

Semiparametric Estimation and Model Selection for Conditional Mixture Copula Models

Guannan Liu¹, Wei Long², Bingduo Yang³, and Zongwu Cai⁴

¹School of Economics and WISE, Xiamen University

²Department of Economics, Tulane University

³Lingnan (University) College, Sun Yat-Sen University

⁴Department of Economics, University of Kansas

January 13, 2021

Abstract

Conditional copula models allow the dependence structure among variables to vary with covariates, and thus can describe the evolution of the dependence structure with those factors. This paper proposes a conditional mixture copula which is a weighted average of several individual conditional copulas. We allow both the weights and copula parameters to vary with a covariate so that the conditional mixture copula offers additional flexibility and accuracy in describing the dependence structure. We propose a two-step semiparametric estimation method and develop asymptotic properties of the estimators. Moreover, we introduce model selection procedures to select the component copulas of the conditional mixture copula model. Simulation results suggest that the proposed procedures have a good performance in estimating and selecting conditional mixture copulas with different model specifications. The proposed model is then applied to investigate how the dependence structures among international equity markets evolve with the volatility in the exchange rate markets.

Keywords: conditional copula; mixture copula; model selection; semiparametric estimation

1 Introduction

The Sklar's theorem by Sklar (1959) enables one to decompose a multivariate joint density into a product of univariate marginal densities and a copula density, so that the latter contains all information about the dependence structure. A copula model has several desirable properties when applied to study dependence. For example, a copula model can catch various types of dependence structures such as linear or nonlinear, symmetric or asymmetric, tail or non-tail dependence. Moreover, unlike the conventional linear correlation, it is invariant to strictly monotonic transformations. Patton (2012) and Fan and Patton (2014) provide excellent summaries of the development in copula.

In many cases, researchers may need a conditional copula to better describe the effect of a covariate on the degree of dependence. In a conditional copula model, the degree of dependence, measured by the copula parameter, is no longer a constant but a function of a covariate. Therefore, compared with copulas with constant parameters, a conditional copula provides another channel to investigate the dependence structure among variables. In the literature, Patton (2006) pioneers the conditional copula model by extending the Sklar's theorem for conditional distributions and sets the copula parameter to be a parametric function of lagged terms. After that, there is a sequence of studies concentrating on the dynamic in copula parameter (e.g., Giacomini *et al.*, 2009, Garcia & Tsafack, 2011, Acar *et al.*, 2011, Hafner & Manner, 2012, Abegaz *et al.*, 2012, and Fermanian & Lopez, 2018).

Another line of extension is to propose a mixture copula that is a linear combination of several individual copulas. The key idea is that, by combining individual copulas with different dependence patterns, a mixture copula can capture dependence structures which do not belong to any individual copula, and thus exhibits greater flexibility to describe dependence structures. Therefore, a mixture copula is more flexible than an individual copula and can be used to specify various dependence structures in data (e.g., Chollete *et al.*, 2005, Hu, 2006, Cai & Wang, 2014, and Liu *et al.*, 2019). However, even though a mixture copula exhibits great flexibility in describing more general dependence structures, the parameters in this model – weights and copula parameters – are usually assumed to be

constants, so researchers still face difficulty in describing how the dependence evolves with certain covariates.

In this paper, we contribute to the literature by proposing an innovative semiparametric conditional mixture copula model which allows both the weights and copula parameters in the mixture copula model to vary with a covariate in a nonparametric way. The superiority of a conditional mixture copula is that it carries the advantages of both the conditional copula and mixture copula discussed above, so empirical practitioners can flexibly describe the dependence structure and effectively mitigate the potential model misspecification problem simultaneously. For example, compared with an individual conditional copula, a conditional mixture copula offers additional flexibility and accuracy by accommodating more copula families whose weights and parameters are both varying with a covariate. To estimate the unknown parameters, we maximize a local log-likelihood function by applying the local polynomial framework (Fan & Gijbels, 1996). We then establish the large sample properties of the nonparametric estimators under some regularity conditions.

When investigating the choice of an appropriate conditional mixture copula model, we suggest two copula model selection methods. The first one follows Huang *et al.* (2013) who propose an information criterion approach to select the components in the nonparametric mixture of regression models. Specifically, for each candidate mixture model consisting of different combinations of individual copulas, we calculate its maximum log-likelihood and then construct an information criterion such as BIC. To implement such an information criterion, we will need an extra step to consider the model complexity, which is measured by the degree of freedom derived by Fan *et al.* (2001). Then, the model with the lowest information criterion value will be selected from the candidate models. The second model selection strategy is in spirit similar to the backward elimination or forward addition procedures in the context of linear regression settings. It involves a sequence of generalized likelihood ratio tests through which component copulas with insignificant weights at the conventional significance levels are filtered out. Using either method, we achieve the goal of selecting an appropriate conditional mixture copula from all candidates to best describe the

dependence structure, and then estimate the unknown parameters in the selected model.

Our simulation results show that the proposed estimation method and model selection procedures exhibit a good performance when the true model is either an individual copula or a mixture copula. On one hand, the estimation errors of copula parameters and weights associated with each component copula decrease remarkably when the sample size increases. On the other hand, the true component copulas are highly likely to be selected even when the sample size is small, and the probability of inaccurate selection declines as the sample size increases.

In an empirical illustration, we apply the proposed estimation and model selection procedures to investigate the dependence structures and comovement patterns among the equity returns in France, Germany, the United States and the United Kingdom along the volatility in the exchange rate markets of the four countries. The empirical results show that, of the Clayton, Gumbel and Frank copulas, the Clayton and Frank copulas are always selected and the weight of the Clayton copula increases when the exchange rates become extremely volatile, indicating a more salient lower tail dependence among the equity markets in the four economies. When examining the magnitude of the dependence measured by Kendall's τ , we find that the lower tail dependence becomes strengthened as volatility in the exchange rate markets increases. Both findings are in line with Garcia & Tsafack (2011): when a sudden shock hits an economy with an active currency market, transmission through the exchange rate market leads to a downside comovement of equity markets more likely than in a tranquil period of the exchange rate market.

The rest of the paper is organized as follows. In Section 2, we propose the estimation method, the asymptotic theory, and the model selection procedures for conditional mixture copula models. We conduct Monte Carlo simulations and discuss the results in Section 3. To highlight the practical usefulness of the proposed methods, in Section 4 we provide an empirical illustration on how the dependence structures among international equity markets evolve with the volatility in exchange rate markets. Section 5 draws the conclusion. In the online appendices, Appendix A provides a stationary bootstrap technique, Appendix B

discusses some practical issues including an EM algorithm, the selection of the bandwidth, and the confidence intervals, Appendix C documents the proofs of the key results, and Appendix D presents additional simulation results.

2 Model and Estimation

In this section, we present a semiparametric conditional mixture copula model and the corresponding estimation and selection procedures.

2.1 A Semiparametric Conditional Mixture Copula Model

Let $\{\mathbf{X}_t\}_{t=1}^T$ be a series of p -dimensional vectors with $\mathbf{X}_t = (X_{1t}, \dots, X_{pt})^\top$ and p being a finite positive integer, and let $\{Z_t\}_{t=1}^T$ be a 1-dimensional vector of the covariate. Denote $F(\mathbf{x}_t|z_t)$ and $f(\mathbf{x}_t|z_t)$ as the joint distribution and the density function of \mathbf{X}_t evaluated at $\mathbf{x}_t \in \mathcal{R}^p$ and conditional on $Z_t = z_t$, $F_s(x_{st}|z_t)$ and $f_s(x_{st}|z_t)$ as the marginal distribution and the density function of X_{st} evaluated at $x_{st} \in \mathcal{R}$ and conditional on $Z_t = z_t$, respectively, where $s = 1, \dots, p$. Our target is to estimate the conditional joint distribution $F(\mathbf{x}_t|z_t)$ based on a conditional mixture copula model. Theoretically, the conditional mixture copula model can be formulated as a linear combination of infinite individual copulas:

$$C\{u(z_t); \omega(z_t), \theta(z_t)\} = \sum_{k=1}^{\infty} \omega_k(z_t) C_k\{u(z_t); \theta_k(z_t)\},$$

where $\{C_k(\cdot; \cdot)\}_{k=1}^{\infty}$ is a set of candidate copulas with unknown parameters $\{\theta_k\}$ and a p -dimensional conditional marginal distribution $u(z_t) = (F_1(x_{1t}|z_t), \dots, F_p(x_{pt}|z_t))$. $\{C_k(\cdot; \cdot)\}_{k=1}^{\infty}$ can be regarded as known basis copula functions so that $C\{u(z_t); \omega(z_t), \theta(z_t)\}$ can be regarded as a series expansion based on the basis copula functions $\{C_k(\cdot; \cdot)\}_{k=1}^{\infty}$. In real applications, we use finite number of d individual copulas to approximate the true model:

$$C\{u(z_t); \omega(z_t), \theta(z_t)\} = \sum_{k=1}^d \omega_k(z_t) C_k\{u(z_t); \theta_k(z_t)\}, \quad (1)$$

where $\omega(z_t) = (\omega_1(z_t), \dots, \omega_d(z_t))^\top$, $\theta(z_t) = (\theta_1(z_t), \dots, \theta_d(z_t))^\top$, and $\{C_1(\cdot; \cdot), \dots, C_d(\cdot; \cdot)\}$ is a set of candidate copulas. Let $\{\omega_k\}_{k=1}^d$ denote the weight parameters satisfying $0 \leq \omega_k \leq 1$ and $\sum_{k=1}^d \omega_k = 1$, and d is the number of candidate copulas. The copula parameters $\{\theta_k(z_t)\}$ and the weight parameters $\{\omega_k(z_t)\}$ are set to be unknown functions of the covariate.

When using (1) to approximate the true model, we may encounter a misspecification problem because some true individual copulas might not be included. To avoid this problem, we can first consider a large set of candidate copulas and then employ a copula model selection procedure discussed in Section 2.4 to filter out the “insignificant” component copulas. Furthermore, even if some true individual copulas are excluded so that the model becomes misspecified, we can still estimate and select the closest mixture copula model by the model selection procedure described in Section 2.4. Therefore, the model in (1) is flexible enough to capture a true copula in real applications.

For model identification, two conditional mixture copulas $C\{u(z_t); \omega(z_t), \theta(z_t)\} = \sum_{k=1}^d \omega_k(z_t) C_k\{u(z_t); \theta_k(z_t)\}$ and $C^*\{u(z_t); \omega^*(z_t), \theta^*(z_t)\} = \sum_{k=1}^{d^*} \omega_k^*(z_t) C_k^*\{u(z_t); \theta_k^*(z_t)\}$ are said to be identified, i.e., $C\{u(z_t); \omega(z_t), \theta(z_t)\} \equiv C^*\{u(z_t); \omega^*(z_t), \theta^*(z_t)\}$, if and only if $d = d^*$ and we can order the summations such that $\omega(z_t) = \omega^*(z_t)$ and $C_k\{u(z_t); \theta_k(z_t)\} = C_k^*\{u(z_t); \theta_k^*(z_t)\}$ for all possible values of z_t , $u(z_t)$ with $k = 1, \dots, d$. Without loss of generality, we follow Cai & Wang (2014) and assume that the conditional mixture copula model under investigation is identified.

Our model setting here has three superiorities. First, instead of imposing assumptions on the functional forms of the unknown weights and copula parameters, we conduct a data-driven method (which will be specified below) to estimate them. Second, compared with an individual conditional copula, our conditional mixture copula model allows not only the copula parameters but also the weight parameters of the component copulas to vary with the covariate. Third, when constructing the model in equation (1), we impose no restrictions on the number of candidate copulas included in the model so that a large copula candidate set can be taken to avoid the copula misspecification problem. We will filter out the “insignificant” component copulas through the copula selection procedures discussed

later in Section 2.4.

2.2 Estimation Procedures

We propose to estimate model (1) in two steps. First, we estimate the unknown marginal distributions in the model by a rescaled empirical distribution function method. Second, after replacing the unknown marginal distributions with the estimates obtained from the first step, we adopt the local polynomial approximation (see Fan and Gijbels, 1996) in a local log-likelihood setting to estimate the weights and copula parameters in model (1). Each step is described specifically as follows:

Step One: We follow Chen & Fan (2006a) and use a rescaled empirical distribution function to estimate the marginal distributions, i.e.

$$\widehat{F}_s(x_{st}) = \frac{1}{T+1} \sum_{t=1}^T I\{X_{st} \leq x_{st}\} \quad \text{for } s = 1, \dots, p.$$

Remark 1. *Ideally, one should use conditional estimators to estimate the marginal distributions, i.e., $\widehat{F}_s(x_{st}|z_t) := \sum_{t=1}^T I\{X_{st} \leq x_{st}\} K_h(Z_t - z_t) / \sum_{t=1}^T K_h(Z_t - z_t)$ (see Abegaz et al., 2012). However, due to the fact that estimators for both the marginals and copula parameters have the same convergence rate of \sqrt{Th} , this setting would make the asymptotic properties of copula parameters more complicated, especially with time series data. Because the main focus of this paper is on the estimation of weights and copula parameters of a conditional mixture copula model and selection of component copula families, we assume that the marginal distributions do not depend on the covariate, which is similar to Acar et al. (2011). Relaxing this assumption for marginal distributions would be an interesting topic for future research.*

Step Two: Given the estimators of the marginals $\widehat{u}_t = \left(\widehat{F}_1(x_{1t}), \dots, \widehat{F}_p(x_{pt}) \right)^\top = (\widehat{u}_{1t}, \dots, \widehat{u}_{pt})^\top$, we next estimate the unknown weight and copula parameters locally by a polynomial. The copula parameter space is restricted for many widely used copula families. For example, for a Gaussian copula, $\theta \in (-1, 1)$, and for a Gumbel Copula, $\theta \in [1, \infty)$.

Moreover, the weight parameters are restricted to take values between 0 and 1. In contrast, the polynomial framework assumes that any points belong to \mathcal{R} can be taken. For this reason, we follow Acar *et al.* (2011) and Abegaz *et al.* (2012) and use some known inverse transformation functions to ensure that the weight and copula parameter space is correct. Specifically, we denote $g_{\omega,k}^{-1} : \mathcal{R} \rightarrow \Omega_k$ and $g_{\theta,k}^{-1} : \mathcal{R} \rightarrow \Theta_k$ as the inverse link functions respectively for the weight and copula parameters of the k th component copula. Therefore, we have $\omega_k(z) = g_{\omega,k}^{-1}(w_k(z)) \in \Omega_k$ and $\theta_k(z) = g_{\theta,k}^{-1}(\vartheta_k(z)) \in \Theta_k$ for $k = 1, \dots, d$. The choice of the link functions is not important for theory development as long as they are monotone. For example, we can choose the inverse link functions $g^{-1}(z) = \exp(z)$ for the Clayton copula, $g^{-1}(z) = z$ for the Frank copula, and $g^{-1}(z) = \exp(z) + 1$ for the Gumbel copula, so that the resulting copula parameter estimates are guaranteed to be in the correct range.

Then, for $k = 1, \dots, d$, assuming that w_k and ϑ_k have the $(q + 1)$ th derivative at point z , we can approximate $w_k(z_t)$ and $\vartheta_k(z_t)$ for data points z_t in the neighborhood of z by the following Taylor expansions:

$$\begin{aligned} w_k(z_t) &\approx w_k(z) + w_k^{(1)}(z)(z_t - z) + \dots + w_k^{(q)}(z)(z_t - z)^q/q! \\ &\equiv \alpha_{k0} + \alpha_{k1}(z_t - z) + \dots + \alpha_{kq}(z_t - z)^q, \end{aligned} \quad (2)$$

$$\begin{aligned} \vartheta_k(z_t) &\approx \vartheta_k(z) + \vartheta_k^{(1)}(z)(z_t - z) + \dots + \vartheta_k^{(q)}(z)(z_t - z)^q/q! \\ &\equiv \beta_{k0} + \beta_{k1}(z_t - z) + \dots + \beta_{kq}(z_t - z)^q, \end{aligned} \quad (3)$$

where $\alpha_{kr} = \alpha_{kr}(z) = w_k^{(r)}(z)/r!$ and $\beta_{kr} = \beta_{kr}(z) = \vartheta_k^{(r)}(z)/r!$ for each $r \in \{0, \dots, q\}$. Then, the local log-likelihood function can be approximated as follows:

$$\frac{1}{T} \sum_{t=1}^T \ln \left(\sum_{k=1}^d g_{\omega,k}^{-1} \{ \alpha_{k0} + \dots + \alpha_{kq}(z_t - z)^q \} c_k [\hat{u}_t; g_{\theta,k}^{-1} \{ \beta_{k0} + \dots + \beta_{kq}(z_t - z)^q \}] \right) \times K_h(z_t - z),$$

where $c_k(\cdot)$ is the copula density function of the k th component copula in model (1), and $K(\cdot)$ is a kernel function with $K_h(\cdot) = K(\cdot/h)/h$ and h being the bandwidth.

For the choice of q , we take $q = 1$ throughout this paper. That is, we apply the commonly

used local linear fitting (see Fan & Gijbels, 1996) in the paper. The local log-likelihood function then reduces to

$$L(\hat{u}, \delta) = \frac{1}{T} \sum_{t=1}^T \ln \left(\sum_{k=1}^d g_{\omega,k}^{-1} \{ \alpha_{k0} + \alpha_{k1}(z_t - z) \} c_k [\hat{u}_t; g_{\theta,k}^{-1} \{ \beta_{k0} + \beta_{k1}(z_t - z) \}] \right) \times K_h(z_t - z), \quad (4)$$

where $\hat{u} = (\hat{u}_1^\top, \dots, \hat{u}_T^\top)^\top$ and $\delta = (\alpha_{10}, \dots, \alpha_{d0}, \beta_{10}, \dots, \beta_{d0}, \alpha_{11}, \dots, \alpha_{d1}, \beta_{11}, \dots, \beta_{d1})^\top$. Note that a maximum likelihood estimator may not have a closed form, so an iterative algorithm should be adopted to find the numerical solution (see Appendix B for details). Then we can obtain the estimators of $w_k^{(r)}(z)$ and $\vartheta_k^{(r)}(z)$ by defining $\hat{w}_k^{(r)}(z) = r! \hat{\alpha}_{kr}$ and $\hat{\vartheta}_k^{(r)}(z) = r! \hat{\beta}_{kr}$ with $r \in \{0, 1\}$. Finally, for any covariate z , the weight and copula parameters in model (1) can be respectively estimated by

$$\begin{aligned} \hat{\omega}_k(z) &= g_{\omega,k}^{-1}(\hat{w}_k(z)) = g_{\omega,k}^{-1}(\hat{\alpha}_{k0}), \quad \text{and} \\ \hat{\theta}_k(z) &= g_{\theta,k}^{-1}(\hat{\vartheta}_k(z)) = g_{\theta,k}^{-1}(\hat{\beta}_{k0}), \end{aligned}$$

for $k = 1, \dots, d$.

2.3 Large Sample Theory

To find the large sample properties of the nonparametric estimators, first, we rewrite the kernel-based local log-likelihood function as

$$L(u, \delta) = \frac{1}{T} \sum_{t=1}^T \ell \left(u_t, g^{-1}(\tilde{Z}_t^\top \delta) \right) K_h(z_t - z),$$

where $\ell(u_t, g^{-1}(\tilde{Z}_t^\top \delta)) = \ln \left(\sum_{k=1}^d g_{\omega,k}^{-1} \{ \alpha_{k0} + \alpha_{k1}(z_t - z) \} c_k [u_t; g_{\theta,k}^{-1} \{ \beta_{k0} + \beta_{k1}(z_t - z) \}] \right)$, $g^{-1}(\cdot) = (g_{\omega,1}^{-1}(\cdot), \dots, g_{\omega,d}^{-1}(\cdot), g_{\theta,1}^{-1}(\cdot), \dots, g_{\theta,d}^{-1}(\cdot))^\top$ is a vector of link functions, and $\tilde{Z}_t = (I, (z_t - z)I)^\top$ with I being a $2d \times 2d$ identity matrix. Next, we define a vector of coefficients $\xi = (\omega^\top, \theta^\top)^\top = (g_{\omega,1}^{-1}(w_1), \dots, g_{\omega,d}^{-1}(w_d), g_{\theta,1}^{-1}(\vartheta_1), \dots, g_{\theta,d}^{-1}(\vartheta_d))^\top$ and its corresponding estima-

tor $\widehat{\xi}$. Similarly, we define $\eta = (w^\top, \vartheta^\top)^\top = (w_1, \dots, w_d, \vartheta_1, \dots, \vartheta_d)^\top$ and its corresponding estimator $\widehat{\eta}$. In addition, let $f(z)$ be the density function of z and ϵ be a small positive constant. Define the domain of z as $\Phi_z = \{z : f(z) \geq \epsilon\}$; there exists a and b such that $z \in [a, b]$, i.e., Φ_z is the set of bounded z whose density is bounded away from 0.

Meanwhile, we introduce some regularity conditions as below:

- C1. The vector of functions η is continuous, bounded and has third order continuous derivatives on Φ_z ;
- C2. There exists two constants a and b such that for any $z \in [a, b]$, the density function $f(z)$ is continuous and $f(z) > \epsilon$ for a small positive constant $\epsilon > 0$;
- C3. The copula log-likelihood function $\ell(u_t, \xi)$ has bounded third derivative with respect to ξ and bounded second derivative with respect to u_t . Further, $\partial\ell(u_t, \xi)/\partial\xi$ and $[g^{-1}]'$ are Lipschitz continuous;
- C4. $0 \leq \omega_k(z) \leq 1$ and $\sum_{k=1}^d \omega_k(z) = 1$ for all $z \in \Phi_z$;
- C5. The kernel function $K(z)$ is twice continuously differentiable on the support $(-1, 1)$, and its second order derivative satisfies a Lipschitz condition. Let $v_0 = \int K^2(z)dz$, $v_2 = \int z^2 K^2(z)dz$ and $\mu_2 = \int z^2 K(z)dz$;
- C6. The bandwidth h satisfies that $h \rightarrow 0$ and $Th \rightarrow \infty$, as $T \rightarrow \infty$;
- C7. Assume that $\{\mathbf{X}_t, Z_t\}_{t=1}^T$ is a strictly stationary α -mixing sequence. Furthermore, assume that there exists some constant $c > 0$ such that $E\|\mathbf{X}_t\|^{2(2+c)} < \infty$ where $\|\cdot\|$ represents the Euclidean norm (L_2 -norm), $E|Z_t|^{2(2+c)} < \infty$, and the mixing coefficient $\alpha(m)$ satisfies $\alpha(m) = O(m^{-c_0})$ with $c_0 = (2+c)(1+c)/c$.

Remark 2. *Conditions in C1 - C3 are for the derivation of the asymptotic properties. Conditions in C4 are mild conditions for identification and conditions in C5 and C6 are commonly employed in nonparametric estimation. Conditions in C7 are the common conditions with weakly dependent data. Most financial models such as ARMA, ARCH and GARCH models satisfy these conditions; see Cai (2002).*

Theorem 1: Let $\{\mathbf{X}_t, Z_t\}_{t=1}^T$ be a strictly stationary and strong mixing sequence. Assume that $\sup_{1 \leq t \leq T} |\widehat{u}_{st} - u_{st}| = O_p(1/\sqrt{T})$ for $s = 1, \dots, p$, $h \rightarrow 0$ and $Th \rightarrow \infty$ as $T \rightarrow \infty$. For a fixed point $z \in \Phi_z$, under conditions C1 - C7, we have

$$D_T \left(\widehat{\delta} - \delta - h^2 B(z) \right) \xrightarrow{d} N \left(0, \begin{pmatrix} \frac{\nu_0}{f(z)} \\ \frac{\nu_2}{\mu_2^2 f(z)} \end{pmatrix} \otimes \{ \Psi(z) \circ \{ [(g^{-1})'(\eta(z))] [(g^{-1})'(\eta(z))]^\top \}^{-1} \} \right),$$

where $D_T = \text{diag}(\sqrt{Th}I, \sqrt{Th}hI)$ with I being an $2d \times 2d$ identity matrix, $B(z) = (\frac{1}{2}\eta''(z)^\top \mu_2, 0^\top)^\top$ is the bias term, \otimes is the Kronecker product, and \circ is the Hadamard product. $\Psi(z) = \Sigma^{-1}(z)\Omega(z)\Sigma^{-1}(z)$ with $\Sigma(z) = -E\{\ell''(u_t, g^{-1}(\eta(z_t))) | z_t = z\}$ and $\Omega(z) = E\{\ell'(u_t, g^{-1}(\eta(z_t)))\ell'(u_t, g^{-1}(\eta(z_t)))^\top | z_t = z\}$.

Remark 3. The condition $\sup_{1 \leq t \leq T} |\widehat{u}_{st} - u_{st}| = O_p(1/\sqrt{T})$ for $s = 1, \dots, p$ can be obtained from Lemma 4.1 in Chen & Fan (2006a). From Theorem 1, as expected, the marginal estimator \widehat{u}_t has little effect on $\widehat{\delta}$ in a large sample, due to the fact that \widehat{u}_t is estimated at a faster convergence rate than the nonparametric estimator $\widehat{\delta}$.

Corollary 1: It follows from Theorem 1 that, for a fixed point $z \in \Phi_z$, as $T \rightarrow \infty$, we have

$$\sqrt{Th} \left(\widehat{\eta}(z) - \eta(z) - \frac{h^2}{2} \eta''(z) \mu_2 \right) \xrightarrow{d} N \left(0, \frac{\nu_0}{f(z)} \Psi(z) \circ \{ [(g^{-1})'(\eta(z))] [(g^{-1})'(\eta(z))]^\top \}^{-1} \right).$$

Corollary 2: By the continuity of the inverse link function $g^{-1}(\cdot)$, for a fixed point $z \in \Phi_z$, as $T \rightarrow \infty$, we have

$$\sqrt{Th} \left(\widehat{\xi}(z) - \xi(z) - h^2 B_\xi(z) \right) \xrightarrow{d} N \left(0, \frac{\nu_0}{f(z)} \Psi(z) \right),$$

where $B_\xi(z) = \frac{1}{2} \mu_2 \frac{1}{g'(\xi(z))} \circ \eta''(z)$.

2.4 Model Selection for Conditional Mixture Copula Models

When a mixture copula model contains too many component copulas, there is a risk of overfitting and efficiency loss. To filter out component copulas with small weights and little contribution to the dependence structure, we consider two model selection procedures.

The first method is to apply the information criterion such as AIC or BIC. However, as argued in Section 1, they can not be directly used to the proposed conditional mixture copula model which has varying coefficients. Huang *et al.* (2013) use the BIC-type selector to identify the number of components in the nonparametric mixture of regression models and find it performs well in numerical studies. This motivates us to select the components in the conditional mixture copula model through the BIC-type selector.

Let $|\Phi_z|$ be the length of the support of z and $K * K$ be the convolution of the kernel K . Define $e_k = K(0) - 0.5 \int K^2(t)dt$ and $m_k = \int (K(t) - 0.5K * K(t))^2 dt$. Following Huang *et al.* (2013), we can calculate the value of the BIC by

$$-2L + \log(T) \times df,$$

where

$$L = \frac{1}{T} \sum_{t=1}^T \ln \left\{ \sum_{k=1}^d \hat{\omega}_k(z_t) c_k [\hat{u}_t; \hat{\theta}_k(z_t)] \right\}$$

and

$$df = (2d - 1)r_k e_k |\Phi_z|/h \quad \text{with} \quad r_k = e_k/m_k.$$

Both \hat{u}_t and $\hat{\xi} = (\hat{\omega}_1, \dots, \hat{\omega}_d, \hat{\theta}_1, \dots, \hat{\theta}_d)^\top$ can be obtained by the estimation procedures discussed in Section 2.2. The degree of freedom, which is originally derived for the generalized likelihood ratio test by Fan *et al.* (2001), can be understood as follows. Suppose that we partition the range of z into $|\Phi_z|/h$ intervals with equispaced length h . Hence, the effective number of each parameter is approximately proportional to $|\Phi_z|/h$. $r_k e_k$ is an adjusting factor that accounts for overlapping intervals due to the local linear fitting. $2d - 1$ is the number of nonparametric estimators in which we minus one to account for the constraint on the weight parameters.

An alternative method to decide which component copulas should be kept or filtered out is to apply a sequence of hypothesis tests. In the classical linear regression models, a sequence of F -tests with the backward elimination or forward addition procedures is used to select the important regressors. We adopt a similar strategy and implement the following test:

$$H_0 : \omega_{i_1}(z) = \dots = \omega_{i_l}(z) = 0 \quad \text{versus} \quad H_1 : \text{not all } \omega_{i_s}(z) = 0,$$

for some $\{i_1, \dots, i_l\} \subset \{1, \dots, d\}$. The model selection is achieved by a sequence of testing procedures above. To simplify the presentation, we only consider the following test:

$$H_0 : \omega_1(z) = \dots = \omega_J(z) = 0 \quad \text{versus} \quad H_1 : \text{not all } \omega_j(z) = 0,$$

and other cases can be implemented in the same manner. Using the local linear fitting with a kernel K and a bandwidth h , we can obtain $\tilde{\xi}(z_t)$ and $\hat{\xi}(z_t)$ under the null hypothesis H_0 and the alternative hypothesis H_1 , respectively.

Define

$$L(H_0) = \sum_{t=1}^T \ell(\hat{u}_t, \tilde{\xi}(z_t)) \quad \text{and} \quad L(H_1) = \sum_{t=1}^T \ell(\hat{u}_t, \hat{\xi}(z_t)),$$

where \hat{u}_t denotes the estimator of the marginal distribution. Fan *et al.* (2001) propose a generalized likelihood ratio test (GLRT) statistic that can be used in many nonparametric testing problems and present the Wilks type of results for various models including the nonparametric regression, varying-coefficient models, generalized varying-coefficient models, varying-coefficient partially linear models, additive models, and spectral density estimation. In the same spirit, we propose a GLRT statistic for the conditional mixture copula models as

$$\lambda_T = L(H_1) - L(H_0).$$

Remark 4. *Acar et al. (2013) show the asymptotic property of the proposed GLRT statistic in the i.i.d. scenario, i.e.,*

$$r_k \lambda_T \xrightarrow{d} \chi_{\mu_T}^2,$$

where $r_k = e_k/m_k$ and $\mu_T = Jr_k e_k |\Phi_z|/h$ with J being the number of testing parameters. However, to the best of our knowledge, the GLRT for copula models with time series data has not been studied, and we leave it to future research.

Because the asymptotic properties of the test statistic require further research, we next propose a bootstrap technology to obtain the p -value of the GLRT statistic as follows:

- (i). Compute the estimators $\tilde{\xi}(z_t)$ and $\hat{\xi}(z_t)$ by using the same bandwidth h under the null hypothesis H_0 and the alternative hypothesis H_1 , respectively. Then, we obtain the GLRT statistic λ_T ;
- (ii). Generate a sample sequence $\{\mathbf{x}_t^*, z_t^*\}_{t=1}^T$ from the original data $\{\mathbf{x}_t, z_t\}_{t=1}^T$ using a stationary bootstrap technique as described in Appendix A;
- (iii). Obtain the marginal distributions estimates $\{\hat{u}_t^*\}_{t=1}^T$ by Step 1 described in Section 2.2;
- (iv). Use the above bootstrap sample to construct the GLRT statistic λ_T^* ; and
- (v). Repeat Steps (ii)-(iv) S times (say, $S=1000$) and obtain S values of the statistic λ_T^* . The p -value of the test is the relative frequency of the event $\{\lambda_T^* > \lambda_T\}$ in the S replications of the bootstrap sampling.

3 Numerical Studies

In this section, we investigate the finite-sample performance of our estimation and model selection procedures through a series of numerical studies. For simplicity, we assume the mixture copula model consists of the Clayton, Gumbel, Frank, and Gaussian copulas. They are widely used in empirical studies because they could describe different dependence structures. Specifically, the Clayton copula exhibits strong lower tail dependence, and can well capture cases such as two markets are likely to crash simultaneously. The Gumbel copula shows strong upper tail dependence and can be an appropriate model when two markets are likely to boom together. The Gaussian copula and the Frank copula exhibit symmetric tail dependence.

The working mixture copula model is then formulated as

$$C(u_1, u_2; \boldsymbol{\omega}(z), \boldsymbol{\theta}(z)) = \omega_{Cl}(z)C_{Cl}(u_1, u_2; \theta_{Cl}(z)) + \omega_{Gu}(z)C_{Gu}(u_1, u_2; \theta_{Gu}(z)) + \omega_{Fr}(z)C_{Fr}(u_1, u_2; \theta_{Fr}(z)) + \omega_{Ga}(z)C_{Ga}(u_1, u_2; \theta_{Ga}(z)),$$

where $C_{Cl}(\cdot)$, $C_{Gu}(\cdot)$, $C_{Fr}(\cdot)$ and $C_{Ga}(\cdot)$ denote the Clayton, Gumbel, Frank, and Gaussian copulas, respectively. Following Abegaz *et al.* (2012), we generate the covariate z from the truncated normal distribution with mean 0 and variance 9, and then consider four different types of copula parameter function $\theta(z)$ with $z \in [-2, 2]$:

- Model 1: $\theta(z) = 10 - 1.5z^2$;
- Model 2: $\theta(z) = 10 - 0.02z^2 + 0.4z^3$;
- Model 3: $\theta(z) = 3 + z + 2e^{-2z^2}$;
- Model 4: $\theta(z) = 5 + 2 \sin(\pi z) + 2e^{-16z^2}$.

For simplicity, we assume that the first marginal distribution u_1 follows the normal distribution $N(1, 0.5)$ and the second marginal distribution u_2 follows the student's t -distribution with 4 degrees of freedom. For each sample we calculate the estimates $\hat{\theta}$ at 101 equally-spaced grid points $z_i = -1.95 + 0.039i$ for $i \in \{0, 1, \dots, 100\}$. Similar to Acar *et al.* (2011) and Abegaz *et al.* (2012), we use the local linear fitting with the regular normal kernel. Each simulation is repeated $M = 1000$ times with the sample size $T \in \{200, 500, 1000\}$.

For comparing purposes, besides the proposed conditional mixture copula method (CM), we additionally consider another popular estimation method for mixture copula. Cai & Wang (2014, CW hereafter) propose a copula selection approach via penalized likelihood plus a shrinkage operator, and establish the asymptotic properties of the proposed penalized likelihood estimator. Similar to CM, this method can also select appropriate copula function and estimate the related parameters simultaneously. The main difference is that CW is only applicable to a mixture copula model with constant weights and copula parameters. In this section, we mainly compare the performance of CM with CW when data are indeed

generated from conditional mixture copulas. For completeness, we will also investigate CM's performance when the true model is a constant mixture copula.

We begin with a simple scenario that data are generated from an individual copula. That is, the true model is an individual copula selected from the four candidates. For each individual copula used to generate data, we assume the function of the parameter follows one of Models 1-4 listed above. Then, we fit the four-component mixture model to the generated data and investigate the performance of the proposed CM method. In Table 1, we conduct both CM and CW, and report the percentage that each copula is correctly (incorrectly) selected and the mean squared errors (MSEs) of the copula parameter estimates over the 101 grid points, which is defined as

$$MSE(\hat{\theta}) = \frac{1}{M} \frac{1}{101} \sum_{j=1}^M \sum_{i=1}^{101} \left(\hat{\theta}_j(z_i) - \theta(z_i) \right)^2.$$

Table 1 shows that the MSEs of copula parameter estimates by both methods decrease for all four functional forms of $\theta(z)$ as the sample size T increases from 200 to 1000. Here, because the Gaussian copula's parameter θ is ranged between -1 and 1, for Gaussian, we recalibrate the four models by dividing 10 for Model 1 and 15 for Models 2-4. As anticipated, MSEs by CM are remarkably lower than those by CW, indicating larger estimation losses produced by CW when parameters in a copula model are indeed conditioning on a covariate. Besides the estimation accuracy measured by MSE, considering that the parameter functions are assumed to follow Models 1-4 which exhibit different patterns, we additionally examine the quality of the CM estimators by checking their estimated paths along the covariate z . Specifically, we plot the estimated paths of Clayton, Gumbel, Frank, and Gaussian with the sample size $T = 1000$ in Figure 1. The black solid curves in the four panels of each row denote the true copula parameter paths $\theta(z)$, which respectively follow Models 1-4, and the other two curves respectively denote the means (red dotted) and medians (blue dashed) of the copula parameter estimates by CM at the 101 grid points from 1000 simulations. The two black dotted-dashed curves connect the 5% and 95% percentiles of the copula

parameter estimates at the 101 grid points. As a comparison, the mean of the estimated copula parameters by CW from 1000 simulations is also plotted and denoted by the brown solid line. For the four candidate copulas, Figure 1 shows that both the mean and median curves by CM are close to the true paths in all four models. Even in Model 4 which contains the complicated sinus function, the performance of the CM estimator is still quite good. On the other hand, the copula parameter estimated by CW is a constant and therefore cannot detect the dynamics in copula parameters conditioning on the covariate. We also examine the results of model selection through the information criterion method and the CW method. In Table 1, values without parentheses represent rates of correctly selected copulas (accurate rates), while values with parentheses indicate rates that copulas are incorrectly selected (inaccurate rates). One can easily observe that the proposed CM method performs reasonably well in selecting the correct individual copula from the mixture model because the true copula is always chosen with 100% chance, and the rates of incorrect selection shrink when the sample size T increases. CW also exhibits good performance in selecting the true candidate copula. In sum, in terms of parameter estimation accuracy exhibited by MSE and model selection accuracy documented by accurate (inaccurate) rate, the proposed CM method displays excellent performance when the true model is an individual copula. Although CW also exhibits high accuracy in copula selection, it fails to capture the dynamic copula parameters in a conditional copula setup. For completeness, we additionally check the results by the proposed hypothesis test procedure with the 0.05 significance level and find similar results. The detailed simulation results are displayed in Table D4 of Appendix D.

[INSERT TABLE 1 AND FIGURE 1 ABOUT HERE.]

Next, we investigate the performance of CM and CW when the true model is a mixture of two copulas. In other words, for the four-component mixture copula, we assume two candidate copulas' weights uniformly equal to zero while the other two copulas' weights respectively equal to $(1+z)^2/29+0.3$ and $1-((1+z)^2/29+0.3)$. For the two component copulas with non-zero weights, we further assume their parameter functions follow different

patterns determined by the four models discussed above. Tables 2 and 3 document the MSEs of copula parameter estimates and the accurate and inaccurate (in parentheses) rates of copula selection by CM and CW, respectively. For example, Panel 1 of Tables 2 and 3 concerns the case that the true mixture model is constructed by Clayton and Gumbel. Comparing results in the two tables, one can observe that the MSEs of the two copulas' parameter estimates by CM decline substantially when the sample size T increases from 200 to 1000, and the magnitudes are remarkably lower than those by CW. In addition, Table 2 indicates that the accurate and inaccurate rates by CM display promising improvement as the sample size increases: in Panel 1, the rates that the Frank and Gaussian copulas are incorrectly selected decrease while the accurate rate for Gumbel increases to about 99% when $T = 1000$. There is a 100% probability that the Clayton copula is correctly selected. CW displays similar patterns in copula selection, as can be seen in Table 3. We have similar findings from the other five combinations in both tables. As in the prior individual copula scenario, in Figure 2 we plot the paths of copula parameter estimates, and compare the estimated paths by CM and CW with the true paths. To save space, here we only demonstrate the six combinations of parameter functions in Panel 2 of Tables 2 and 3 when $T = 1000$. In Figure 2, the two plots in each column represent a combination of two copulas with different parameter functions. For example, in Figure 2(a), the upper plot demonstrates the true path (Model 1), the mean and median of the estimated paths by CM, and the mean of the estimated path by CW for Clayton, while the lower plot contains the corresponding results for Frank. In general, Figure 2 shows that the copula parameters of the Clayton-Frank mixture can be well estimated by the proposed CM method in all six combinations, while the estimates by CW are constants and unable to detect how copula parameters varies with the covariate.

[INSERT TABLES 2 AND 3 AND FIGURE 2 ABOUT HERE.]

In addition to copula parameters, it is also worth examining the performance of the proposed CM method in estimating weights of each candidate copula. Using the same 4-component mixture model, without loss of generality, we assume the weight of the first

copula in the true mixture model follows $(1 + z)^2/29 + 0.3$, and that of the second copula follows $1 - ((1 + z)^2/29 + 0.3)$. We display the MSEs of the weight estimates by both CM and CW for all six mixture models in Table 4. As can be seen therein, the MSEs of the weight estimates by both methods decrease in all cases when sample size T increases, and CM uniformly exhibits lower MSEs than CW. Similar to copula parameter estimates, we also draw the true paths, the mean and median of the estimated paths by CM, and the mean of the estimated paths by CW for weight parameters in Figure 3. It corresponds to the weights for the Clayton-Frank combination in Panel 2 of Table 4 when $T = 1000$. Figure 3 shows that both the mean and median paths of the weight estimates by CM track the true paths closely, implying a good performance of the proposed CM method in estimating the weights of the Clayton-Frank combination. Weight estimates by CW are constants and thus cannot detect how weights vary with the covariate.

[INSERT TABLE 4 AND FIGURE 3 ABOUT HERE.]

Finally, we compare the performance of CM and CW when the true mixture copula model exhibits constant parameters. To save space, we only consider two scenarios. First, we assume data are generated from an individual Clayton copula with the dependence parameter equals either 5 or 7. Second, we generate data from a combination of the Clayton and Gumbel copulas. For simplicity, we assume the two copulas are equally weighted with two pairs of constant copula parameters, $(\theta_{Cl} = 5, \theta_{Gu} = 4)$ and $(\theta_{Cl} = 7, \theta_{Gu} = 6)$. Table 5 shows that, when the true copula model exhibits constant parameters, the CW method exhibits better performance than the proposed CM method because the MSEs produced by CW are slightly lower than those by CM. This should be expected because CW exhibits higher estimation efficiency when parameters in a mixture copula are indeed constant. In terms of picking up the correct copula functions, both methods exhibit similar performance.

[INSERT TABLE 5 ABOUT HERE.]

We additionally conduct simulations to investigate the performance of our method when: (i) the conditional mixture model contains 3- and 4-dimensional copulas, and (ii) data are

generated from copulas not included in the candidate set (i.e., the mixture copula is misspecified). These additional simulation results, displayed in Appendix D, provide further evidence that the proposed CM method still displays good performance in the two scenarios.

4 An Empirical Illustration

In this section we apply the proposed methods to investigate how the dependence structures among the international equity markets evolve with the volatility in exchange rate markets. Equity price and exchange rate are two important financial variables that are closely linked to each other. Shocks in the exchange rate market cause fluctuations in the value of a domestic currency, impacting trade flows, capital movements and equity prices. Therefore, understanding the relation between exchange rate markets and equity markets and the spillover effect of exchange rate markets on equity prices has substantive implications in terms of risk management.

From CRSP, we collect the weekly MSCI equity prices in four developed economies (France, Germany, the United States and the United Kingdom) and the weekly exchange rates among the U.S. dollar (USD), the British pound (GBP) and the euro (EUR). The observations are between 01/07/1999 and 11/07/2018. We transform the weekly equity prices and exchange rates into log returns by taking the first order differences on their logarithmic levels. The first panel in Table 6 documents some summary statistics of the weekly log returns of the equity prices and exchange rates. One can observe that the European stock markets exhibit larger fluctuations than the United States market during the sample period, while the latter gives relatively higher average returns. Compared with the equity markets, both returns and fluctuations are lower in these exchange rate markets. The Jarque-Bera test results show that the null hypothesis of normality is rejected for all six return series. The second panel in Table 6 displays that the linear (Pearson) correlation coefficients across the four equity markets are very high, which is expected considering the economic synchronization of the four developed economies.

[INSERT TABLE 6 ABOUT HERE.]

Preliminary examinations suggest that the autocorrelation and conditional heteroscedasticity exist in these log return series. Thus, we follow Chen & Fan (2006b) and use an AR-GARCH process to model the conditional mean and conditional variance. Specifically, we fit the series of returns to an AR(1)-GARCH(1,1) process written as

$$x_{it} = \gamma_{i0} + \gamma_{i1}x_{i,t-1} + e_{it}, \quad e_{it} = \sigma_{it}\epsilon_{it},$$

where x_{it} denotes the return at time t for country i . The innovations ϵ_{it} are assumed to be i.i.d. The conditional variance is defined as

$$\sigma_{it}^2 = \alpha_{i0} + \alpha_{i1}\sigma_{i,t-1}^2 + \alpha_{i2}e_{i,t-1}^2,$$

where α_{i0} , α_{i1} and α_{i2} are parameters of GARCH(1,1) for country i with $\alpha_{i0} > 0$, $\alpha_{i1} \geq 0$, $\alpha_{i2} \geq 0$ and $\alpha_{i1} + \alpha_{i2} < 1$. Table 7 summarizes the coefficients of the AR(1)-GARCH(1,1) filtering and shows that most estimates are statistically significant. For all cases, the Ljung-Box test statistics are not significant at any conventional levels, implying the effectiveness of the AR-GARCH procedure in filtering out the linear dependence in the series of returns.

[INSERT TABLE 7 ABOUT HERE.]

Because our target is to investigate the dynamic pattern of the dependence among the international equity markets along the path of the exchange rate volatility, we first respectively estimate the volatility of USD-EUR and USD-GBP by the AR(1)-GARCH(1,1) model discussed above. Figure 4 demonstrates the time series plots of the four countries' weekly equity prices along the estimated volatility of the exchange rates (black dashed) for USD-GBP and USD-EUR. Unlike the equity prices which substantially fluctuate during the sample period, the volatility paths of exchange rates are relatively stable. As a matter of fact, the exchange rate markets witnessed a tranquil period during 1999-2007 with the estimated volatility of USD-EUR and USD-GBP ranged between 0.007 and 0.02, whereas the equity markets in

these countries were shocked by a sequence of events such as the recession induced by the burst of the dotcom bubbles, the 9/11 terrorism attack, and two military operations against Afghanistan and Iraq, etc. However, sometimes the tranquility in the exchange rate markets could also be interrupted by a domestic or international event. As can be seen in the first two panels in Figure 4, there is a remarkable peak on the volatility path of USD-EUR in 2008 when the global financial crisis caused severe recession and stock markets crashed in all developed economies. For example, Figure 4(a) shows that just in 2008 the stock price dropped by 46% in France (blue solid) and 40% in the United States (red solid), while the volatility of the USD-EUR exchange rate is nearly tripled at the end of 2008, jumping from 0.012 to 0.034. One could observe similar patterns from the other two pairs. One distinctive feature of the USD-GBP volatility path is that, besides the spike in 2008, there is another remarkable peak in June of 2017 due to panics among investors induced by the passage of the Brexit referendum.

[INSERT FIGURE 4 ABOUT HERE.]

Next, we fit the filtered equity returns (i.e., the residuals) to the conditional mixture copula model that contains the Clayton, Gumbel, and Frank copulas. In the first step, we apply the two proposed model selection procedures to determine which candidate copula(s) should be included in the mixture model. We firstly consider the information criterion method. Because the mixture copula model contains three components, for each pair of markets, we need to respectively estimate $2^3 - 1 = 7$ copula models and calculate their corresponding BIC values. Then, as discussed in Section 2.4, we make a comparison and choose the model with the lowest BIC value among the seven BICs. Table 8 documents the BIC values of the seven models for the three pairs of equity markets. It shows that the mixture model with the Clayton and Frank copulas are selected for all three pairs as it exhibits the lowest BIC among the seven BIC values. In other words, based on the comparison among the BICs associated with the seven candidate models, the final mixture copula should be constructed as $\omega_{Cl}C_{Cl} + \omega_{Fr}C_{Fr}$.

[INSERT TABLE 8 ABOUT HERE.]

For comparing purposes, we alternatively apply the hypothesis test method to check which component copulas should be kept. Similar to the procedures of the backward elimination, we start with a mixture model with all three component copulas. Then we test whether the three weight parameters ω_{Cl} , ω_{Gu} and ω_{Fr} respectively equal to zero, and remove the copula whose weight parameter's p -value is the highest among those greater than 0.05, the significance level. We then refit the model until the p -values of the weight parameters of the remaining component copulas are all lower than 0.05. The sequence of the hypothesis tests and the results are displayed in Table 9. In Panel 1, the hypothesis test results show that the p -values for the Gumbel copula's weight parameters are remarkably greater than 0.05 in all three market pairs, indicating that the weight parameters of the Gumbel copula are insignificantly different from zero and the Gumbel copula should be filtered out from the mixture models in the first step. Next, we refit the mixture model which only contains the Clayton and Frank copulas. Panel 2 of Table 9 suggests that both component copulas should be kept as the p -values of their weight parameters are lower than 0.05 in all three market pairs. That is, the final mixture copula selected by the hypothesis test method is consistent with the one selected by the information criterion method.

[INSERT TABLE 9 ABOUT HERE.]

When a mixture copula contains many component copulas, it would be efficient if we can simultaneously filter out several candidate copulas in one step. To this end, in this example we also apply the hypothesis test by firstly examining the joint hypothesis tests. Specifically, in the first step, we respectively test whether each two of the three copulas' weight parameters are simultaneously equal to zero. Failing to reject the null hypothesis of any of these joint tests indicates that the mixture copula only contains an individual copula. In our example, we reject the null hypotheses of all the joint tests for the three market pairs at the 0.05 significance level, suggesting that the mixture model should include more than one component copula. Next, we implement hypothesis tests to examine whether the three weight parameters are respectively equal to zero. The test results show that, for all three pairs, we cannot reject $\omega_{Gu} = 0$ at any conventional levels. Therefore, the final choice of

the mixture model should be constructed by the Clayton and Frank copulas. The detailed results for this sequence of hypothesis tests are available upon request.

Our finding is in line with Garcia & Tsafack (2011) who argue that a combination of the Gaussian and rotated Gumbel copulas outperforms the other models in examining comovements among stock returns along the exchange rate volatility. Given the observed downward comovements among the equity markets when the exchange rates become extremely volatile, it is expected that the Clayton copula should be kept in the mixture model to capture the lower tail dependence. The inclusion of the Frank copula is also intuitive because, as Figure 4 shows, except for the extreme scenario in 2008, in general we do not observe obvious tail dependence of equity returns along the exchange rate volatility over the sample period.

Subsequently, we fit the data to the selected conditional mixture copula model and obtain the estimated weight and copula parameters through the proposed method. Since parameters from different copulas are not directly comparable, here we transform the copulas' parameter estimates into Kendall's τ s. As a measure of concordance between random variables, Kendall's τ is invariant to nonlinear transformations and thus can capture nonlinear dependence which is unable to be measured by the linear correlation coefficient. For example, for X_1 and X_2 with their respective CDFs u_1 and u_2 , it is possible to express Kendall's τ in terms of a copula which connects the two random variables as $\tau = 4 \int_0^1 \int_0^1 C(u_1, u_2; \theta) dC(u_1, u_2; \theta) - 1$. For Clayton, Kendall's $\tau_{Cl} = \frac{\theta}{\theta+2}$. For Gumbel, $\tau_{Gu} = 1 - \frac{1}{\theta}$. For Frank, $\tau_{Fr} = 1 - \frac{4}{\theta} \left(1 - \int_0^\theta \frac{t}{e^t - 1} dt \right)$.

Figure 5 demonstrates the estimates (solid curves) of Kendall's τ s and the weights of the Clayton and Frank copulas for the U.S.-Germany equity returns along the volatility of USD-EUR, and the 5% and 95% percentiles (dashed curves) obtained through the proposed block bootstrap method. Figure 5(a) shows that the magnitude of the lower tail dependence, measured by Kendall's τ of Clayton, doubles from 0.15 to about 0.3 as the volatility of USD-EUR is more than tripled from about 0.01 to 0.035. Such strengthened asymmetric dependence is further amplified by the increasing weight associated with the Clayton copula, as displayed by Figure 5(b): as the exchange rate between the U.S. dollar and the euro

becomes increasingly volatile, the effect of the lower tail dependence turns to be more dominant. On the contrary, both the estimates of Kendall's τ and the weight of the Frank copula decrease as the volatility of USD-EUR increases. Even though the magnitude of the decline in Kendall's τ for the Frank copula is relatively small, its dominance in the dependence structure is remarkably weakened due to the decreased weight in the mixture model. For the pairs of U.S.-France and U.S.-UK, Figures 6 and 7 show quite similar patterns: when the exchange rate becomes increasingly volatile, the lower tail dependence is strengthened as the estimates of Kendall's τ and the weight associated with the Clayton copula simultaneously increase, while both Kendall's τ and the weight of the Frank copula shrink. Given the close economic and political connections of the three countries to the United States, the similar patterns of Kendall's τ s along the exchange rate volatility should be expected. When the exchange rate becomes extremely volatile (e.g., during the global economic recession), the weight of the Clayton copula exceeds that of the Frank copula, indicating that the lower tail dependence dominates the mixture dependence structure so that the two equity markets exhibit a higher probability to crash simultaneously. The fact that higher volatility in exchange rate markets is associated with more extreme asymmetric dependence among equity markets is not only in line with our observations in Figures 5-7, but also consistent with the findings in the literature such as Longin & Solnik (2001) and Garcia & Tsafack (2011). An extreme jump in the exchange rate usually leads to an extreme comovement in equity markets. As Garcia & Tsafack (2011) argue, when a sudden and unexpected shock hits an economy with a very active currency market, transmission through the latter makes a downside comovement of equity markets more likely than in a calm period of the exchange rate market. This may partially explain why the Clayton copula's weight and Kendall's τ both increase as the exchange rate market becomes more volatile.

[INSERT FIGURE 5-7 ABOUT HERE.]

5 Conclusion

This paper proposes a semiparametric conditional mixture copula model in which both weight and copula parameters can vary with a covariate in a nonparametric way. The conditional mixture copula exploits the advantages of both the conditional copula which can capture a covariates impact on the degree of dependence (see Acar *et al.*, 2011, Patton, 2012, and Fermanian & Lopez, 2018), and the mixture copula which can combine copula families with different dependence patterns (see Hu, 2006, and Cai & Wang, 2014). Therefore, it provides extra flexibility and an unified way for practitioners to measure the dependence pattern and the degree of dependence.

In the theoretical part, we provide a two-step estimation procedure to separately estimate the marginal distributions and the weight and copula parameters in the model, and the large sample properties of these estimators are derived. Moreover, we introduce two model selection approaches to choose an appropriate conditional mixture copula model from a large copula candidate set. Monte Carlo simulation results confirm that the proposed estimation and model selection procedures perform reasonably well in estimating unknown parameters and selecting component copulas. The only exception is when weights and copula parameters in a mixture copula are constants: simulation results show that although the proposed conditional mixture copula estimation method still provides accurate copula selection, its estimation accuracy, measured by MSE, becomes slightly lower than the constant mixture copula estimation method proposed by Cai & Wang (2014). In an empirical illustration, we apply the proposed methods to examine how the dependence structures among the international equity markets evolve with the volatility in the exchange rate markets and find that both the weight of the Clayton copula and the degree of the lower tail dependence among the equity markets remarkably increase when the exchange rate markets become more volatile.

In practice, because whether weights and copula parameters are constants or not is unknown *ex ante*, the conditional mixture copula is an ideal model if practitioners have strong belief that both the pattern of dependence (summarized by weights of component copulas) and degree of dependence (measured by copula parameters) are affected by certain

covariate. For example, applying the proposed conditional mixture copula, practitioners can extend Zimmer (2012) and investigate how strength and direction of comovement among housing markets in the United States evolved with certain economic indicator such as per capita disposable income in the past four decades. In this analysis, as in the empirical illustration in Section 4, we need the conditional mixture copula to detect effects of the covariate on dependence structure. Alternatively, if practitioners need a quick examination on tail dependence and degree of dependence, the CW method would be an useful model with fewer parameters to estimate. In this study, we skip testing the irrelevance of the covariate and refer interested readers to Acar *et al.* (2013), Gijbels *et al.* (2017), and Derumigny & Fermanian (2017) for discussions and empirical illustrations in the i.i.d. scenario. Testing the simplifying assumption with weakly dependent data deserves a separate study in the future.

Some interesting future research topics related to this article should be mentioned. First, the proposed method can be extended to a higher dimension of Z because the dependence structure could be potentially affected by several economic variables simultaneously. In other words, the copula parameter $\theta_k(\cdot)$ in equation (1) can be written as $\theta_k(\gamma^\top \mathbf{Z}_t)$, where $\theta_k(\gamma^\top \mathbf{Z}_t)$ now is a flexible function of the so-called single-index $\gamma^\top \mathbf{Z}_t$, i.e., a linear combination of p_z -dimensional economics variable \mathbf{Z}_t with γ being a p_z -dimensional vector loading, and $\mathbf{Z}_t = (Z_{1t}, \dots, Z_{p_z t})^\top$. Second, our model can be applied by empirical practitioners to study how dependence structures among the international stock markets abruptly changed amid the outbreak of the COVID-19 pandemic when equity prices crashed in global markets.

Acknowledgements

The authors would like to thank the Editor-in-Chief Håkon Gjessing, one associate editor, and two anonymous reviewers for their guidance and insightful comments which substantially improve the paper. Liu's research is supported by the National Natural Science Foundation of China (Grant No. 71803160). Long's research is partly supported by the Kurzius Family

Early Career Professorship in Economics at Tulane University. Yang's research is partly supported by the Major Program of NSFC (71991474), Humanity and Social Science Youth Foundation of Ministry of Education of China (19C10421028), the Fundamental Research Funds for the Central Universities (19wkpy61) and the Major Program of the National Social Science Foundation of China (17ZDA073). Cais research is partially supported by the National Natural Science Foundation of China (Grant Nos. 71631004 and 72033008).

Supporting information

Appendices for this article are available online. In the Appendices we describe a stationary bootstrap technique which is adopted in Sections 2.4, and discuss some practical issues including an EM algorithm, the selection of the bandwidth, and confidence intervals. We also present the proofs of the key results and additional simulation results in the Appendices.

References

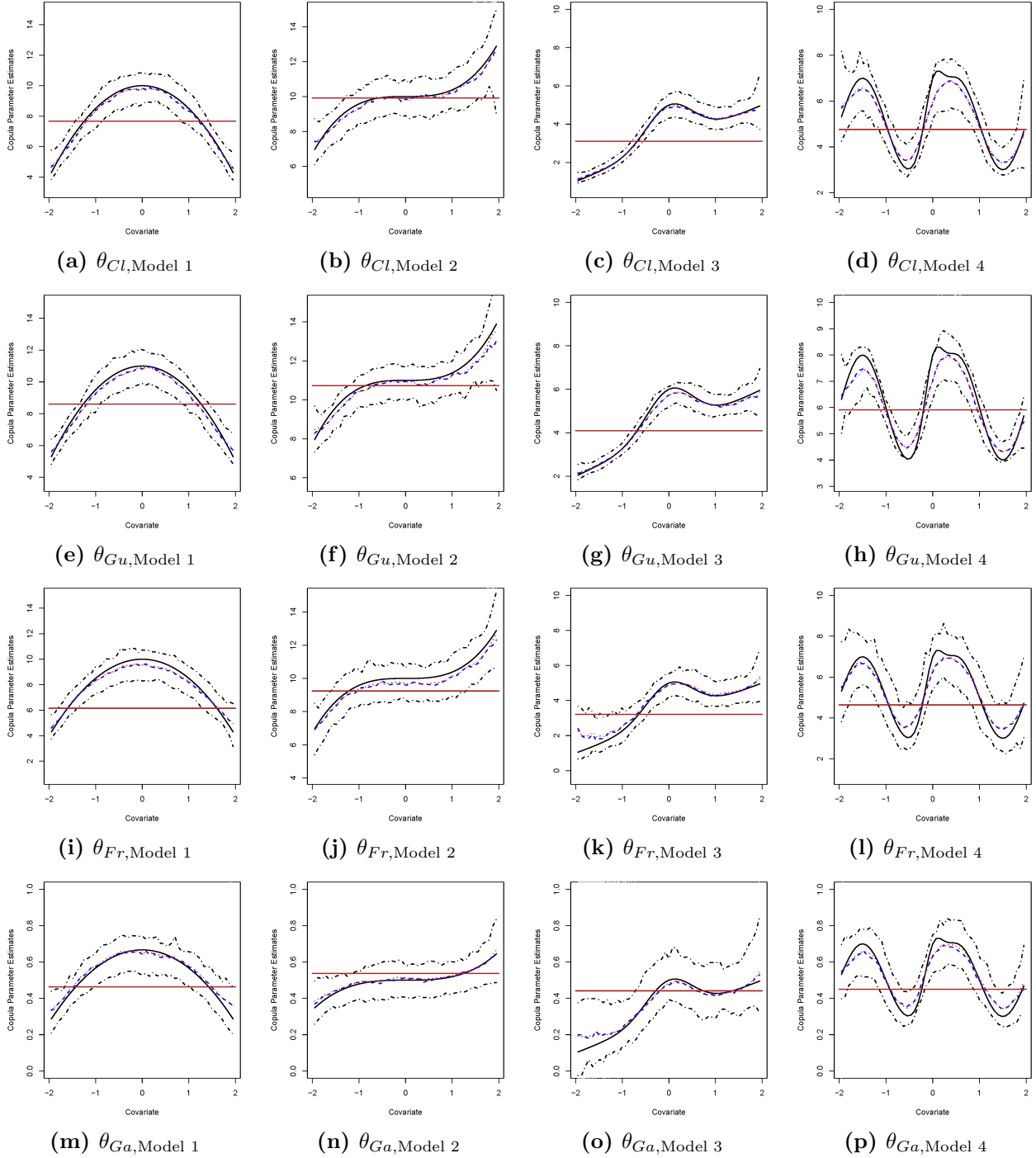
- [1] Abegaz, F., Gijbels, I., & Veraverbeke, N. (2012). Semiparametric Estimation of Conditional Copulas. *Journal of Multivariate Analysis*, **110**, 43-73.
- [2] Acar, E., Craiu, R., & Yao, F. (2011). Dependence Calibration in Conditional Copulas: A Nonparametric Approach. *Biometrics*, **67**, 445-453.
- [3] Acar, E., Craiu, R., & Yao, F. (2013). Statistical Testing of Covariate Effects in Conditional Copula Models. *Electronic Journal of Statistics*, **7**, 2822-2850.
- [4] Cai, Z. (2002). Regression Quantiles for Time Series. *Econometric Theory*, **18**, 169-192.
- [5] Cai, Z., & Wang, X. (2014). Selection of Mixed Copula Model via Penalized Likelihood. *Journal of the American Statistical Association*, **109**, 788-801.
- [6] Chen, X., & Fan, Y. (2006a). Estimation of Copula-Based Semiparametric Time Series Models. *Journal of Econometrics*, **130**, 307-335.

- [7] Chen, X., & Fan, Y. (2006b). Estimation and Model Selection of Semiparametric Copula-Based Multivariate Dynamic Models under Copula Misspecification. *Journal of Econometrics*, **135**, 125-154.
- [8] Chollete, L., Peña, V., & Lu, C. (2005). Comovement of International Financial Markets. Unpublished Manuscript.
- [9] Derumigny, A., & Fermanian, J. 2017. About tests of the “simplifying” assumption for conditional copulas. *Dependence Modeling*, **5**, 154-197.
- [10] Fan, J., & Gijbels, I. (1996). *Local Polynomial Modelling and Its Applications*. Volume 66. London: Chapman & Hall.
- [11] Fan, J., & Li, R. (2001). Variable Selection via Nonconcave Penalized Likelihood and its Oracle Properties”. *Journal of the American Statistical Association*, **96**, 1348-1360
- [12] Fan, J., Zhang, C., & Zhang, J. (2001). Generalized Likelihood Ratio Statistics and Wilks Phenomenon. *Annals of Statistics*, **29**, 153-193.
- [13] Fan, Y., & Patton, A. J. (2014). Copulas in Econometrics. *Annual Review of Economics*, **6**, 179-200.
- [14] Fermanian, J., & Lopez, O. (2018). Single-Index Copulas. *Journal of Multivariate Analysis*, **165**, 27-55.
- [15] Garcia, R., & Tsafack, G. (2011). Dependence Structure and Extreme Comovements in International Equity and Bond Markets. *Journal of Banking & Finance*, **35**, 1954-1970.
- [16] Giacomini, E., Hardle, W., & Spokoiny, V. (2009). Inhomogeneous Dependency Modelling with Time Varying Copulae. *Journal of Business and Economic Statistics*, **27**, 224-234.
- [17] Gijbels, I., Omelka, M., & Veraverbeke, N. (2017). Nonparametric testing for no covariate effects in conditional copulas. *Statistics*, **51**, 475-509.

- [18] Hafner, C., & Manner, H. (2012). Dynamic Stochastic Copula Models: Estimation, Inference and Applications. *Journal of Applied Econometrics*, **27**, 269-295.
- [19] Hu, L. (2006). Dependence Patterns Across Financial Markets: A Mixed Copula Approach. *Applied Financial Economics*, **16**, 717-729.
- [20] Huang, M., Li, R., & Wang, S. (2013). Nonparametric Mixture of Regression Models. *Journal of the American Statistical Association*, **108**, 929-941.
- [21] Liu, G., Long, W., Zhang, X., & Li, Q. (2019). Detecting Financial Data Dependence Structure by Averaging Mixture Copulas. *Econometric Theory*, **35**, 777-815.
- [22] Patton, A. (2006). Modeling Asymmetric Exchange Rate Dependence. *International Economic Review*, **47**, 527-556.
- [23] Patton, A. (2012). A Review of Copula Models for Economic Time Series. *Journal of Multivariate Analysis*, **110**, 4-18.
- [24] Sklar, A. (1959). Fonctions de répartition à n dimensions et leurs marges. *Publications de l'Institut de Statistique de L'Université de Paris*, **8**, 229-231.
- [25] Zimmer, D. (2012). The Role of Copulas in the Housing Crisis. *Review of Economics and Statistics*, **94**, 607-620.

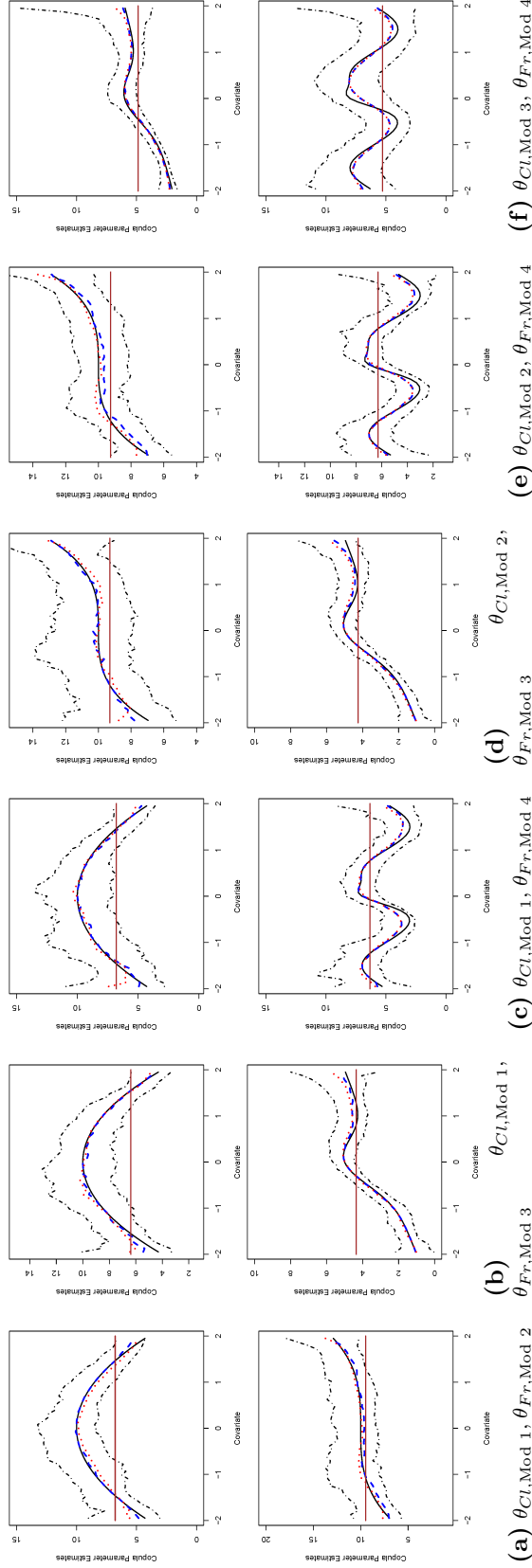
Corresponding author: Bingduo Yang, Lingnan (University) College, Sun Yat-Sen University, Guangzhou, Guangdong 510275, China. Email: bdyang2006@sina.com

Figure 1: Estimated paths for copula parameters (θ) when the true model is an individual copula.



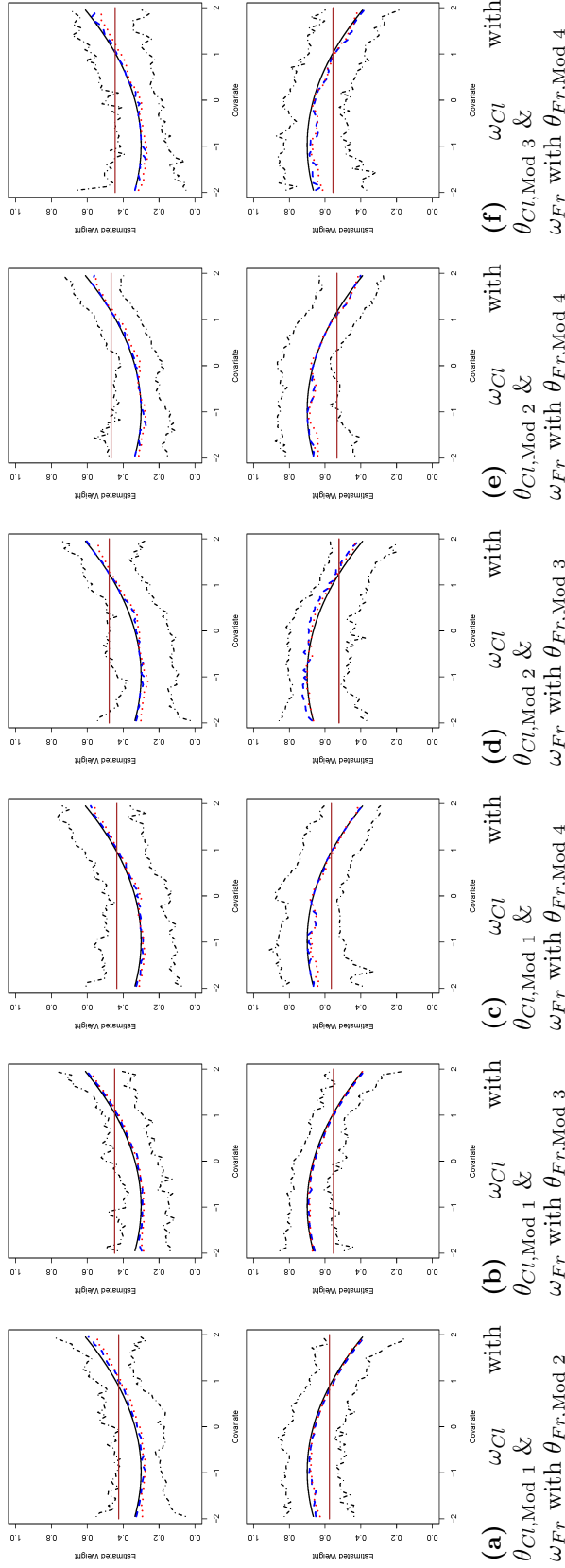
Notes. Panels (a)-(d) denote the estimates of the Clayton parameters which respectively follow Models 1-4. Panels (e)-(h) denote the estimates of the Gumbel parameters which respectively follow Models 1-4. Panels (i)-(l) denote the estimates of the Frank parameters which respectively follow Models 1-4. Panels (m)-(p) denote the estimates of the Gaussian parameters which respectively follow Models 1-4. Model 1: $\theta(z) = 10 - 1.5z^2$. Model 2: $\theta(z) = 10 - 0.02z^2 + 0.4z^3$. Model 3: $\theta(z) = 3 + z + 2e^{-2z^2}$. Model 4: $\theta(z) = 5 + 2\sin(\pi z) + 2e^{-16z^2}$. In each panel, the black solid line denotes the true path of $\theta(z)$. The red dotted line and the blue dashed line respectively denote the mean and median of the copula parameter function estimates at the grid points with 1000 simulations. The brown solid line denotes the mean of the estimates with 1000 simulations by Cai & Wang (2014). The black dotted-dashed lines denote the 5% and 95% percentiles of the copula parameter estimates at the grid points. The sample size $T = 1000$ in all panels.

Figure 2: Estimated paths for copula parameters (θ) when the true model is a combination of the Clayton and Frank copulas.



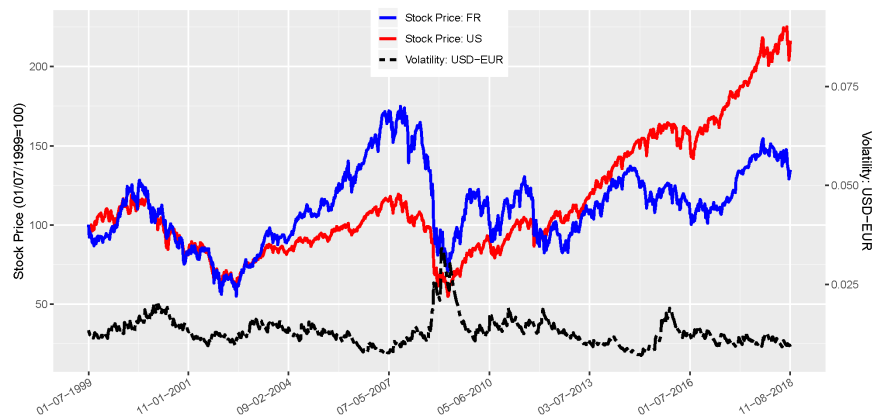
Notes. Panel (a) displays the parameter estimates for the Clayton (up) and Frank (down) copulas when $\theta_{Cl} = \text{Model 1}$ and $\theta_{Fr} = \text{Model 2}$. Panel (b) displays the parameter estimates for the Clayton (up) and Frank (down) copulas when $\theta_{Cl} = \text{Model 1}$ and $\theta_{Fr} = \text{Model 3}$. Panel (c) displays the parameter estimates for the Clayton (up) and Frank (down) copulas when $\theta_{Cl} = \text{Model 1}$ and $\theta_{Fr} = \text{Model 4}$. Panel (d) displays the parameter estimates for the Clayton (up) and Frank (down) copulas when $\theta_{Cl} = \text{Model 2}$ and $\theta_{Fr} = \text{Model 3}$. Panel (e) displays the parameter estimates for the Clayton (up) and Frank (down) copulas when $\theta_{Cl} = \text{Model 2}$ and $\theta_{Fr} = \text{Model 4}$. Panel (f) displays the parameter estimates for the Clayton (up) and Frank (down) copulas when $\theta_{Cl} = \text{Model 3}$ and $\theta_{Fr} = \text{Model 4}$. Model 1: $\theta(z) = 10 - 1.5z^2$. Model 2: $\theta(z) = 10 - 0.02z^2 + 0.4z^3$. Model 3: $\theta(z) = 3 + z + 2e^{-2z^2}$. Model 4: $\theta(z) = 5 + 2 \sin(\pi z) + 2e^{-16z^2}$. The black solid line denotes the true path of $\theta(z)$. The red dotted line and the blue dashed line respectively denote the mean and median of the copula parameter function estimates at the grid points with 1000 simulations. The brown solid line denotes the mean of the estimates with 1000 simulations by Cai & Wang (2014). The black dotted-dashed lines denote the 5% and 95% percentiles of the copula parameter estimates at the grid points. The sample size $T = 1000$ in all panels.

Figure 3: Estimated paths for weights (ω) when the true model is a combination of the Clayton and Frank copulas.

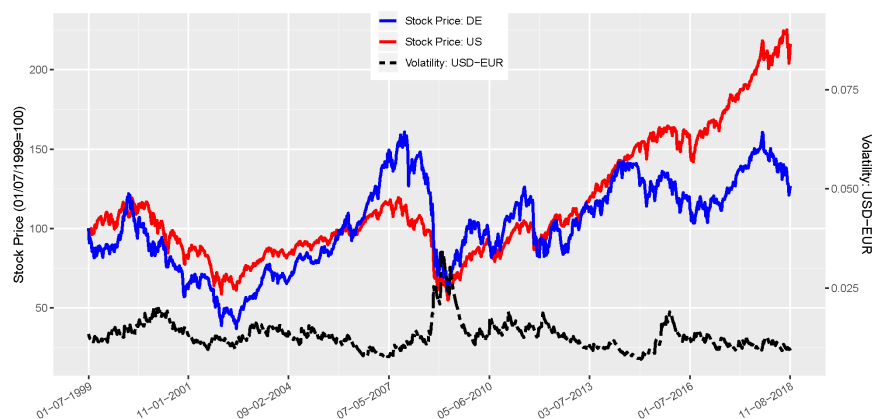


Notes. Panel (a) displays the weight estimates for the Clayton (up) and Frank (down) copulas when $\theta_{Cl} = \text{Model 1}$ and $\theta_{Fr} = \text{Model 2}$. Panel (b) displays the weight estimates for the Clayton (up) and Frank (down) copulas when $\theta_{Cl} = \text{Model 1}$ and $\theta_{Fr} = \text{Model 3}$. Panel (c) displays the weight estimates for the Clayton (up) and Frank (down) copulas when $\theta_{Cl} = \text{Model 2}$ and $\theta_{Fr} = \text{Model 3}$. Panel (d) displays the weight estimates for the Clayton (up) and Frank (down) copulas when $\theta_{Cl} = \text{Model 2}$ and $\theta_{Fr} = \text{Model 4}$. Panel (e) displays the weight estimates for the Clayton (up) and Frank (down) copulas when $\theta_{Cl} = \text{Model 3}$ and $\theta_{Fr} = \text{Model 4}$. Panel (f) displays the weight estimates for the Clayton (up) and Frank (down) copulas when $\theta_{Cl} = \text{Model 3}$ and $\theta_{Fr} = \text{Model 4}$. The black solid line denotes the true path of $\omega(z)$. The red dotted line and the blue dashed line respectively denote the mean and median of the weight estimates at the grid points with 1000 simulations. The brown solid line denotes the mean of the estimates with 1000 simulations by Cai & Wang (2014). The black dotted-dashed lines denote the 5% and 95% percentiles of the weight estimates at the grid points. The sample size $T = 1000$ in all panels.

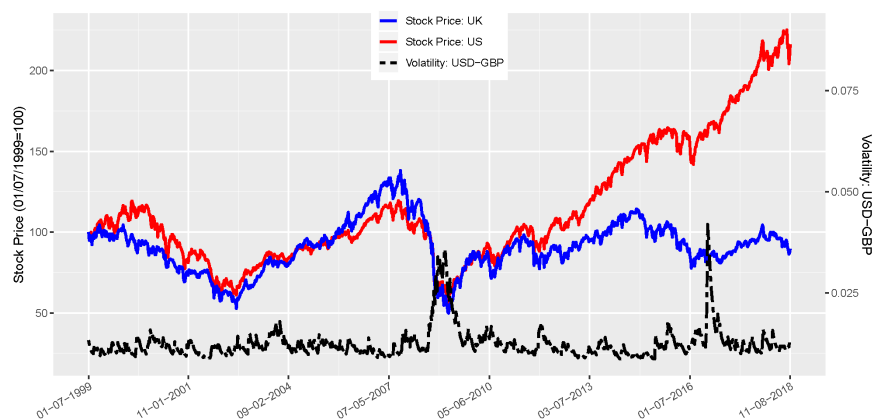
Figure 4: Time series plots for weekly equity prices and estimated volatility of exchange rates.



(a) U.S. - France



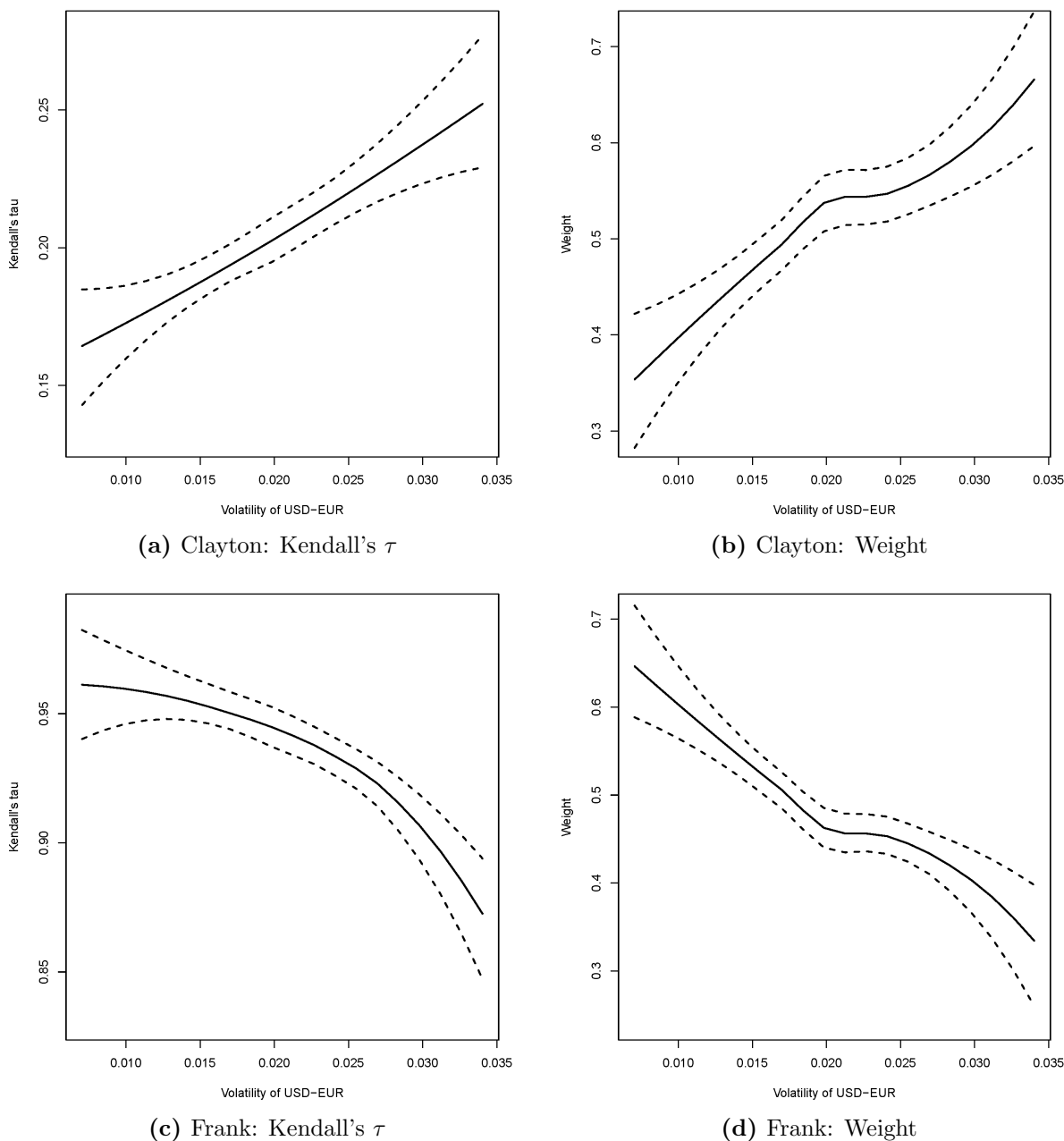
(b) U.S. - Germany



(c) U.S. - UK

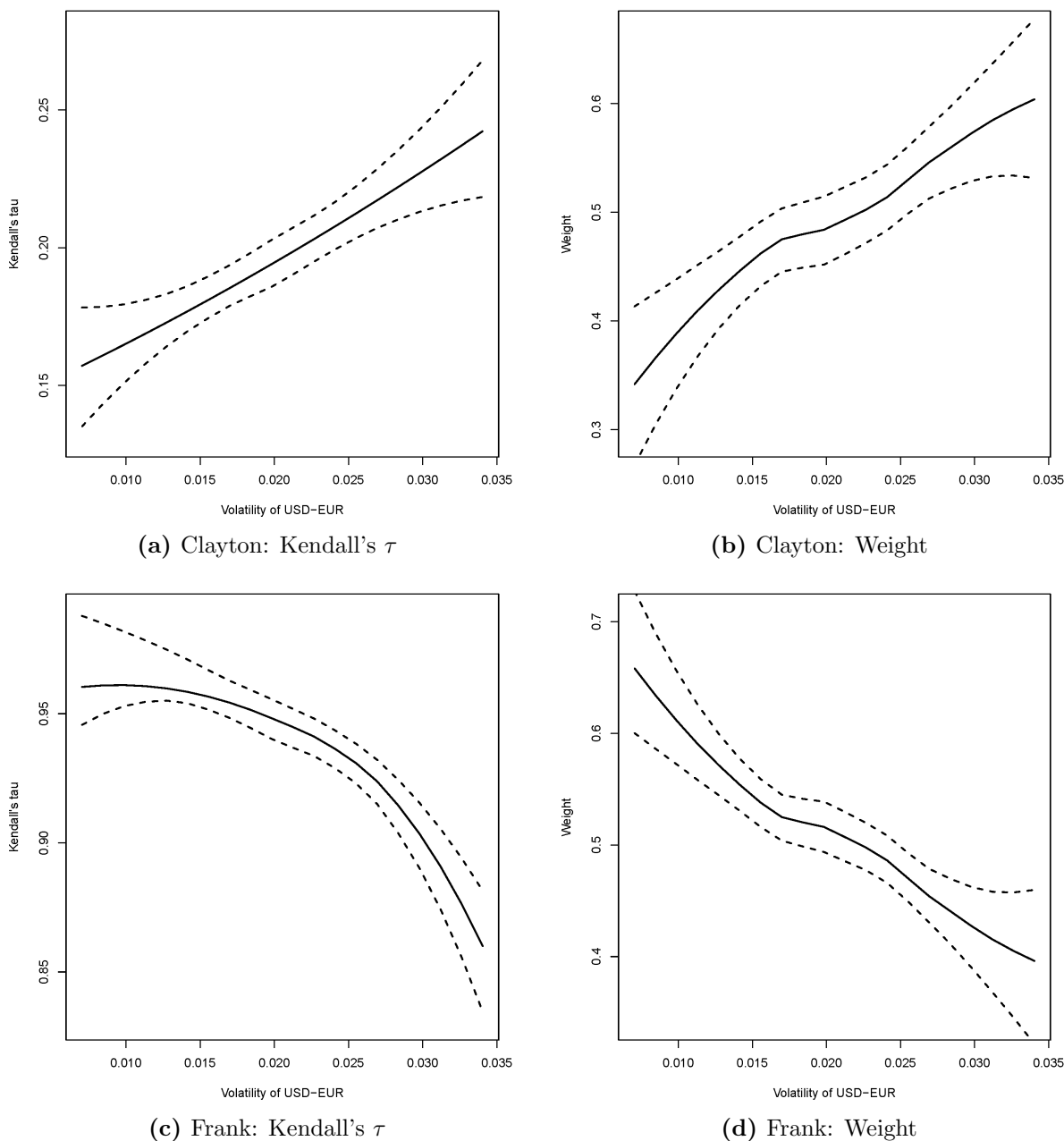
Notes. This figure displays four developed economies' weekly MSCI stock prices (01/07/1999 = 100) and the volatility of exchange rates estimated by the AR(1)-GARCH(1,1) during 01/07/1999 – 11/07/2018. Panel(a) plots the volatility of the exchange rate between the US dollar and the euro (black dashed) and the MSCI stock prices in France (blue solid) and the United States (red solid). Panel(b) plots the volatility of the exchange rate between the US dollar and the euro (black dashed) and the MSCI stock prices in Germany (blue solid) and the United States (red solid). Panel(c) plots the volatility of the exchange rate between the US dollar and the British pound (black dashed) and the MSCI stock prices in the United Kingdom (blue solid) and the United States (red solid).

Figure 5: U.S.-Germany: The estimated paths for Kenall's τ s and weights.



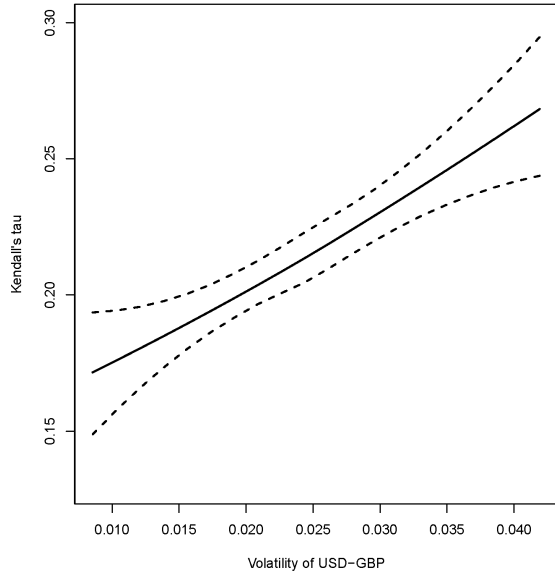
Notes. This figure displays the estimates of Kendall's τ s and the weights of the Clayton and Frank copulas for equity returns between U.S. and Germany, along the estimated volatility of USD-EUR. Panel(a) shows the Kendall's τ estimate of Clayton along the estimated volatility of USD-EUR. Panel(b) shows the weight estimate of Clayton along the estimated volatility of USD-EUR. Panel(c) shows the Kendall's τ estimate of Frank along the estimated volatility of USD-EUR. Panel(d) shows the weight estimate of Frank along the estimated volatility of USD-EUR. The two dashed lines in all four panels denote the 5% and 95% percentiles. The data are at weekly frequency and span from 01/07/1999 to 11/07/2018.

Figure 6: U.S.-France: The estimated paths for Kendall's τ s and weights.

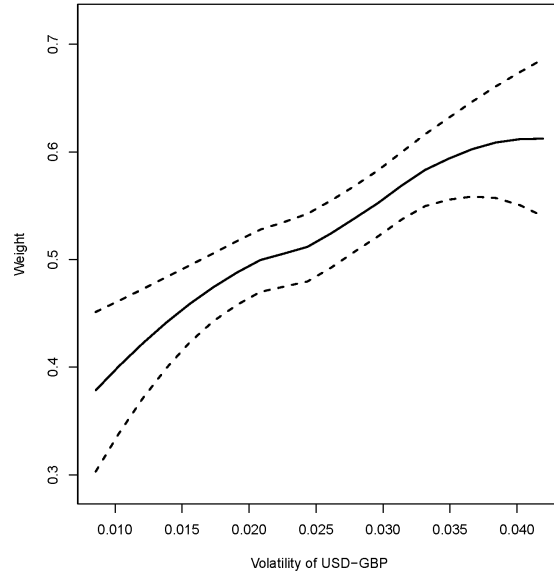


Notes. This figure displays the estimates of Kendall's τ s and the weights of the Clayton and Frank copulas for equity returns between U.S. and France, along the estimated volatility of USD-EUR. Panel(a) shows the Kendall's τ estimate of Clayton along the estimated volatility of USD-EUR. Panel(b) shows the weight estimate of Clayton along the estimated volatility of USD-EUR. Panel(c) shows the Kendall's τ estimate of Frank along the estimated volatility of USD-EUR. Panel(d) shows the weight estimate of Frank along the estimated volatility of USD-EUR. The two dashed lines in all four panels denote the 5% and 95% percentiles. The data are at weekly frequency and span from 01/07/1999 to 11/07/2018.

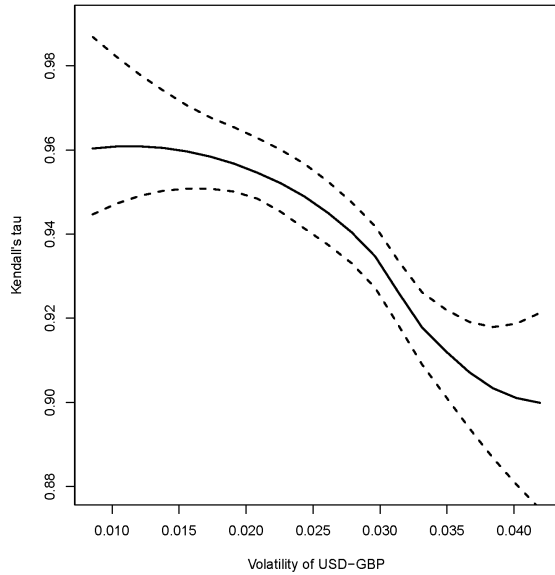
Figure 7: U.S.-UK: The estimated paths for Kenall's τ s and weights.



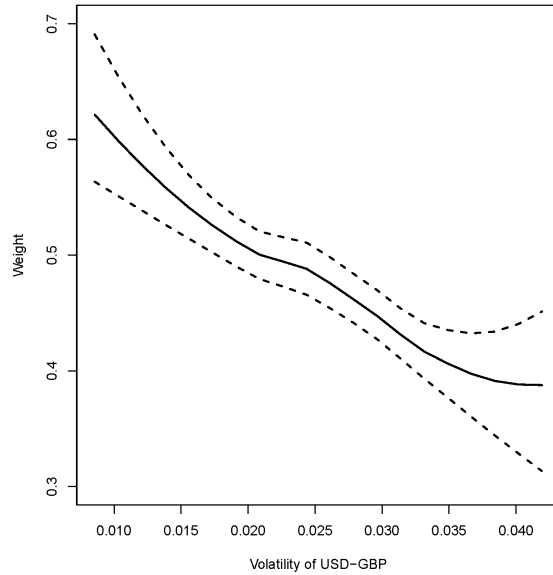
(a) Clayton: Kendall's τ



(b) Clayton: Weight



(c) Frank: Kendall's τ



(d) Frank: Weight

Notes. This figure displays the estimates of Kendall's τ s and the weights of the Clayton and Frank copulas for equity returns between U.S. and UK, along the estimated volatility of USD-GBP. Panel(a) shows the Kendall's τ estimate of Clayton along the estimated volatility of USD-GBP. Panel(b) shows the weight estimate of Clayton along the estimated volatility of USD-GBP. Panel(c) shows the Kendall's τ estimate of Frank along the estimated volatility of USD-GBP. Panel(d) shows the weight estimate of Frank along the estimated volatility of USD-GBP. The two dashed lines in all four panels denote the 5% and 95% percentiles. The data are at weekly frequency and span from 01/07/1999 to 11/07/2018.

Table 1: MSEs of copula parameter estimates and accurate (inaccurate) rates of selection by the proposed conditional mixture copula (CM) and Cai & Wang's (2014) constant mixture copula (CW) when the true model is an individual conditional copula.

		True Copula: Clayton																
		T = 200				T = 500				T = 1000								
		Clayton		Gumbel		Frank		Gaussian		Clayton		Gumbel		Frank		Gaussian		
		CM	CW	CM	CW	CM	CW	CM	CW	CM	CW	CM	CW	CM	CW	CM	CW	
Panel 1	Model 1	MSE	0.731	3.955	-	-	-	-	-	-	-	0.354	1.949	-	-	-	-	
		Rate	1.000	1.000	(0.000)	(0.000)	(0.066)	(0.073)	(0.052)	(0.051)	1.000	1.000	(0.000)	(0.000)	(0.025)	(0.027)	(0.011)	(0.014)
	Model 2	MSE	0.764	2.783	-	-	-	-	-	-	0.456	1.732	-	-	0.252	1.132	-	-
		Rate	1.000	1.000	(0.000)	(0.000)	(0.091)	(0.091)	(0.094)	(0.098)	1.000	1.000	(0.000)	(0.000)	(0.019)	(0.022)	(0.020)	(0.025)
Model 3	MSE	0.801	3.106	-	-	-	-	-	-	0.437	2.008	-	-	0.273	1.576	-	-	
	Rate	1.000	1.000	(0.000)	(0.000)	(0.064)	(0.092)	(0.061)	(0.079)	1.000	1.000	(0.000)	(0.000)	(0.000)	(0.006)	(0.017)	(0.025)	
Model 4	MSE	0.714	2.854	-	-	-	-	-	-	0.495	2.154	-	-	0.236	1.274	-	-	
	Rate	1.000	1.000	(0.000)	(0.000)	(0.061)	(0.062)	(0.079)	(0.091)	1.000	1.000	(0.000)	(0.000)	(0.011)	(0.014)	(0.009)	(0.014)	
Panel 2		True Copula: Gumbel																
		Clayton		Gumbel		Frank		Gaussian		Clayton		Gumbel		Frank		Gaussian		
		CM	CW	CM	CW	CM	CW	CM	CW	CM	CW	CM	CW	CM	CW	CM	CW	
Model 1	MSE	-	-	0.753	3.719	-	-	-	-	0.409	2.481	-	-	0.230	1.716	-	-	
	Rate	(0.000)	(0.000)	1.000	1.000	(0.074)	(0.086)	(0.035)	(0.037)	(0.000)	(0.000)	(0.043)	(0.054)	(0.021)	(0.024)	(0.000)	(0.013)	
Model 2	MSE	-	-	0.929	3.004	-	-	-	-	0.672	2.559	-	-	0.329	1.453	-	-	
	Rate	(0.000)	(0.000)	1.000	1.000	(0.084)	(0.094)	(0.085)	(0.096)	(0.000)	(0.000)	(0.049)	(0.057)	(0.043)	(0.055)	(0.000)	(0.017)	
Model 3	MSE	-	-	0.625	2.387	-	-	-	-	0.433	1.721	-	-	0.248	1.385	-	-	
	Rate	(0.000)	(0.000)	1.000	1.000	(0.071)	(0.076)	(0.072)	(0.070)	(0.000)	(0.000)	(0.030)	(0.044)	(0.048)	(0.050)	(0.000)	(0.019)	
Model 4	MSE	-	-	0.813	3.776	-	-	-	-	0.441	2.568	-	-	0.215	2.070	-	-	
	Rate	(0.000)	(0.000)	1.000	1.000	(0.058)	(0.057)	(0.055)	(0.051)	(0.000)	(0.000)	(0.037)	(0.041)	(0.025)	(0.030)	(0.000)	(0.012)	
Panel 3		True Copula: Frank																
		Clayton		Gumbel		Frank		Gaussian		Clayton		Gumbel		Frank		Gaussian		
		CM	CW	CM	CW	CM	CW	CM	CW	CM	CW	CM	CW	CM	CW	CM	CW	
Model 1	MSE	-	-	-	-	0.835	3.555	-	-	-	-	-	-	0.231	1.711	-	-	
	Rate	(0.000)	(0.000)	(0.011)	(0.012)	1.000	1.000	(0.183)	(0.221)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.026)	(0.038)	
Model 2	MSE	-	-	-	-	0.787	2.867	-	-	-	-	-	-	0.257	1.454	-	-	
	Rate	(0.000)	(0.000)	(0.001)	(0.005)	1.000	1.000	(0.136)	(0.162)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.026)	(0.027)	
Model 3	MSE	-	-	-	-	0.820	3.068	-	-	-	-	-	-	0.211	1.221	-	-	
	Rate	(0.000)	(0.000)	(0.000)	(0.002)	1.000	1.000	(0.219)	(0.220)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.063)	(0.072)	
Model 4	MSE	-	-	-	-	0.846	3.695	-	-	-	-	-	-	0.226	1.387	-	-	
	Rate	(0.000)	(0.000)	(0.000)	(0.000)	1.000	1.000	(0.261)	(0.289)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.058)	(0.067)	
Panel 4		True Copula: Gaussian																
		Clayton		Gumbel		Frank		Gaussian		Clayton		Gumbel		Frank		Gaussian		
		CM	CW	CM	CW	CM	CW	CM	CW	CM	CW	CM	CW	CM	CW	CM	CW	
Model 1	MSE	-	-	-	-	-	-	0.011	0.031	-	-	-	-	-	-	0.002	0.010	
	Rate	(0.000)	(0.000)	(0.000)	(0.000)	(0.107)	(0.176)	1.000	1.000	(0.000)	(0.000)	(0.000)	(0.000)	(0.033)	(0.039)	1.000	1.000	
Model 2	MSE	-	-	-	-	-	-	0.005	0.018	-	-	-	-	-	-	0.001	0.006	
	Rate	(0.000)	(0.000)	(0.000)	(0.000)	(0.082)	(0.115)	1.000	1.000	(0.000)	(0.000)	(0.000)	(0.000)	(0.021)	(0.021)	1.000	1.000	
Model 3	MSE	-	-	-	-	-	-	0.004	0.019	-	-	-	-	-	-	0.001	0.005	
	Rate	(0.000)	(0.000)	(0.000)	(0.000)	(0.099)	(0.120)	1.000	1.000	(0.000)	(0.000)	(0.000)	(0.000)	(0.049)	(0.053)	1.000	1.000	
Model 4	MSE	-	-	-	-	-	-	0.005	0.023	-	-	-	-	-	-	0.001	0.007	
	Rate	(0.000)	(0.000)	(0.000)	(0.000)	(0.049)	(0.088)	1.000	1.000	(0.000)	(0.000)	(0.000)	(0.000)	(0.011)	(0.031)	1.000	1.000	

Table 2: MSEs of copula parameter estimates and accurate (inaccurate) rates of selection by the proposed conditional mixture copula (CM) when the true model is a conditional mixture copula.

Panel 1		Combination: Clayton + Gumbel											
		$T = 200$				$T = 500$				$T = 1000$			
		Clayton	Gumbel	Frank	Gaussian	Clayton	Gumbel	Frank	Gaussian	Clayton	Gumbel	Frank	Gaussian
Model 1 + Model 2	MSE	0.903	0.850	-	-	0.633	0.521	-	-	0.303	0.287	-	-
	Rate	1.000	0.859	(0.151)	(0.234)	1.000	0.936	(0.064)	(0.125)	1.000	0.986	(0.012)	(0.053)
Model 1 + Model 3	MSE	0.893	0.775	-	-	0.472	0.405	-	-	0.288	0.197	-	-
	Rate	1.000	0.901	(0.078)	(0.212)	1.000	0.927	(0.022)	(0.101)	1.000	0.973	(0.007)	(0.061)
Model 1 + Model 4	MSE	0.941	0.901	-	-	0.507	0.476	-	-	0.273	0.286	-	-
	Rate	1.000	0.925	(0.039)	(0.277)	1.000	0.941	(0.011)	(0.121)	1.000	0.991	(0.003)	(0.067)
Model 2 + Model 3	MSE	1.032	0.957	-	-	0.688	0.530	-	-	0.301	0.251	-	-
	Rate	1.000	0.863	(0.136)	(0.209)	1.000	0.923	(0.034)	(0.091)	1.000	0.985	(0.016)	(0.044)
Model 2 + Model 4	MSE	1.005	0.833	-	-	0.605	0.497	-	-	0.331	0.244	-	-
	Rate	1.000	0.874	(0.130)	(0.291)	1.000	0.945	(0.031)	(0.139)	1.000	0.987	(0.005)	(0.078)
Model 3 + Model 4	MSE	0.843	0.860	-	-	0.512	0.490	-	-	0.346	0.268	-	-
	Rate	1.000	0.869	(0.142)	(0.226)	1.000	0.954	(0.047)	(0.116)	1.000	0.994	(0.008)	(0.047)
Panel 2		Combination: Clayton + Frank											
		$T = 200$				$T = 500$				$T = 1000$			
		Clayton	Gumbel	Frank	Gaussian	Clayton	Gumbel	Frank	Gaussian	Clayton	Gumbel	Frank	Gaussian
Model 1 + Model 2	MSE	1.036	-	0.925	-	0.633	-	0.447	-	0.380	-	0.251	-
	Rate	1.000	(0.015)	0.874	(0.183)	1.000	(0.003)	0.908	(0.098)	1.000	(0.000)	0.993	(0.043)
Model 1 + Model 3	MSE	1.047	-	0.859	-	0.629	-	0.517	-	0.291	-	0.266	-
	Rate	1.000	(0.022)	0.891	(0.194)	1.000	(0.006)	0.939	(0.090)	1.000	(0.000)	0.997	(0.036)
Model 1 + Model 4	MSE	0.931	-	0.906	-	0.424	-	0.601	-	0.230	-	0.325	-
	Rate	1.000	(0.017)	0.923	(0.207)	1.000	(0.000)	0.971	(0.107)	1.000	(0.000)	1.000	(0.029)
Model 2 + Model 3	MSE	0.935	-	0.923	-	0.591	-	0.459	-	0.372	-	0.245	-
	Rate	1.000	(0.021)	0.856	(0.251)	1.000	(0.000)	0.911	(0.137)	1.000	(0.000)	0.984	(0.061)
Model 2 + Model 4	MSE	0.817	-	0.852	-	0.456	-	0.427	-	0.296	-	0.289	-
	Rate	1.000	(0.014)	0.904	(0.191)	1.000	(0.003)	0.975	(0.112)	1.000	(0.000)	0.994	(0.054)
Model 3 + Model 4	MSE	0.831	-	0.874	-	0.533	-	0.512	-	0.230	-	0.222	-
	Rate	1.000	(0.009)	0.899	(0.233)	1.000	(0.000)	0.964	(0.128)	1.000	(0.000)	1.000	(0.072)
Panel 3		Combination: Gumbel + Frank											
		$T = 200$				$T = 500$				$T = 1000$			
		Clayton	Gumbel	Frank	Gaussian	Clayton	Gumbel	Frank	Gaussian	Clayton	Gumbel	Frank	Gaussian
Model 1 + Model 2	MSE	-	0.814	0.933	-	-	0.513	0.481	-	-	0.289	0.227	-
	Rate	(0.003)	0.975	0.874	(0.151)	(0.001)	0.987	0.957	(0.078)	(0.000)	0.998	0.983	(0.023)
Model 1 + Model 3	MSE	-	0.853	0.891	-	-	0.503	0.465	-	-	0.275	0.201	-
	Rate	(0.000)	0.982	0.901	(0.187)	(0.000)	0.986	0.989	(0.084)	(0.000)	0.993	0.981	(0.051)
Model 1 + Model 4	MSE	-	0.930	0.942	-	-	0.535	0.526	-	-	0.208	0.267	-
	Rate	(0.004)	0.974	0.887	(0.174)	(0.000)	0.983	0.946	(0.081)	(0.000)	1.000	0.992	(0.038)
Model 2 + Model 3	MSE	-	0.727	0.865	-	-	0.517	0.574	-	-	0.305	0.264	-
	Rate	(0.002)	0.986	0.939	(0.152)	(0.000)	0.977	0.951	(0.073)	(0.000)	0.997	0.998	(0.047)
Model 2 + Model 4	MSE	-	0.958	0.771	-	-	0.466	0.397	-	-	0.268	0.203	-
	Rate	(0.001)	0.967	0.853	(0.131)	(0.000)	0.981	0.976	(0.062)	(0.000)	1.000	0.993	(0.013)
Model 3 + Model 4	MSE	-	0.783	0.732	-	-	0.463	0.402	-	-	0.326	0.340	-
	Rate	(0.000)	0.958	0.944	(0.147)	(0.001)	0.967	0.983	(0.064)	(0.000)	0.987	0.987	(0.022)
Panel 4		True Copula: Clayton + Gaussian											
		$T = 200$				$T = 500$				$T = 1000$			
		Clayton	Gumbel	Frank	Gaussian	Clayton	Gumbel	Frank	Gaussian	Clayton	Gumbel	Frank	Gaussian
Model 1 + Model 2	MSE	0.707	-	-	0.005	0.411	-	-	0.003	0.231	-	-	0.001
	Rate	1.000	(0.011)	(0.193)	0.961	1.000	(0.004)	(0.081)	0.998	1.000	(0.000)	(0.042)	1.000
Model 1 + Model 3	MSE	0.735	-	-	0.007	0.398	-	-	0.004	0.250	-	-	0.003
	Rate	1.000	(0.005)	(0.152)	0.954	1.000	(0.001)	(0.078)	0.972	1.000	(0.000)	(0.039)	0.995
Model 1 + Model 4	MSE	0.852	-	-	0.008	0.501	-	-	0.004	0.222	-	-	0.001
	Rate	1.000	(0.011)	(0.183)	0.957	1.000	(0.000)	(0.075)	0.984	1.000	(0.000)	(0.037)	1.000
Model 2 + Model 3	MSE	0.753	-	-	0.006	0.443	-	-	0.004	0.205	-	-	0.002
	Rate	1.000	(0.016)	(0.179)	0.947	1.000	(0.001)	(0.083)	0.988	1.000	(0.000)	(0.041)	1.000
Model 2 + Model 4	MSE	0.849	-	-	0.005	0.529	-	-	0.003	0.230	-	-	0.001
	Rate	1.000	(0.021)	(0.182)	0.957	1.000	(0.003)	(0.086)	0.997	1.000	(0.000)	(0.047)	1.000
Model 3 + Model 4	MSE	0.831	-	-	0.006	0.477	-	-	0.003	0.209	-	-	0.001
	Rate	1.000	(0.008)	(0.176)	0.944	1.000	(0.000)	(0.081)	0.989	1.000	(0.000)	(0.039)	1.000
Panel 5		Combination: Gumbel + Gaussian											
		$T = 200$				$T = 500$				$T = 1000$			
		Clayton	Gumbel	Frank	Gaussian	Clayton	Gumbel	Frank	Gaussian	Clayton	Gumbel	Frank	Gaussian
Model 1 + Model 2	MSE	-	0.883	-	0.008	-	0.474	-	0.005	-	0.208	-	0.002
	Rate	(0.000)	0.931	(0.245)	0.964	(0.000)	0.974	(0.121)	1.000	(0.000)	0.998	(0.052)	1.000
Model 1 + Model 3	MSE	-	0.920	-	0.007	-	0.448	-	0.005	-	0.233	-	0.002
	Rate	(0.000)	0.945	(0.196)	0.947	(0.000)	0.961	(0.089)	0.995	(0.000)	0.994	(0.037)	1.000
Model 1 + Model 4	MSE	-	0.844	-	0.008	-	0.503	-	0.006	-	0.220	-	0.003
	Rate	(0.000)	0.927	(0.203)	0.953	(0.000)	0.979	(0.117)	0.989	(0.000)	0.987	(0.066)	0.994
Model 2 + Model 3	MSE	-	0.797	-	0.008	-	0.427	-	0.005	-	0.269	-	0.002
	Rate	(0.000)	0.958	(0.192)	0.961	(0.000)	0.978	(0.102)	0.993	(0.000)	0.993	(0.051)	1.000
Model 2 + Model 4	MSE	-	0.835	-	0.006	-	0.474	-	0.004	-	0.283	-	0.002
	Rate	(0.000)	0.944	(0.235)	0.955	(0.000)	0.980	(0.115)	0.987	(0.000)	0.998	(0.064)	1.000
Model 3 + Model 4	MSE	-	0.937	-	0.007	-	0.539	-	0.004	-	0.317	-	0.002
	Rate	(0.000)	0.936	(0.197)	0.941	(0.000)	0.979	(0.099)	0.979	(0.000)	0.987	(0.033)	0.998
Panel 6		Combination: Frank + Gaussian											
		$T = 200$				$T = 500$				$T = 1000$			
		Clayton	Gumbel	Frank	Gaussian	Clayton	Gumbel	Frank	Gaussian	Clayton	Gumbel	Frank	Gaussian
Model 1 + Model 2	MSE	-	-	0.879	0.011	-	-	0.512	0.008	-	-	0.274	0.003
	Rate	(0.000)	(0.011)	0.875	0.983	(0.000)	(0.000)	0.913	1.000	(0.000)	(0.000)	0.983	1.000
Model 1 + Model 3	MSE	-	-	0.742	0.009	-	-	0.442	0.006	-	-	0.260	0.002
	Rate	(0.000)	(0.013)	0.854	0.987	(0.000)	(0.000)	0.907	1.000	(0.000)	(0.000)	0.992	1.000
Model 1 + Model 4	MSE	-	-	0.833	0.013	-	-	0.435	0.009	-	-	0.279	0.005
	Rate	(0.000)	(0.015)	0.886	0.991	(0.000)	(0.000)	0.914	1.000	(0.000)	(0.000)	0.985	1.000
Model 2 + Model 3	MSE	-	-	0.858	0.008	-	-	0.408	0.005	-	-	0.251	0.003
	Rate	(0.000)	(0.012)	0.873	0.989	(0.000)	(0.000)	0.905	0.993	(0.000)	(0.000)	0.990	1.000
Model 2 + Model 4	MSE	-	-	0.903	0.011	-	-	0.511	0.006	-	-	0.281	0.003
	Rate	(0.000)	(0.011)	0.862	0.991	(0.000)	(0.000)	0.916	1.000	(0.000)	(0.000)	0.989	1.000
Model 3 + Model 4	MSE	-	-	0.944	0.013	-	-	0.553	0.007	-	-	0.302	0.003
	Rate	(0.000)	(0.015)	0.887	0.983	(0.000)	(0.000)	0.920	1.000	(0.000)	(0.000)	0.983	1.000

Table 3: MSEs of copula parameter estimates and accurate (inaccurate) rates of selection by Cai & Wangs (2014) constant mixture copula (CW) when the true model is a conditional mixture copula.

Panel 1		Combination: Clayton + Gumbel											
		T = 200				T = 500				T = 1000			
		Clayton	Gumbel	Frank	Gaussian	Clayton	Gumbel	Frank	Gaussian	Clayton	Gumbel	Frank	Gaussian
Model 1 + Model 2	MSE	4.836	3.129	-	-	3.166	2.646	-	-	2.401	1.480	-	-
	Rate	1.000	0.931	(0.174)	(0.254)	1.000	0.955	(0.106)	(0.183)	1.000	0.993	(0.033)	(0.104)
Model 1 + Model 3	MSE	3.991	3.528	-	-	2.597	2.445	-	-	1.931	1.558	-	-
	Rate	1.000	0.921	(0.112)	(0.287)	1.000	0.948	(0.054)	(0.113)	1.000	0.984	(0.013)	(0.065)
Model 1 + Model 4	MSE	3.821	3.681	-	-	2.638	2.353	-	-	1.514	1.231	-	-
	Rate	1.000	0.930	(0.104)	(0.296)	1.000	0.964	(0.035)	(0.152)	1.000	1.000	(0.007)	(0.077)
Model 2 + Model 3	MSE	4.658	4.436	-	-	2.701	2.411	-	-	1.512	1.382	-	-
	Rate	1.000	0.916	(0.158)	(0.253)	1.000	0.958	(0.049)	(0.093)	1.000	0.989	(0.014)	(0.039)
Model 2 + Model 4	MSE	4.819	4.627	-	-	2.706	2.568	-	-	1.705	1.588	-	-
	Rate	1.000	0.948	(0.129)	(0.301)	1.000	0.962	(0.034)	(0.154)	1.000	1.000	(0.015)	(0.088)
Model 3 + Model 4	MSE	3.158	3.687	-	-	2.546	2.610	-	-	1.447	1.658	-	-
	Rate	1.000	0.905	(0.157)	(0.258)	1.000	0.932	(0.079)	(0.128)	1.000	1.000	(0.019)	(0.041)
Panel 2		Combination: Clayton + Frank											
		T = 200				T = 500				T = 1000			
		Clayton	Gumbel	Frank	Gaussian	Clayton	Gumbel	Frank	Gaussian	Clayton	Gumbel	Frank	Gaussian
Model 1 + Model 2	MSE	3.769	-	3.197	-	2.477	-	2.653	-	1.296	-	1.424	-
	Rate	1.000	(0.023)	0.931	(0.205)	1.000	(0.005)	0.923	(0.134)	1.000	(0.000)	1.000	(0.072)
Model 1 + Model 3	MSE	3.825	-	3.003	-	2.402	-	1.847	-	1.838	-	1.097	-
	Rate	1.000	(0.025)	0.925	(0.237)	1.000	(0.000)	0.944	(0.118)	1.000	(0.000)	1.000	(0.063)
Model 1 + Model 4	MSE	3.770	-	3.568	-	2.581	-	2.419	-	1.403	-	1.370	-
	Rate	1.000	(0.020)	0.895	(0.262)	1.000	(0.003)	0.946	(0.120)	1.000	(0.000)	1.000	(0.055)
Model 2 + Model 3	MSE	3.704	-	3.239	-	2.688	-	2.172	-	1.522	-	1.165	-
	Rate	1.000	(0.018)	0.907	(0.289)	1.000	(0.001)	0.957	(0.154)	1.000	(0.000)	1.000	(0.069)
Model 2 + Model 4	MSE	2.911	-	2.941	-	2.249	-	2.447	-	1.239	-	1.384	-
	Rate	1.000	(0.015)	0.922	(0.264)	1.000	(0.003)	0.955	(0.116)	1.000	(0.000)	1.000	(0.048)
Model 3 + Model 4	MSE	3.335	-	3.569	-	2.137	-	2.398	-	1.393	-	1.371	-
	Rate	1.000	(0.013)	0.873	(0.262)	1.000	(0.000)	0.931	(0.173)	1.000	(0.000)	0.998	(0.087)
Panel 3		Combination: Gumbel + Frank											
		T = 200				T = 500				T = 1000			
		Clayton	Gumbel	Frank	Gaussian	Clayton	Gumbel	Frank	Gaussian	Clayton	Gumbel	Frank	Gaussian
Model 1 + Model 2	MSE	-	3.584	3.276	-	2.473	2.443	-	-	1.397	1.266	-	-
	Rate	(0.005)	0.985	0.903	(0.187)	(0.001)	1.000	0.992	(0.085)	(0.000)	1.000	0.995	(0.031)
Model 1 + Model 3	MSE	-	3.747	3.255	-	-	2.465	2.677	-	-	1.378	1.523	-
	Rate	(0.000)	0.957	0.942	(0.195)	(0.000)	0.994	1.000	(0.082)	(0.000)	1.000	1.000	(0.050)
Model 1 + Model 4	MSE	-	3.726	3.580	-	-	2.461	2.062	-	-	1.385	1.037	-
	Rate	(0.005)	0.989	0.896	(0.193)	(0.000)	0.990	0.977	(0.088)	(0.000)	1.000	1.000	(0.045)
Model 2 + Model 3	MSE	-	3.105	3.175	-	-	2.562	2.403	-	-	1.502	1.143	-
	Rate	(0.001)	0.971	0.944	(0.188)	(0.000)	0.991	0.984	(0.073)	(0.000)	1.000	1.000	(0.017)
Model 2 + Model 4	MSE	-	3.716	3.528	-	-	2.527	2.508	-	-	1.468	1.395	-
	Rate	(0.000)	0.963	0.907	(0.172)	(0.000)	1.000	0.997	(0.089)	(0.000)	1.000	1.000	(0.013)
Model 3 + Model 4	MSE	-	3.454	3.721	-	-	2.409	2.116	-	-	1.443	1.222	-
	Rate	(0.002)	0.970	0.905	(0.133)	(0.001)	0.987	1.000	(0.061)	(0.000)	0.998	1.000	(0.024)
Panel 4		Combination: Clayton + Gaussian											
		T = 200				T = 500				T = 1000			
		Clayton	Gumbel	Frank	Gaussian	Clayton	Gumbel	Frank	Gaussian	Clayton	Gumbel	Frank	Gaussian
Model 1 + Model 2	MSE	3.894	-	-	0.027	2.502	-	-	0.016	1.299	-	-	0.008
	Rate	1.000	(0.017)	(0.227)	0.957	1.000	(0.008)	(0.103)	1.000	1.000	(0.000)	(0.034)	1.000
Model 1 + Model 3	MSE	3.780	-	-	0.032	3.495	-	-	0.019	3.306	-	-	0.009
	Rate	1.000	(0.004)	(0.204)	0.974	1.000	(0.000)	(0.094)	0.986	1.000	(0.000)	(0.062)	1.000
Model 1 + Model 4	MSE	3.827	-	-	0.044	3.553	-	-	0.022	3.269	-	-	0.008
	Rate	1.000	(0.016)	(0.176)	0.977	1.000	(0.000)	(0.068)	0.982	1.000	(0.000)	(0.049)	1.000
Model 2 + Model 3	MSE	2.820	-	-	0.023	1.849	-	-	0.013	1.031	-	-	0.006
	Rate	1.000	(0.015)	(0.213)	0.965	1.000	(0.000)	(0.091)	0.975	1.000	(0.000)	(0.063)	1.000
Model 2 + Model 4	MSE	2.917	-	-	0.023	2.041	-	-	0.011	1.665	-	-	0.005
	Rate	1.000	(0.029)	(0.204)	0.973	1.000	(0.005)	(0.085)	0.984	1.000	(0.000)	(0.059)	0.997
Model 3 + Model 4	MSE	3.606	-	-	0.035	2.563	-	-	0.012	1.717	-	-	0.004
	Rate	1.000	(0.005)	(0.198)	0.968	1.000	(0.002)	(0.094)	0.970	1.000	(0.000)	(0.045)	1.000
Panel 5		Combination: Gumbel + Gaussian											
		T = 200				T = 500				T = 1000			
		Clayton	Gumbel	Frank	Gaussian	Clayton	Gumbel	Frank	Gaussian	Clayton	Gumbel	Frank	Gaussian
Model 1 + Model 2	MSE	-	3.548	-	0.048	-	2.517	-	0.027	-	1.422	-	0.008
	Rate	(0.000)	0.944	(0.278)	0.974	(0.000)	0.988	(0.133)	1.000	(0.000)	1.000	(0.048)	1.000
Model 1 + Model 3	MSE	-	3.569	-	0.033	-	2.562	-	0.020	-	1.835	-	0.006
	Rate	(0.000)	0.941	(0.212)	0.931	(0.000)	0.955	(0.120)	1.000	(0.000)	1.000	(0.046)	1.000
Model 1 + Model 4	MSE	-	3.730	-	0.044	-	2.377	-	0.032	-	1.348	-	0.009
	Rate	(0.000)	0.908	(0.174)	0.948	(0.000)	0.976	(0.119)	1.000	(0.000)	0.997	(0.057)	1.000
Model 2 + Model 3	MSE	-	3.026	-	0.042	-	2.220	-	0.019	-	1.534	-	0.009
	Rate	(0.000)	0.923	(0.228)	0.977	(0.000)	0.983	(0.125)	1.000	(0.000)	1.000	(0.061)	1.000
Model 2 + Model 4	MSE	-	4.082	-	0.024	-	2.626	-	0.011	-	1.488	-	0.007
	Rate	(0.000)	0.911	(0.293)	0.934	(0.000)	0.987	(0.127)	0.998	(0.000)	1.000	(0.073)	1.000
Model 3 + Model 4	MSE	-	3.430	-	0.035	-	2.435	-	0.017	-	1.415	-	0.010
	Rate	(0.000)	0.950	(0.204)	0.950	(0.000)	0.959	(0.110)	1.000	(0.000)	1.000	(0.041)	1.000
Panel 6		Combination: Frank + Gaussian											
		T = 200				T = 500				T = 1000			
		Clayton	Gumbel	Frank	Gaussian	Clayton	Gumbel	Frank	Gaussian	Clayton	Gumbel	Frank	Gaussian
Model 1 + Model 2	MSE	-	-	3.625	0.059	-	-	2.382	0.036	-	-	1.217	0.015
	Rate	(0.000)	(0.016)	0.891	0.974	(0.000)	(0.000)	0.932	1.000	(0.000)	(0.000)	0.998	1.000
Model 1 + Model 3	MSE	-	-	3.770	0.043	-	-	2.587	0.030	-	-	1.226	0.008
	Rate	(0.000)	(0.014)	0.863	0.954	(0.000)	(0.000)	0.894	1.000	(0.000)	(0.000)	1.000	1.000
Model 1 + Model 4	MSE	-	-	3.445	0.065	-	-	2.282	0.042	-	-	1.102	0.020
	Rate	(0.000)	(0.010)	0.895	0.987	(0.000)	(0.000)	0.907	1.000	(0.000)	(0.000)	0.991	1.000
Model 2 + Model 3	MSE	-	-	4.311	0.043	-	-	2.730	0.029	-	-	1.755	0.008
	Rate	(0.000)	(0.017)	0.914	0.980	(0.000)	(0.000)	0.913	1.000	(0.000)	(0.000)	0.986	1.000
Model 2 + Model 4	MSE	-	-	4.430	0.054	-	-	2.993	0.030	-	-	1.944	0.010
	Rate	(0.000)	(0.014)	0.882	0.976	(0.000)	(0.000)	0.896	1.000	(0.000)	(0.000)	0.993	1.000
Model 3 + Model 4	MSE	-	-	3.709	0.064	-	-	2.443	0.031	-	-	1.675	0.009
	Rate	(0.000)	(0.012)	0.854	0.963	(0.000)	(0.000)	0.919	1.000	(0.000)	(0.000)	0.980	1.000

Table 4: MSEs of weight estimates by the proposed conditional mixture copula (CM) and Cai & Wang's (2014) constant mixture copula (CW) when the true model is a conditional mixture copula.

		True Copula: Clayton + Gumbel																							
		$T = 200$						$T = 500$						$T = 1000$											
		Clayton		Gumbel		Frank		Gaussian		Clayton		Gumbel		Frank		Gaussian		Clayton		Gumbel		Frank		Gaussian	
Model 1 + Model 2	0.009	0.085	0.010	0.091	-	-	0.005	0.056	0.009	0.057	-	-	-	-	0.002	0.031	0.003	0.035	-	-	-	-	-	-	
Model 1 + Model 3	0.010	0.087	0.011	0.084	-	-	0.006	0.048	0.008	0.050	-	-	-	-	0.002	0.024	0.001	0.021	-	-	-	-	-	-	
Model 1 + Model 4	0.011	0.091	0.012	0.086	-	-	0.004	0.055	0.009	0.049	-	-	-	-	0.001	0.029	0.003	0.026	-	-	-	-	-	-	
Model 2 + Model 3	0.009	0.078	0.007	0.083	-	-	0.004	0.048	0.005	0.041	-	-	-	-	0.001	0.023	0.002	0.019	-	-	-	-	-	-	
Model 2 + Model 4	0.013	0.084	0.010	0.078	-	-	0.006	0.040	0.004	0.047	-	-	-	-	0.002	0.026	0.001	0.021	-	-	-	-	-	-	
Model 3 + Model 4	0.011	0.086	0.013	0.087	-	-	0.006	0.041	0.007	0.055	-	-	-	-	0.001	0.019	0.002	0.022	-	-	-	-	-	-	
Panel 1		True Copula: Clayton + Frank																							
		$T = 200$						$T = 500$						$T = 1000$											
		Clayton		Gumbel		Frank		Gaussian		Clayton		Gumbel		Frank		Gaussian		Clayton		Gumbel		Frank		Gaussian	
		CM	CW	CM	CW	CM	CW	CM	CW	CM	CW	CM	CW	CM	CW	CM	CW	CM	CW	CM	CW	CM	CW	CM	CW
Model 1 + Model 2	0.013	0.112	-	-	0.009	0.107	-	-	0.007	0.071	-	-	0.004	0.066	-	-	0.002	0.048	-	-	0.002	0.034	-	-	
Model 1 + Model 3	0.012	0.088	-	-	0.008	0.116	-	-	0.007	0.052	-	-	0.003	0.063	-	-	0.003	0.039	-	-	0.002	0.040	-	-	
Model 1 + Model 4	0.012	0.093	-	-	0.010	0.102	-	-	0.006	0.058	-	-	0.004	0.059	-	-	0.004	0.037	-	-	0.001	0.031	-	-	
Model 2 + Model 3	0.010	0.091	-	-	0.011	0.113	-	-	0.003	0.057	-	-	0.006	0.053	-	-	0.001	0.030	-	-	0.004	0.027	-	-	
Model 2 + Model 4	0.012	0.102	-	-	0.009	0.098	-	-	0.005	0.054	-	-	0.004	0.061	-	-	0.003	0.026	-	-	0.002	0.035	-	-	
Model 3 + Model 4	0.010	0.118	-	-	0.009	0.086	-	-	0.003	0.062	-	-	0.005	0.050	-	-	0.002	0.032	-	-	0.002	0.028	-	-	
Panel 2		True Copula: Gumbel + Frank																							
		$T = 200$						$T = 500$						$T = 1000$											
		Clayton		Gumbel		Frank		Gaussian		Clayton		Gumbel		Frank		Gaussian		Clayton		Gumbel		Frank		Gaussian	
		CM	CW	CM	CW	CM	CW	CM	CW	CM	CW	CM	CW	CM	CW	CM	CW	CM	CW	CM	CW	CM	CW	CM	CW
Model 1 + Model 2	-	-	0.011	0.104	0.009	0.117	-	-	0.006	0.063	0.004	0.064	-	-	0.007	0.066	0.005	0.062	-	-	0.003	0.038	0.003	0.036	
Model 1 + Model 3	-	-	0.011	0.105	0.008	0.106	-	-	0.007	0.057	0.005	0.053	-	-	0.007	0.057	0.005	0.053	-	-	0.002	0.035	0.002	0.030	
Model 1 + Model 4	-	-	0.014	0.097	0.009	0.093	-	-	0.008	0.069	0.003	0.051	-	-	0.008	0.069	0.003	0.051	-	-	0.003	0.034	0.002	0.039	
Model 2 + Model 3	-	-	0.015	0.099	0.008	0.094	-	-	0.005	0.051	-	-	-	-	0.005	0.045	0.005	0.050	-	-	0.001	0.031	0.003	0.032	
Model 2 + Model 4	-	-	0.009	0.081	0.011	0.109	-	-	0.005	0.045	0.006	0.047	-	-	0.005	0.045	0.006	0.047	-	-	0.001	0.031	0.003	0.032	
Model 3 + Model 4	-	-	0.010	0.093	0.011	0.104	-	-	0.005	0.045	0.006	0.047	-	-	0.005	0.045	0.006	0.047	-	-	0.002	0.035	0.004	0.030	
Panel 3		True Copula: Clayton + Gaussian																							
		$T = 200$						$T = 500$						$T = 1000$											
		Clayton		Gumbel		Frank		Gaussian		Clayton		Gumbel		Frank		Gaussian		Clayton		Gumbel		Frank		Gaussian	
		CM	CW	CM	CW	CM	CW	CM	CW	CM	CW	CM	CW	CM	CW	CM	CW	CM	CW	CM	CW	CM	CW	CM	CW
Model 1 + Model 2	0.014	0.106	-	-	0.013	0.098	0.006	0.063	-	-	0.006	0.059	-	-	0.003	0.032	-	-	0.003	0.036	-	-	0.003	0.036	
Model 1 + Model 3	0.013	0.097	-	-	0.011	0.072	0.007	0.057	-	-	0.008	0.044	0.004	0.029	-	-	0.004	0.029	-	-	0.005	0.037	-	-	
Model 1 + Model 4	0.011	0.103	-	-	0.012	0.069	0.006	0.056	-	-	0.007	0.042	0.002	0.032	-	-	0.002	0.032	-	-	0.004	0.031	-	-	
Model 2 + Model 3	0.013	0.094	-	-	0.015	0.083	0.005	0.051	-	-	0.007	0.041	0.003	0.027	-	-	0.003	0.027	-	-	0.005	0.033	-	-	
Model 2 + Model 4	0.009	0.095	-	-	0.014	0.102	0.005	0.053	-	-	0.008	0.052	0.003	0.029	-	-	0.003	0.029	-	-	0.004	0.035	-	-	
Model 3 + Model 4	0.012	0.102	-	-	0.010	0.095	0.007	0.052	-	-	0.004	0.045	0.004	0.033	-	-	0.004	0.033	-	-	0.003	0.034	-	-	
Panel 4		True Copula: Gumbel + Gaussian																							
		$T = 200$						$T = 500$						$T = 1000$											
		Clayton		Gumbel		Frank		Gaussian		Clayton		Gumbel		Frank		Gaussian		Clayton		Gumbel		Frank		Gaussian	
		CM	CW	CM	CW	CM	CW	CM	CW	CM	CW	CM	CW	CM	CW	CM	CW	CM	CW	CM	CW	CM	CW	CM	CW
Model 1 + Model 2	-	-	0.007	0.093	-	-	0.008	0.069	-	-	0.004	0.049	-	-	0.004	0.031	-	-	0.002	0.032	-	-	0.002	0.017	
Model 1 + Model 3	-	-	0.008	0.102	-	-	0.009	0.083	-	-	0.005	0.058	-	-	0.005	0.047	-	-	0.002	0.031	-	-	0.003	0.029	
Model 1 + Model 4	-	-	0.008	0.078	-	-	0.008	0.090	-	-	0.004	0.044	-	-	0.004	0.047	-	-	0.002	0.029	-	-	0.002	0.025	
Model 2 + Model 3	-	-	0.012	0.089	-	-	0.010	0.102	-	-	0.008	0.047	-	-	0.006	0.055	-	-	0.003	0.027	-	-	0.004	0.031	
Model 2 + Model 4	-	-	0.009	0.083	-	-	0.009	0.091	-	-	0.005	0.047	-	-	0.004	0.051	-	-	0.002	0.028	-	-	0.002	0.027	
Model 3 + Model 4	-	-	0.010	0.094	-	-	0.011	0.097	-	-	0.006	0.051	-	-	0.006	0.055	-	-	0.003	0.025	-	-	0.004	0.026	
Panel 5		True Copula: Frank + Gaussian																							
		$T = 200$						$T = 500$						$T = 1000$											
		Clayton		Gumbel		Frank		Gaussian		Clayton		Gumbel		Frank		Gaussian		Clayton		Gumbel		Frank		Gaussian	
		CM	CW	CM	CW	CM	CW	CM	CW	CM	CW	CM	CW	CM	CW	CM	CW	CM	CW	CM	CW	CM	CW	CM	CW
Model 1 + Model 2	-	-	0.011	0.113	0.009	0.106	-	-	0.010	0.105	0.008	0.101	-	-	0.007	0.060	0.005	0.052	-	-	0.003	0.035	0.003	0.031	
Model 1 + Model 3	-	-	0.010	0.105	0.014	0.085	-	-	0.006	0.057	0.007	0.049	-	-	0.006	0.057	0.007	0.049	-	-	0.003	0.032	0.003	0.026	
Model 1 + Model 4	-	-	0.011	0.102	0.008	0.101	-	-	0.006	0.056	0.006	0.050	-	-	0.006	0.056	0.005	0.050	-	-	0.004	0.031	0.004	0.027	
Model 2 + Model 3	-	-	0.013	0.098	0.013	0.092	-	-	0.008	0.055	0.006	0.053	-	-	0.008	0.055	0.006	0.053	-	-	0.005	0.036	0.003	0.033	
Model 2 + Model 4	-	-	0.009	0.097	0.012	0.089	-	-	0.004	0.053	0.006	0.045	-	-	0.004	0.053	0.006	0.045	-	-	0.002	0.035	0.003	0.028	
Model 3 + Model 4	-	-	0.012	0.102	0.012	0.091	-	-	0.007	0.052	0.007	0.045	-	-	0.007	0.053	0.008	0.041	-	-	0.004	0.029	0.006	0.022	
Panel 6		True Copula: Frank + Gaussian																							
		$T = 200$						$T = 500$						$T = 1000$											
		Clayton		Gumbel		Frank		Gaussian		Clayton		Gumbel		Frank		Gaussian		Clayton		Gumbel		Frank		Gaussian	
		CM	CW	CM	CW	CM	CW	CM	CW	CM	CW	CM	CW	CM	CW	CM	CW	CM	CW	CM	CW	CM	CW	CM	CW
Model 1 + Model 2	-	-	0.011	0.113	0.009	0.106	-	-	0.007	0.052	0.007	0.045	-	-	0.007	0.053	0.008	0.041	-	-	0.004	0.029	0.006	0.022	
Model 1 + Model 3	-	-	0.010	0.105	0.014	0.085	-	-	0.006	0.057	0.006	0.050	-	-	0.006	0.057	0.007	0.049	-	-	0.003	0.032	0.003	0.026	
Model 1 + Model 4	-	-	0.011	0.102	0.008	0.101	-	-	0.006	0.056	0.006	0.050	-	-	0.006	0.056	0.005	0.050	-	-	0.004	0.031	0.004	0.027	
Model 2 + Model 3	-	-	0.013	0.098	0.013	0.092	-	-	0.008	0.055	0.006	0.053	-	-	0.008	0.055	0.006	0.053	-	-	0.005	0.036	0.003	0.033	
Model 2 + Model 4	-	-	0.009	0.097	0.012	0.089	-	-																	

Table 5: MSEs of copula parameter estimates and accurate (inaccurate) rates of selection by the proposed conditional mixture copula (CM) and Cai & Wang’s (2014) constant mixture copula (CW) when the true model is an individual constant copula (Panel 1) and a constant mixture copula (Panel 2).

Panel 1		True Copula: Clayton							
		$T = 200$							
		Clayton		Gumbel		Frank		Gaussian	
		CM	CW	CM	CW	CM	CW	CM	CW
$\theta_{Cl} = 5$	MSE	0.173	0.121	–	–	–	–	–	–
	Rate	1.000	1.000	(0.000)	(0.000)	(0.000)	(0.000)	(0.013)	(0.000)
$\theta_{Cl} = 7$	MSE	0.135	0.119	–	–	–	–	–	–
	Rate	1.000	1.000	(0.000)	(0.000)	(0.000)	(0.000)	(0.004)	(0.000)
		$T = 500$							
		Clayton		Gumbel		Frank		Gaussian	
		CM	CW	CM	CW	CM	CW	CM	CW
$\theta_{Cl} = 5$	MSE	0.104	0.075	–	–	–	–	–	–
	Rate	1.000	1.000	(0.000)	(0.000)	(0.000)	(0.000)	(0.002)	(0.000)
$\theta_{Cl} = 7$	MSE	0.071	0.062	–	–	–	–	–	–
	Rate	1.000	1.000	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
		$T = 1000$							
		Clayton		Gumbel		Frank		Gaussian	
		CM	CW	CM	CW	CM	CW	CM	CW
$\theta_{Cl} = 5$	MSE	0.047	0.033	–	–	–	–	–	–
	Rate	1.000	1.000	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
$\theta_{Cl} = 7$	MSE	0.034	0.029	–	–	–	–	–	–
	Rate	1.000	1.000	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Panel 2		True Copula: Clayton + Gumbel (equally weighted)							
		$T = 200$							
		Clayton		Gumbel		Frank		Gaussian	
		CM	CW	CM	CW	CM	CW	CM	CW
$\theta_{Cl} = 5, \theta_{Gu} = 4$	MSE	0.253	0.194	0.367	0.231	–	–	–	–
	Rate	1.000	1.000	0.859	0.831	(0.118)	(0.125)	(0.236)	(0.221)
$\theta_{Cl} = 7, \theta_{Gu} = 6$	MSE	0.385	0.253	0.513	0.277	–	–	–	–
	Rate	1.000	1.000	0.893	0.865	(0.127)	(0.167)	(0.262)	(0.258)
		$T = 500$							
		Clayton		Gumbel		Frank		Gaussian	
		CM	CW	CM	CW	CM	CW	CM	CW
$\theta_{Cl} = 5, \theta_{Gu} = 4$	MSE	0.172	0.110	0.218	0.138	–	–	–	–
	Rate	1.000	1.000	0.941	0.922	(0.053)	(0.051)	(0.115)	(0.107)
$\theta_{Cl} = 7, \theta_{Gu} = 6$	MSE	0.243	0.183	0.324	0.206	–	–	–	–
	Rate	1.000	1.000	0.949	0.958	(0.044)	(0.049)	(0.127)	(0.139)
		$T = 1000$							
		Clayton		Gumbel		Frank		Gaussian	
		CM	CW	CM	CW	CM	CW	CM	CW
$\theta_{Cl} = 5, \theta_{Gu} = 4$	MSE	0.106	0.071	0.133	0.079	–	–	–	–
	Rate	1.000	1.000	0.981	0.991	(0.009)	(0.006)	(0.035)	(0.022)
$\theta_{Cl} = 7, \theta_{Gu} = 6$	MSE	0.166	0.105	0.218	0.133	–	–	–	–
	Rate	1.000	1.000	0.983	0.977	(0.005)	(0.002)	(0.030)	(0.031)

Table 6: Summary Statistics

	FR	DE	US	UK	USD-EUR	USD-GBP
Panel 1: Summary statistics						
Mean (%)	0.037	0.031	0.078	-0.007	0.003	0.023
Median (%)	0.266	0.282	0.244	0.155	0.039	-0.038
Min (%)	-17.581	-17.504	-16.748	-15.220	-9.010	-5.547
Max (%)	12.829	13.977	10.344	10.915	6.085	10.222
Std. Dev	0.033	0.035	0.023	0.028	0.013	0.013
Skewness	-0.514	-0.698	-0.678	-0.521	-0.256	0.704
Kurtosis	2.916	3.081	5.043	3.079	2.788	5.377
<i>JB</i>	415***	497***	1184***	459***	349***	1341***
Panel 2: Linear correlation coefficients						
	DE	US	UK			
FR	0.932	0.740	0.878			
DE	–	0.740	0.830			
US	–	–	0.735			

Notes. Panel 1 documents the summary statistics of the weekly log returns of the MSCI equity prices in France, Germany, the United States and the United Kingdom, and the weekly log returns of the U.S. dollar - euro and U.S. dollar - British pound exchange rates. *JB* denotes the statistic of the Jarque-Bera test with the null hypothesis of normality. *** indicates rejection of the null at 1%. Panel 2 documents the linear correlation coefficients among the four international equity markets' weekly log returns. The sample period are between 01/07/1999 and 11/07/2018.

Table 7: The estimates of AR(1)-GARCH(1,1)

	AR(1) Part	GARCH(1,1) Part			<i>LB</i>
	γ_1 (<i>p</i> -value)	α_0 (<i>p</i> -value)	α_1 (<i>p</i> -value)	α_2 (<i>p</i> -value)	
Panel 1: Equity Return					
France	-0.119*** (0.000)	0.000*** (0.000)	0.805*** (0.000)	0.161*** (0.000)	0.755 (0.385)
Germany	-0.079** (0.018)	0.000*** (0.006)	0.836*** (0.000)	0.129*** (0.000)	0.127 (0.722)
U.S.	-0.122*** (0.000)	0.000*** (0.002)	0.795*** (0.000)	0.169*** (0.000)	1.763 (0.184)
UK	-0.088*** (0.009)	0.000*** (0.000)	0.755*** (0.000)	0.181*** (0.000)	1.716 (0.190)
Panel 2: Exchange Rate					
USD-EUR	0.043 (0.176)	0.000** (0.039)	0.898*** (0.000)	0.086*** (0.000)	0.029 (0.864)
USD-GBP	0.013 (0.706)	0.000*** (0.008)	0.799*** (0.000)	0.126*** (0.000)	2.160 (0.142)

Notes. This table summarizes the results of the AR(1)-GARCH(1,1) filtering. *LB* denotes the statistic of the Ljung-Box test with the null of zero autocorrelation for the residuals filtered by AR(1)-GARCH(1,1). Values in parentheses are corresponding *p*-values. ** and *** respectively indicates rejection of the null at 5% and 1%.

Table 8: BIC values for different copula models

	U.S.-Germany	U.S.-France	U.S.-UK
Clayton	-11800.282	-11508.044	-11789.353
Gumbel	-11043.538	-10738.931	-11056.510
Frank	-10747.739	-10763.693	-10789.578
Clayton+Gumbel	-12692.052	-12448.875	-12705.331
Clayton+Frank	-12940.750	-12685.836	-12965.381
Gumbel+Frank	-12579.391	-12356.165	-12590.426
Clayton+Gumbel+Frank	-11743.994	-11542.665	-11701.498

Notes. This table reports the BIC values of seven individual and mixture copula models for the pairs of U.S.-Germany, U.S.-France and U.S.-UK. The best models are in bold.

Table 9: p -values of the hypothesis tests

	U.S. - France	U.S. - Germany	U.S. - UK
	p -value	p -value	p -value
Panel 1: Mixture model = $\omega_{Cl}C_{Cl} + \omega_{Gu}C_{Gu} + \omega_{Fr}C_{Fr}$			
$H_0 : \omega_{Cl} = 0$ v.s. $H_1 : \omega_{Cl} \neq 0$	0.000	0.000	0.000
$H_0 : \omega_{Gu} = 0$ v.s. $H_1 : \omega_{Gu} \neq 0$	0.217	0.335	0.241
$H_0 : \omega_{Fr} = 0$ v.s. $H_1 : \omega_{Fr} \neq 0$	0.000	0.000	0.000
Panel 2: Mixture model = $\omega_{Cl}C_{Cl} + \omega_{Fr}C_{Fr}$			
$H_0 : \omega_{Cl} = 0$ v.s. $H_1 : \omega_{Cl} \neq 0$	0.000	0.000	0.000
$H_0 : \omega_{Fr} = 0$ v.s. $H_1 : \omega_{Fr} \neq 0$	0.000	0.000	0.000

Notes. This table displays the p -values for the estimates of the weight parameters in a sequence of hypothesis tests with the 0.05 significance level. In Panel 1, the mixture model is assumed to contain all three component copulas and three hypothesis tests are respectively implemented. The component copula with the highest p -value among those greater than 0.05 is filtered out. In Panel 2, we refit the mixture model with two component copulas and implement two hypothesis tests. We will exclude the copula whose weight estimate's p -value is greater than 0.05. Otherwise we will keep both component copulas.

Appendix

Appendix A describes a stationary bootstrap technique which is adopted in Sections 2.4 of the main text, Appendix B discusses some practical issues including an EM algorithm, the selection of the bandwidth, and confidence intervals, Appendix C shows the proofs of the key results, and Appendix D presents additional simulation results.

F The Stationary Bootstrap Resampling Scheme

Define $y_t \equiv (\mathbf{x}_t^\top, z_t)$, suppose $\{y_t\}_{t=1}^T$ is a strictly stationary and weakly dependent time series. Let

$$B_{t,b} = \{y_t, y_{t+1}, \dots, y_{t+b-1}\}$$

be the block consisting of b observations starting from y_t to y_{t+b-1} . When $j > T$, y_j is defined to be y_t , where $t = j \pmod{T}$ and $y_0 = y_T$. Let p be a constant such that $p \in [0, 1]$. Independent of $\{y_t\}_{t=1}^T$, let L_1, L_2, \dots be a sequence of i.i.d random variables which have the geometric distribution, i.e.,

$$P\{L_k = m\} = (1 - p)^{m-1}p, \quad m = 1, 2, \dots,$$

where $p = T^{-1/3}$. Independent of $\{y_t\}_{t=1}^T$ and L_k , let I_1, I_2, \dots be a sequence of i.i.d random variables which have the discrete uniform distribution on $\{1, \dots, T\}$.

A pseudo time series $\{y_t^*\}_{t=1}^T$ is generated in the following way. Sampling a sequence of blocks of random length by the prescription $B_{I_1, L_1}, B_{I_2, L_2}, \dots$, where I_k is generated from a uniform distribution on $\{1, \dots, T\}$ and L_k is generated from the distribution as defined earlier. The first L_1 observations in the pseudo time series $\{y_t^*\}_{t=1}^T$ are determined by the first block B_{I_1, L_1} of observations $y_{I_1}, \dots, y_{I_1+L_1-1}$, the next L_2 observations in the pseudo time series are the observations in the second sampled block B_{I_2, L_2} , namely $y_{I_2}, \dots, y_{I_2+L_2-1}$. This process is not stopped until T observations in the pseudo time series have been generated.

By randomly varying the block length, Politis and Romano (1994) show that the pseudo time series $\{y_t^*\}_{t=1}^T$, conditional on the original data $\{y_t\}_{t=1}^T$, is actually stationary. Hence, this resampling method is applicable for stationary and weakly dependent time series.

G Practical Issues

G.1 An EM Algorithm

An important computational issue in practice is that the maximum likelihood estimator of the local log-likelihood function (4) in the main text may not have an explicit expression, especially when d , the number of candidate copulas, is large. Therefore, we propose to use the expectation maximization (EM) algorithm (see Dempster *et al.*, 1977) to find the numerical solution of the maximum likelihood estimation. The EM algorithm decomposes the optimization process into two steps: for each $k \in \{1, \dots, d\}$, the E-step computes and updates α_{k0} , with given α_{k1} and $\beta_k = (\beta_{k0}, \beta_{k1})^\top$, and the M-step maximizes the local log-likelihood function to estimate α_{k1} and β_k .

Specifically, we define

$$c_c(\hat{u}_t, \delta) = \sum_{k=1}^d g_{\omega,k}^{-1}(\mathbf{z}_{t,z}^\top \alpha_k) c_k[\hat{u}_t; g_{\theta,k}^{-1}(\mathbf{z}_{t,z}^\top \beta_k)]$$

and the objective function

$$Q(\hat{u}, \delta) = \sum_{t=1}^T \ln c_c(\hat{u}_t, \delta) K_h(z_t - z) + \rho \left(1 - \sum_{k=1}^d g_{\omega,k}^{-1}(\alpha_{k0}) \right), \quad (5)$$

where $\mathbf{z}_{t,z} = (1, z_t - z)^\top$, $\alpha_k = (\alpha_{k0}, \alpha_{k1})^\top$, and the last term in equation (5) is for the constraint on the weight parameters.

To maximize the objective function $Q(\hat{u}, \delta)$, we take the first derivative of $Q(\hat{u}, \delta)$ with respect to α_{k0} and set it equal to zero:

$$\sum_{t=1}^T \frac{c_k[\hat{u}_t; g_{\theta,k}^{-1}(\mathbf{z}_{t,z}^\top \beta_k)]}{c_c(\hat{u}_t, \delta)} (g_{\omega,k}^{-1})'(\mathbf{z}_{t,z}^\top \alpha_k) K_h(z_t - z) - \rho (g_{\omega,k}^{-1})'(\alpha_{k0}) = 0.$$

Therefore, we have

$$\sum_{t=1}^T \frac{c_k[\hat{u}_t; g_{\theta,k}^{-1}(\mathbf{z}_{t,z}^\top \beta_k)]}{c_c(\hat{u}_t, \delta) (g_{\omega,k}^{-1})'(\alpha_{k0})} (g_{\omega,k}^{-1})'(\mathbf{z}_{t,z}^\top \alpha_k) K_h(z_t - z) = \rho. \quad (6)$$

By multiplying $g_{\omega,k}^{-1}(\alpha_{k0})$ on both sides of equation (6), we obtain

$$\sum_{t=1}^T \frac{c_k[\hat{u}_t; g_{\theta,k}^{-1}(\mathbf{z}_{t,z}^\top \beta_k)]}{c_c(\hat{u}_t, \delta)(g_{\omega,k}^{-1})'(\alpha_{k0})} g_{\omega,k}^{-1}(\alpha_{k0}) (g_{\omega,k}^{-1})'(\mathbf{z}_{t,z}^\top \alpha_k) K_h(z_t - z) = \rho g_{\omega,k}^{-1}(\alpha_{k0}). \quad (7)$$

Taking sums on both sides of the above equation over all k leads to

$$\rho = \sum_{k=1}^d \sum_{t=1}^T \frac{c_k[\hat{u}_t; g_{\theta,k}^{-1}(\mathbf{z}_{t,z}^\top \beta_k)]}{c_c(\hat{u}_t, \delta)(g_{\omega,k}^{-1})'(\alpha_{k0})} g_{\omega,k}^{-1}(\alpha_{k0}) (g_{\omega,k}^{-1})'(\mathbf{z}_{t,z}^\top \alpha_k) K_h(z_t - z). \quad (8)$$

We plug ρ back to equation (7) and then obtain

$$\alpha_{k0} = g_{\omega,k} \left\{ \rho^{-1} \sum_{t=1}^T \frac{c_k[\hat{u}_t; g_{\theta,k}^{-1}(\mathbf{z}_{t,z}^\top \beta_k)]}{c_c(\hat{u}_t, \delta)(g_{\omega,k}^{-1})'(\alpha_{k0})} g_{\omega,k}^{-1}(\alpha_{k0}) (g_{\omega,k}^{-1})'(\mathbf{z}_{t,z}^\top \alpha_k) K_h(z_t - z) \right\}. \quad (9)$$

Expectation Step

Let $\delta^{(0)}(z) = (\alpha_{10}^{(0)}, \dots, \alpha_{d0}^{(0)}, \beta_{10}^{(0)}, \dots, \beta_{d0}^{(0)}, \alpha_{11}^{(0)}, \dots, \alpha_{d1}^{(0)}, \beta_{11}^{(0)}, \dots, \beta_{d1}^{(0)})^\top$ be the initial estimators at each iterative. For a given grid point z , we can iteratively update α_{k0} by

$$\alpha_{k0}^{(1)} = g_{\omega,k} \left\{ \frac{1}{\rho^{(0)}} \sum_{t=1}^T \frac{c_k[\hat{u}_t; g_{\theta,k}^{-1}(\mathbf{z}_{t,z}^\top \beta_k^{(0)})]}{c_c(\hat{u}_t, \delta^{(0)}(z))(g_{\omega,k}^{-1})'(\alpha_{k0}^{(0)})} g_{\omega,k}^{-1}(\alpha_{k0}^{(0)}) (g_{\omega,k}^{-1})'(\mathbf{z}_{t,z}^\top \alpha_k^{(0)}) K_h(z_t - z) \right\},$$

for $k = 1, \dots, d$, where

$$\rho^{(0)} = \sum_{k=1}^d \sum_{t=1}^T \frac{c_k[\hat{u}_t; g_{\theta,k}^{-1}(\mathbf{z}_{t,z}^\top \beta_k^{(0)})]}{c_c(\hat{u}_t, \delta^{(0)}(z))(g_{\omega,k}^{-1})'(\alpha_{k0}^{(0)})} g_{\omega,k}^{-1}(\alpha_{k0}^{(0)}) (g_{\omega,k}^{-1})'(\mathbf{z}_{t,z}^\top \alpha_k^{(0)}) K_h(z_t - z).$$

Maximization Step

Let $\alpha_0 = (\alpha_{10}, \dots, \alpha_{d0})^\top$, $\alpha_1 = (\alpha_{11}, \dots, \alpha_{d1})^\top$, and $\beta = (\beta_{10}, \dots, \beta_{d0}, \beta_{11}, \dots, \beta_{d1})^\top$. Given $\{\alpha_{k0}^{(1)}\}$ updated from the E-step, we respectively maximize the objective function $Q(\hat{u}, \delta)$ with respect to α_1 and β . Following Acar *et al.* (2011), we find the numerical solutions through the Newton-Raphson iteration method

$$\alpha_1^{(1)} = \alpha_1^{(0)} - \frac{Q'_{\alpha_1}(\hat{u}, \alpha_0^{(1)}, \alpha_1^{(0)}, \beta^{(0)})}{Q''_{\alpha_1}(\hat{u}, \alpha_0^{(1)}, \alpha_1^{(0)}, \beta^{(0)})},$$

$$\beta^{(1)} = \beta^{(0)} - \frac{Q'_{\beta}(\hat{u}, \alpha_0^{(1)}, \alpha_1^{(1)}, \beta^{(0)})}{Q''_{\beta}(\hat{u}, \alpha_0^{(1)}, \alpha_1^{(1)}, \beta^{(0)})}, \quad (10)$$

where Q'_{α_1} and Q''_{α_1} are the first and second derivatives of $Q(\hat{u}, \delta)$ with respect to α_1 , and Q'_{β} and Q''_{β} are the first and second derivatives of $Q(\hat{u}, \delta)$ with respect to β , respectively.

G.2 Bandwidth Selection

When doing the nonparametric estimation, one needs to choose an appropriate bandwidth parameter h to balance the trade-off between the bias and variance of the estimator. In literature, researchers have proposed various approaches such as the cross-validation technique and the plug-in methods. Because we assume the sequence $\{\mathbf{x}_t, z_t\}_{t=1}^T$ is serially dependent (see Conditions in C7 of the main text), the *forward leave-one-out* cross-validation method can be used to select the bandwidth parameter.

Specifically, let $\hat{\xi}_h(\cdot) = (\hat{\omega}_h(\cdot)^\top, \hat{\theta}_h(\cdot)^\top)^\top$ denote the estimates of the weight and copula parameter functions with a known bandwidth parameter h . We obtain $\hat{\xi}_{h,t^*} = (\hat{\omega}_{h,1,t^*}, \dots, \hat{\omega}_{h,d,t^*}, \hat{\theta}_{h,1,t^*}, \dots, \hat{\theta}_{h,d,t^*})^\top$ at the sample point $\{\mathbf{x}_{t^*}, z_{t^*}\}$ using the data $\{\mathbf{x}_t, z_t, t < t^*\}$ for each $t_0 + 1 \leq t^* \leq T$, where t_0 is the minimum window size used to estimate $\hat{\xi}_{h,t_0+1}$. The sequential estimators $\{\hat{\xi}_{h,t^*}\}_{t^*=t_0+1}^T$ are constructed through this forward recursive scheme, and the optimal h^* can be achieved by maximizing the following objective function

$$h^* = \arg \max_h \sum_{t^*=t_0+1}^T \ln \left\{ \sum_{k=1}^d \hat{\omega}_{h,k,t^*}(z_{t^*}) c_k[\hat{u}_{t^*}; \hat{\theta}_{h,k,t^*}(z_{t^*})] \right\}.$$

The obtained h^* is the *forward leave-one-out* cross-validation estimator in terms of the log-likelihood. In practice, to ensure that the obtained *forward leave-one-out* cross-validation estimator h^* satisfies condition C6 of the main text, we can first pick an initial range $[h_1, h_2]$ with $h_1 = c_1 T^{-1/5}$ and $h_2 = c_2 T^{-1/5}$ for some constants c_1 and c_2 . Then the optimal h^* is selected within the range $[h_1, h_2]$ by maximizing the aforementioned objective function. See Hansen (2009) for details.

G.3 Confidence Intervals

Because most economic and financial data are serially dependent, the i.i.d bootstrap procedures may not be applicable here. Therefore, we follow the block bootstrap method proposed by Patton (2012) to calculate the pointwise confidence intervals on weight and copula parameters for serially dependent data. By dividing the data into several blocks, we can preserve

the original time series structure within a block. To construct the confidence intervals, the block bootstrap can be implemented as follows:

- (i). Generate a sample sequence $\{\mathbf{x}_t^*, z_t^*\}_{t=1}^T$ from the original data $\{\mathbf{x}_t, z_t\}_{t=1}^T$ using a stationary bootstrap technique described in Appendix A;
- (ii). Obtain the estimate of the marginal distribution $\{\widehat{u}_t^*\}_{t=1}^T$ by Step One described in Section 2.2;
- (iii). Calculate new estimator $\widehat{\xi}^*(z)$ at the grid point z by equation (4) in the main text with $\{\widehat{u}_t^*\}_{t=1}^T$;
- (iv). Repeat Steps (i)-(iii) B times (say, $B = 1000$), and get B values of the estimator $\widehat{\xi}^*(z)$ as an empirical sample at each grid point z . Let the $\alpha/2$ -th and $(1-\alpha/2)$ -th percentiles of the sample sequence $\{\widehat{\xi}^*(z)\}$ be $q_{\alpha/2}$ and $q_{1-\alpha/2}$, respectively; and
- (v). The empirical $100(1-\alpha)\%$ confidence interval for $\widehat{\xi}(z)$ is $[q_{\alpha/2}, q_{1-\alpha/2}]$.

H Proofs of the Key Results

In this section, we prove the main results in Section 2.3 of the main text.

Lemma 1: Define

$$\widetilde{Z}_t^\top \delta = \begin{pmatrix} \alpha_{10} + \alpha_{11}(z_t - z) \\ \vdots \\ \alpha_{d0} + \alpha_{d1}(z_t - z) \\ \beta_{10} + \beta_{11}(z_t - z) \\ \vdots \\ \beta_{d0} + \beta_{d1}(z_t - z) \end{pmatrix} \quad \text{and} \quad g^{-1}(\widetilde{Z}_t^\top \delta) = \begin{pmatrix} g_{\omega,1}^{-1}(\alpha_{10} + \alpha_{11}(z_t - z)) \\ \vdots \\ g_{\omega,d}^{-1}(\alpha_{d0} + \alpha_{d1}(z_t - z)) \\ g_{\theta,1}^{-1}(\beta_{10} + \beta_{11}(z_t - z)) \\ \vdots \\ g_{\theta,d}^{-1}(\beta_{d0} + \beta_{d1}(z_t - z)) \end{pmatrix}.$$

Let

$$\begin{aligned} A_T &= h \sum_{t=1}^T D_T^{-1} \widetilde{Z}_t [\ell g^{-1}]'(\widehat{u}_t, g^{-1}(\widetilde{Z}_t^\top \delta)) K_h(z_t - z), \\ A_T^* &= h \sum_{t=1}^T D_T^{-1} \widetilde{Z}_t [\ell g^{-1}]'(u_t, g^{-1}(\widetilde{Z}_t^\top \delta)) K_h(z_t - z), \\ B_T &= h \sum_{t=1}^T D_T^{-1} \widetilde{Z}_t [\ell g^{-1}]''(\widehat{u}_t, g^{-1}(\widetilde{Z}_t^\top \delta)) (D_T^{-1} \widetilde{Z}_t)^\top K_h(z_t - z), \quad \text{and} \end{aligned}$$

$$B_T^* = h \sum_{t=1}^T D_T^{-1} \tilde{Z}_t [\ell g^{-1}]''(u_t, g^{-1}(\tilde{Z}_t^\top \delta)) (D_T^{-1} \tilde{Z}_t)^\top K_h(z_t - z),$$

where $[\ell g^{-1}]'(\cdot) = \ell'(u_t, g^{-1}(\cdot)) \circ (g^{-1})'(\cdot)$ is the first derivative of $\ell(u_t, g^{-1}(\eta))$ with respect to η . $[\ell g^{-1}]''(\cdot) = [\ell'(u_t, g^{-1}(\cdot)) \circ (g^{-1})'(\cdot)]' = \ell''(u_t, g^{-1}(\cdot)) \circ \{[(g^{-1})'(\cdot)][(g^{-1})'(\cdot)]^\top\} + \text{diag}\{\ell'(u_t, g^{-1}(\cdot))\} \circ (g^{-1})''(\cdot)$ and $A \circ B$ is the Hadamard product between A and B . The marginal distribution estimator \hat{u}_t is obtained from the two-step procedures described in Section 2.2. Lemma 4.1 in Chen & Fan (2006) shows that $\sup_{1 \leq t \leq T} |\hat{u}_{st} - u_{st}| = O_p(1/\sqrt{T})$ for $s = 1, \dots, p$. Under the conditions C1 - C7, we have

$$A_T = A_T^* + o_p(1) \quad \text{and} \quad B_T = B_T^* + o_p(1).$$

Lemma 1 suggests that we can derive the asymptotic distribution of $\hat{\delta}$ without considering errors from the marginal distribution estimation. The estimator \hat{u}_t of the marginal part has little effect on the estimation of $\hat{\delta}$ if the sample size T is large.

Proof of Lemma 1:

The Lipschitz continuity of ℓ' and $[g^{-1}]'$, and the condition $\sup_{1 \leq t \leq T} |\hat{u}_{st} - u_{st}| = O_p(1/\sqrt{T})$ for $s = 1, \dots, p$ imply that

$$\begin{aligned} (|A_T - A_T^*|)_r &\leq \left(h \sum_{t=1}^T |D_T^{-1} \tilde{Z}_t ([\ell g^{-1}]'(\hat{u}_t, g^{-1}(\tilde{Z}_t^\top \delta)) - [\ell g^{-1}]'(u_t, g^{-1}(\tilde{Z}_t^\top \delta)))| K_h(z_t - z) \right)_r \\ &\leq C_1 \sup_{1 \leq t \leq T, 1 \leq s \leq p} |\hat{u}_{st} - u_{st}| \left(h \sum_{t=1}^T |D_T^{-1} \tilde{Z}_t \mathbf{1} K_h(z_t - z)| \right)_r \\ &= C_1 O_p(1/\sqrt{T}) O_p(\sqrt{Th}) = O_p(\sqrt{h}) = o_p(1), \end{aligned}$$

where $(\cdot)_r$ denotes r th element of a vector, C_1 is a constant and $\mathbf{1}$ is a $2d$ -dimensional vector with each element equals to one. $h \sum_{t=1}^T |D_T^{-1} \tilde{Z}_t \mathbf{1} K_h(z_t - z)|$ is of order $O_p(\sqrt{Th})$ because the first $2d$ elements are $(h/\sqrt{Th}) \sum_{t=1}^T K_h(z_t - z) = (Th/\sqrt{Th}) \frac{1}{T} \sum_{t=1}^T K_h(z_t - z) = O_p(\sqrt{Th})$ and the last $2d$ elements are $(Th)^{-1/2} h \sum_{t=1}^T |z_t - z| K_h(z_t - z) = \sqrt{Th} \frac{1}{Th} \sum_{t=1}^T |z_t - z| K_h(z_t - z) = \sqrt{Th} f(z) \int |u| k(u) du = O_p(\sqrt{Th})$.

In the same vein, we can show $|B_T - B_T^*| = o_p(1)$.

This completes the proof. □

Proof of Theorem 1:

To minimize the objective function $L(\hat{u}, \hat{\delta})$ given \hat{u} , it is equivalent to minimize $Th\{L(\hat{u}, \hat{\delta}) -$

$L(\hat{u}, \delta)$. And $Th\{L(\hat{u}, \hat{\delta}) - L(\hat{u}, \delta)\}$ can be written as

$$\begin{aligned}
& Th\{L(\hat{u}, \hat{\delta}) - L(\hat{u}, \delta)\} \\
&= h \sum_{t=1}^T \left\{ \ell(\hat{u}_t, g^{-1}(\tilde{Z}_t^\top \hat{\delta})) - \ell(\hat{u}_t, g^{-1}(\tilde{Z}_t^\top \delta)) \right\} K_h(z_t - z) \\
&= h \sum_{t=1}^T (\hat{\delta} - \delta)^\top \tilde{Z}_t [\ell g^{-1}]'(\hat{u}_t, g^{-1}(\tilde{Z}_t^\top \delta)) K_h(z_t - z) \\
&+ \frac{1}{2} h \sum_{t=1}^T (\hat{\delta} - \delta)^\top \tilde{Z}_t [\ell g^{-1}]''(\hat{u}_t, g^{-1}(\tilde{Z}_t^\top \delta)) \tilde{Z}_t^\top (\hat{\delta} - \delta) K_h(z_t - z) + o_p(1) \\
&= (D_T(\hat{\delta} - \delta))^\top A_T + \frac{1}{2} (D_T(\hat{\delta} - \delta))^\top B_T (D_T(\hat{\delta} - \delta)) + o_p(1) \\
&= (D_T(\hat{\delta} - \delta))^\top A_T^* + \frac{1}{2} (D_T(\hat{\delta} - \delta))^\top B_T^* (D_T(\hat{\delta} - \delta)) + o_p(1),
\end{aligned}$$

where A_T and B_T are defined in Lemma 1, and the last equation follows Lemma 1. After taking the first derivative with respect to $D_T(\hat{\delta} - \delta)$, we obtain

$$D_T(\hat{\delta} - \delta) = -B_T^{*-1} A_T^* + o_p(1).$$

Note that $E(-B_T^*)$ can be written as

$$\begin{aligned}
E(-B_T^*) &= E \left(-Th D_T^{-1} \tilde{Z}_t [\ell g^{-1}]''(u_t, g^{-1}(\tilde{Z}_t^\top \delta)) (D_T^{-1} \tilde{Z}_t)^\top K_h(z_t - z) \right) \\
&= \begin{pmatrix} G_1 & G_2 \\ G_2 & G_3 \end{pmatrix}.
\end{aligned}$$

Here,

$$\begin{aligned}
G_1 &= E \left(-[\ell g^{-1}]''(u_t, g^{-1}(\tilde{Z}_t^\top \delta)) K_h(z_t - z) \right) \\
&= E \left(-[\ell g^{-1}]''(u_t, g^{-1}(\delta^0 + \delta^1(z_t - z))) K_h(z_t - z) \right) \\
&= E \left(-[\ell g^{-1}]''(u_t, g^{-1}(\delta^0)) K_h(z_t - z) \right) + E \left([\ell g^{-1}]'''(u_t, g^{-1}(\delta^{0*})) \delta^1(z_t - z) K_h(z_t - z) \right) \\
&= f(z) \Sigma(z) \circ [(g^{-1})'(\eta(z))] [(g^{-1})'(\eta(z))]^\top + o_p(1),
\end{aligned}$$

where $\delta = (\delta^{0\top}, \delta^{1\top})^\top$ with $\delta^0 = \eta(z)$, and δ^{0*} is a point between δ^0 and $\tilde{Z}_t^\top \delta$. $\Sigma(z) = -E\{\ell''(u_t, g^{-1}(\eta(z_t))) | z_t = z\}$ with $\eta(z_t) = (w(z_t)^\top, \vartheta(z_t)^\top)^\top$. The third equation follows the Taylor expansion and the equation $[\ell g^{-1}]''(\cdot) = [\ell'(u_t, g^{-1}(\cdot)) \circ (g^{-1})'(\cdot)]' = \ell''(u_t, g^{-1}(\cdot)) \circ$

$$\{[(g^{-1})'(\cdot)][(g^{-1})'(\cdot)]^\top\} + \text{diag}\{\ell'(u_t, g^{-1}(\cdot))\} \circ (g^{-1})''(\cdot).$$

$$\begin{aligned} G_2 &= E\left(-[\ell g^{-1}]''(u_t, g^{-1}(\tilde{Z}_t^\top \delta))((z_t - z)/h)K_h(z_t - z)\right) \\ &= E\left(-[\ell g^{-1}]''(u_t, g^{-1}(\delta^0 + \delta^1(z_t - z)))((z_t - z)/h)K_h(z_t - z)\right) \\ &= E\left(-[\ell g^{-1}]''(u_t, g^{-1}(\delta^0))((z_t - z)/h)K_h(z_t - z)\right) + o_p(1) \\ &= \Sigma(z) \circ [((g^{-1})'(\eta(z)))[(g^{-1})'(\eta(z))]^\top] E\{((z_t - z)/h)K_h(z_t - z)\} + o_p(1) \\ &= o_p(1), \end{aligned}$$

and

$$\begin{aligned} G_3 &= E\left(-[\ell g^{-1}]''(u_t, g^{-1}(\tilde{Z}_t^\top \delta))((z_t - z)^2/h^2)K_h(z_t - z)\right) \\ &= E\left(-[\ell g^{-1}]''(u_t, g^{-1}(\delta^0))((z_t - z)^2/h^2)K_h(z_t - z)\right) + o_p(1) \\ &= \mu_2 f(z) \Sigma(z) \circ [((g^{-1})'(\eta(z)))[(g^{-1})'(\eta(z))]^\top] + o_p(1). \end{aligned}$$

It follows that

$$E(-B_T^*) = \text{diag}(1, \mu_2) \otimes \{f(z) \Sigma(z) \circ [((g^{-1})'(\eta(z)))[(g^{-1})'(\eta(z))]^\top]\} + o_p(1).$$

Similarly, we can show that $\text{Var}(-B_T^*) = o_p(1)$. Hence, $-B_T^* = \text{diag}(1, \mu_2) \otimes \{f(z) \Sigma(z) \circ [((g^{-1})'(\eta(z)))[(g^{-1})'(\eta(z))]^\top]\} + o_p(1)$.

In the same way, we can calculate

$$\begin{aligned} E(A_T^*) &= hE \sum_{t=1}^T D_T^{-1} \tilde{Z}_t [\ell g^{-1}]'(u_t, g^{-1}(\tilde{Z}_t^\top \delta)) K_h(z_t - z) \\ &= D_T \begin{pmatrix} \frac{h^2}{2} \mu_2 f(z) \Sigma(z) \circ [((g^{-1})'(\eta(z)))[(g^{-1})'(\eta(z))]^\top] \eta''(z) \\ \mathbf{0} \end{pmatrix} + o_p(D_T h^2), \end{aligned}$$

where $\mathbf{0}$ is a $2d$ -dimensional vector with all elements equal 0.

Further, following the same technology on pages 251-252 of Fan & Gijbels (1996), we have

$$\begin{aligned} E(A_T^* A_T^{*\top}) &= h^2 E \sum_{t=1}^T D_T^{-1} \tilde{Z}_t [\ell g^{-1}]'(u_t, g^{-1}(\tilde{Z}_t^\top \delta)) [\ell g^{-1}]'(u_t, g^{-1}(\tilde{Z}_t^\top \delta))^\top \tilde{Z}_t^\top D_T^{-1} K_h^2(z_t - z) \\ &= \text{diag}(\nu_0, \nu_2) \otimes \{f(z) \Omega(z) \circ [((g^{-1})'(\eta(z)))[(g^{-1})'(\eta(z))]^\top]\} + o_p(1), \end{aligned}$$

where $\Omega(z) = E\{\ell'(u_t, g^{-1}(\eta(z_t)))\ell'(u_t, g^{-1}(\eta(z_t)))^\top | z_t = z\}$.

It follows that

$$\begin{aligned} \text{Var}(D_T(\widehat{\delta} - \delta)) &= B_T^{*-1} E(A_T^* A_T^{*\top})(B_T^{*-1})^\top \\ &= \begin{pmatrix} \frac{\nu_0}{f(z)} \\ \frac{\nu_2}{\mu_2^2 f(z)} \end{pmatrix} \otimes \{\Sigma^{-1}(z)\Omega(z)\Sigma^{-1}(z) \circ \{[(g^{-1})'(\eta(z))][(g^{-1})'(\eta(z))]^\top\}^{-1}\}. \end{aligned}$$

To establish the asymptotic normality for $\widehat{\delta}$, we use Doob's small-block and large-block method, whose proof is similar to the counterparts on pages 252-255 of Fan & Gijbels (1996). Details are omitted here. This completes the proof. □

I Additional Simulation Results

I.1 High Dimensional Conditional Mixture Copula

We extend the bivariate mixture model to the 3- and 4-dimensional cases. For the former, we assume one marginal distribution follows $N(1, 0.5)$ while the other two follow the student's t -distributions with the degrees of freedom respectively equal to 4 and 5. For the latter, we assume two marginal distributions respectively follow $N(1, 0.5)$ and $N(0, 2)$ while the other two follow the student's t -distributions with the degrees of freedom respectively equal to 4 and 5. For simplicity, we only examine the scenario that data are generated from the mixture of Clayton and Gumbel.

Table D1 documents the MSEs of the copula parameter estimates and the selection rates of the four candidate copulas in the mixture model by the propose conditional mixture copula method. We consider the 3-dimensional copula (Panel 1) and 4-dimensional copula (Panel 2) respectively. Table D2 displays the MSEs of the weight estimates for the four candidate copulas under the 3-dimensional (Panel 1) and 4-dimensional (Panel 2) cases.

Table D1: MSEs of copula parameter estimates and accurate (inaccurate) rates of selection when the true model is a mixture of 3- and 4-dimensional Clayton and Gumbel copulas.

Panel 1		Combination: Clayton + Gumbel (3-dimension)											
		$T = 200$				$T = 500$				$T = 1000$			
		Clayton	Gumbel	Frank	Gaussian	Clayton	Gumbel	Frank	Gaussian	Clayton	Gumbel	Frank	Gaussian
Model 1 + Model 2	MSE	0.928	0.901	–	–	0.693	0.612	–	–	0.305	0.291	–	–
	Rate	1.000	0.827	(0.156)	(0.268)	1.000	0.952	(0.041)	(0.145)	1.000	0.973	(0.017)	(0.041)
Model 1 + Model 3	MSE	0.945	0.917	–	–	0.536	0.506	–	–	0.323	0.221	–	–
	Rate	1.000	0.919	(0.069)	(0.258)	1.000	0.929	(0.034)	(0.131)	1.000	0.962	(0.008)	(0.077)
Model 1 + Model 4	MSE	0.991	0.929	–	–	0.518	0.503	–	–	0.293	0.324	–	–
	Rate	1.000	0.938	(0.041)	(0.271)	1.000	0.935	(0.013)	(0.133)	1.000	0.993	(0.004)	(0.085)
Model 2 + Model 3	MSE	1.022	1.014	–	–	0.757	0.618	–	–	0.368	0.270	–	–
	Rate	1.000	0.873	(0.145)	(0.227)	1.000	0.943	(0.031)	(0.132)	1.000	0.982	(0.013)	(0.056)
Model 2 + Model 4	MSE	1.106	0.886	–	–	0.682	0.539	–	–	0.377	0.286	–	–
	Rate	1.000	0.891	(0.133)	(0.318)	1.000	0.962	(0.029)	(0.146)	1.000	0.992	(0.005)	(0.067)
Model 3 + Model 4	MSE	0.883	0.920	–	–	0.571	0.555	–	–	0.340	0.275	–	–
	Rate	1.000	0.857	(0.154)	(0.249)	1.000	0.944	(0.042)	(0.121)	1.000	0.991	(0.010)	(0.054)
Panel 2		Combination: Clayton + Gumbel (4-dimension)											
		$T = 200$				$T = 500$				$T = 1000$			
		Clayton	Gumbel	Frank	Gaussian	Clayton	Gumbel	Frank	Gaussian	Clayton	Gumbel	Frank	Gaussian
Model 1 + Model 2	MSE	0.945	0.922	–	–	0.714	0.631	–	–	0.320	0.303	–	–
	Rate	1.000	0.876	(0.132)	(0.276)	1.000	0.966	(0.059)	(0.143)	1.000	0.992	(0.010)	(0.074)
Model 1 + Model 3	MSE	1.027	0.933	–	–	0.624	0.517	–	–	0.377	0.284	–	–
	Rate	1.000	0.933	(0.069)	(0.231)	1.000	0.915	(0.023)	(0.132)	1.000	0.977	(0.008)	(0.067)
Model 1 + Model 4	MSE	0.983	0.996	–	–	0.603	0.683	–	–	0.313	0.342	–	–
	Rate	1.000	0.926	(0.043)	(0.264)	1.000	0.933	(0.014)	(0.154)	1.000	0.989	(0.005)	(0.088)
Model 2 + Model 3	MSE	1.054	1.136	–	–	0.840	0.702	–	–	0.432	0.368	–	–
	Rate	1.000	0.882	(0.127)	(0.221)	1.000	0.925	(0.030)	(0.128)	1.000	0.990	(0.013)	(0.053)
Model 2 + Model 4	MSE	1.127	0.872	–	–	0.753	0.628	–	–	0.386	0.340	–	–
	Rate	1.000	0.869	(0.118)	(0.305)	1.000	0.937	(0.035)	(0.160)	1.000	0.988	(0.008)	(0.076)
Model 3 + Model 4	MSE	0.909	0.955	–	–	0.631	0.657	–	–	0.379	0.354	–	–
	Rate	1.000	0.903	(0.166)	(0.256)	1.000	0.951	(0.041)	(0.127)	1.000	0.990	(0.009)	(0.068)

Notes. This table documents the mean square errors (MSEs) of the copula parameter estimates and the rates of selection for each copula in the mixture model when the true model includes two copulas. The sample size $T \in \{200, 500, 1000\}$ and each simulation is repeated 1000 times.

Table D2: MSEs of copula weight estimates when the true model is a mixture of 3- and 4-dimensional Clayton and Gumbel copulas.

Panel 1	Combination: Clayton + Gumbel (3-dimension)											
	T = 200				T = 500				T = 1000			
	Clayton	Gumbel	Frank	Gaussian	Clayton	Gumbel	Frank	Gaussian	Clayton	Gumbel	Frank	Gaussian
Model 1 + Model 2	0.013	0.012	-	-	0.006	0.010	-	-	0.002	0.003	-	-
Model 1 + Model 3	0.011	0.017	-	-	0.007	0.010	-	-	0.003	0.002	-	-
Model 1 + Model 4	0.014	0.014	-	-	0.005	0.009	-	-	0.003	0.003	-	-
Model 2 + Model 3	0.011	0.009	-	-	0.004	0.004	-	-	0.002	0.002	-	-
Model 2 + Model 4	0.015	0.018	-	-	0.007	0.009	-	-	0.002	0.002	-	-
Model 3 + Model 4	0.014	0.016	-	-	0.008	0.008	-	-	0.002	0.003	-	-

Panel 2	Combination: Clayton + Gumbel (4-dimension)											
	T = 200				T = 500				T = 1000			
	Clayton	Gumbel	Frank	Gaussian	Clayton	Gumbel	Frank	Gaussian	Clayton	Gumbel	Frank	Gaussian
Model 1 + Model 2	0.019	0.020	-	-	0.010	0.012	-	-	0.002	0.003	-	-
Model 1 + Model 3	0.015	0.022	-	-	0.008	0.012	-	-	0.003	0.004	-	-
Model 1 + Model 4	0.016	0.015	-	-	0.007	0.010	-	-	0.003	0.003	-	-
Model 2 + Model 3	0.012	0.011	-	-	0.005	0.005	-	-	0.002	0.003	-	-
Model 2 + Model 4	0.017	0.021	-	-	0.009	0.011	-	-	0.003	0.002	-	-
Model 3 + Model 4	0.017	0.017	-	-	0.010	0.009	-	-	0.003	0.004	-	-

Notes. The table documents the mean square errors (MSEs) of the mixture copula weight estimates. The sample size $T \in \{200, 500, 1000\}$ and each simulation is repeated 1000 times.

I.2 Misspecified Conditional Mixture Model

We still hold the working mixture copulas model as a combination the Clayton, Gumbel, Frank and Gaussian copulas, but consider two cases in which the working model is misspecified: (1). data are generated from the survival Gumbel copula; (2). data are generated from the mixture of Gumbel and survival Gumbel copulas. Table D3 displays the results of CM in the two misspecified scenarios.

Table D3: Percentage of selected copula(s) when the working model is misspecified.

Panel A	True Copula: Survival Gumbel											
	T = 200				T = 500				T = 1000			
	Clayton	Gumbel	Frank	Gaussian	Clayton	Gumbel	Frank	Gaussian	Clayton	Gumbel	Frank	Gaussian
Model 1	(0.996)	(0.000)	(0.021)	(0.093)	(1.000)	(0.000)	(0.016)	(0.039)	(1.000)	(0.000)	(0.010)	(0.025)
Model 2	(0.983)	(0.000)	(0.023)	(0.108)	(0.995)	(0.000)	(0.011)	(0.050)	(1.000)	(0.000)	(0.013)	(0.031)
Model 3	(0.996)	(0.000)	(0.027)	(0.079)	(0.994)	(0.000)	(0.014)	(0.042)	(0.998)	(0.000)	(0.008)	(0.023)
Model 4	(0.995)	(0.000)	(0.012)	(0.085)	(0.996)	(0.000)	(0.015)	(0.041)	(0.997)	(0.000)	(0.010)	(0.027)

Panel B	True Copula: Survival Gumbel + Gumbel											
	T = 200				T = 500				T = 1000			
	Clayton	Gumbel	Frank	Gaussian	Clayton	Gumbel	Frank	Gaussian	Clayton	Gumbel	Frank	Gaussian
Model 1 + Model 2	(0.901)	0.871	(0.100)	(0.231)	(0.947)	0.927	(0.050)	(0.139)	(0.951)	0.993	(0.030)	(0.061)
Model 1 + Model 3	(0.915)	0.845	(0.084)	(0.226)	(0.911)	0.932	(0.041)	(0.115)	(0.963)	0.968	(0.031)	(0.071)
Model 1 + Model 4	(0.915)	0.903	(0.055)	(0.315)	(0.903)	0.946	(0.048)	(0.162)	(0.974)	0.983	(0.022)	(0.088)
Model 2 + Model 3	(0.891)	0.918	(0.168)	(0.247)	(0.955)	0.937	(0.078)	(0.138)	(0.928)	0.990	(0.047)	(0.063)
Model 2 + Model 4	(0.947)	0.892	(0.143)	(0.281)	(0.948)	0.959	(0.077)	(0.149)	(0.990)	0.961	(0.038)	(0.059)
Model 3 + Model 4	(0.960)	0.934	(0.132)	(0.262)	(0.986)	0.971	(0.081)	(0.155)	(0.982)	0.984	(0.044)	(0.067)

Notes. This table documents the rates of selection for each copula in the mixture model when the working mixture model is misspecified. In Panel 1, data are generated from the Survival Gumbel copula. In Panel 2, data are generated from the mixture of Survival Gumbel and Gumbel copulas. Values with parentheses are the rates that copulas in the mixture copula are incorrectly selected. The sample size $T \in \{200, 500, 1000\}$ and each simulation is repeated 1000 times.

I.3 Model Selection by Hypothesis Test

Table D4 displays the accurate (inaccurate) rate of selection by the proposed hypothesis test procedure with the 0.05 significance level when the true model is an individual copula. The simulation setup here is the same with that of Table 1 in the main text.

Table D4: Accurate (inaccurate) rates of selection by the proposed hypothesis test procedure with the 0.05 significance level when the true model is an individual copula.

Panel 1												
True Copula: Clayton												
T = 200				T = 500				T = 1000				
	Clayton	Gumbel	Frank	Gaussian	Clayton	Gumbel	Frank	Gaussian	Clayton	Gumbel	Frank	Gaussian
Model 1	1.000	(0.000)	(0.071)	(0.055)	1.000	(0.000)	(0.051)	(0.043)	1.000	(0.000)	(0.027)	(0.029)
Model 2	1.000	(0.000)	(0.091)	(0.115)	1.000	(0.000)	(0.037)	(0.061)	1.000	(0.000)	(0.014)	(0.019)
Model 3	1.000	(0.000)	(0.060)	(0.071)	1.000	(0.000)	(0.003)	(0.031)	1.000	(0.000)	(0.000)	(0.013)
Model 4	1.000	(0.000)	(0.066)	(0.075)	1.000	(0.000)	(0.031)	(0.033)	1.000	(0.000)	(0.024)	(0.024)
Panel 2												
True Copula: Gumbel												
T = 200				T = 500				T = 1000				
	Clayton	Gumbel	Frank	Gaussian	Clayton	Gumbel	Frank	Gaussian	Clayton	Gumbel	Frank	Gaussian
Model 1	(0.000)	1.000	(0.065)	(0.044)	(0.000)	1.000	(0.054)	(0.031)	(0.000)	1.000	(0.021)	(0.030)
Model 2	(0.000)	1.000	(0.094)	(0.081)	(0.000)	1.000	(0.063)	(0.049)	(0.000)	1.000	(0.010)	(0.044)
Model 3	(0.000)	1.000	(0.069)	(0.071)	(0.000)	1.000	(0.002)	(0.039)	(0.000)	1.000	(0.000)	(0.014)
Model 4	(0.000)	1.000	(0.061)	(0.039)	(0.000)	1.000	(0.022)	(0.039)	(0.000)	1.000	(0.022)	(0.015)
Panel 3												
True Copula: Frank												
T = 200				T = 500				T = 1000				
	Clayton	Gumbel	Frank	Gaussian	Clayton	Gumbel	Frank	Gaussian	Clayton	Gumbel	Frank	Gaussian
Model 1	(0.000)	(0.012)	1.000	(0.193)	(0.000)	(0.000)	1.000	(0.079)	(0.000)	(0.000)	1.000	(0.027)
Model 2	(0.000)	(0.005)	1.000	(0.232)	(0.000)	(0.000)	1.000	(0.047)	(0.000)	(0.000)	1.000	(0.060)
Model 3	(0.000)	(0.001)	1.000	(0.243)	(0.000)	(0.000)	1.000	(0.132)	(0.000)	(0.000)	1.000	(0.073)
Model 4	(0.000)	(0.002)	1.000	(0.194)	(0.000)	(0.000)	1.000	(0.155)	(0.000)	(0.000)	1.000	(0.087)
Panel 4												
True Copula: Gaussian												
T = 200				T = 500				T = 1000				
	Clayton	Gumbel	Frank	Gaussian	Clayton	Gumbel	Frank	Gaussian	Clayton	Gumbel	Frank	Gaussian
Model 1	(0.000)	(0.009)	(0.087)	1.000	(0.000)	(0.000)	(0.055)	1.000	(0.000)	(0.000)	(0.031)	1.000
Model 2	(0.000)	(0.000)	(0.094)	1.000	(0.000)	(0.000)	(0.037)	1.000	(0.000)	(0.000)	(0.020)	1.000
Model 3	(0.000)	(0.001)	(0.069)	1.000	(0.000)	(0.000)	(0.051)	1.000	(0.000)	(0.000)	(0.028)	1.000
Model 4	(0.000)	(0.000)	(0.057)	1.000	(0.000)	(0.000)	(0.036)	1.000	(0.000)	(0.000)	(0.029)	1.000

Notes. This table documents the rates of selection for each copula in the mixture model when the true model is an individual copula. In each panel, values without parentheses are the rates that the true copula is correctly selected, while values with parentheses are the rates that copulas in the mixture copula are incorrectly selected.

References

- [1] Acar, E., Craiu, R., & Yao, F. (2011). Dependence Calibration in Conditional Copulas: A Nonparametric Approach. *Biometrics*, **67**, 445-453.
- [2] Chen, X., & Fan, Y. (2006). Estimation of Copula-Based Semiparametric Time Series Models. *Journal of Econometrics*, **130**, 307-335.
- [3] Dempster, A., Laird, N., & Rubin, D. (1977). Maximum Likelihood from Incomplete Data via the EM Algorithm. *Journal of the Royal Statistical Society, Series B*, **39**, 1-38.

- [4] Fan, J., & Gijbels, I. (1996). *Local Polynomial Modelling and Its Applications*. Volume 66. London: Chapman & Hall.
- [5] Hansen, B. (2009). Lecture Notes on Nonparametrics. Technical report, University of Wisconsin.
- [6] Patton, A. (2012). Copula Methods for Forecasting Multivariate Time Series in Handbook of Economic Forecasting, Vol. 2, Elsevier, Oxford, 2011.
- [7] Politis, D., & Romano, J. (1994). The Stationary Bootstrap. *Journal of the American Statistical Association*, **89**, 1303-1313.