hull, J. and White, A. (1998). Value at risk when daily changes in market variables are not normally distributed. *J. Derivatives* **5**, 9-19.

Horváth, L., Kokoszka, P. and Teyssiére, G. (2001). Empirical process of the squared residuals of an ARCH sequence. *Ann. Statist.* **29**, 445-469.

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

Klugman, S. A. and Parsa, R. (1999). Fitting bivariate loss distributions with copulas. *Insurance: Mathematics and Economics* **24**, 139-148.

Koul, H. L. and Ling, S. (2006). Fitting an error distribution in some heteroscedastic time series models. *Ann. Statist.* **34**, 994-1012.

McNeil, A. J., Frey, R. and Embrechts, P. (2005). *Quantitative Risk Management: Concepts, Techniques and Tools.* Princeton University Press, New Jersey.

Nelsen, R. B. (2005). *An Introduction to Copulas.* 2nd edition. Springer, New York.

Scaillet, O. (2007). Kernel based goodness-of-fit tests for copulas with fixed smoothing parameters. *J. Multivariate Anal.* **98**, 533-543.

Vervaat, W. (1972). Functional central limit theorems for processes with positive drift and their inverses. *Z. Wahrsch. Verw. Gebiete* **23**, 249-253.

Department of Statistics, The Chinese University of Hong Kong, Shatin, N.T., Hong Kong.

E-mail: nhchan@sta.cuhk.edu.hk

Sage SB Inc., 1715 N. Brown Road, Lawrensville, GA 30043, USA.

E-mail: chen_jian@sage.com

Department of Economics, Yale University, Box 208281, New Haven, CT 06520-8281, USA and Shanghai University of Finance and Economics, Shanghai, PRC.

E-mail: xiaohong.chen@yale.edu

Department of Economics, Vanderbilt University, VU Station B #351819, 2301 Vanderbilt Place, Nashville, TN 37235, USA.

E-mail: yanqin.fan@vanderbilt.edu

School of Mathematics, Georgia Institute of Technology, Atlanta, GA 30332-0160, USA.

E-mail: peng@math.gatech.edu

(Received October 2006; accepted June 2007)