I N S T I T U T D E S T A T I S T I Q U E

B I O S T A T I S T I Q U E E T

S C I E N C E S A C T U A R I E L L E S

(I S B A)

UNIVERSITÉ CATHOLIQUE DE LOUVAIN

D I S C U S S I O N
P A P E R

2013/23

Semiparametric Conditional Quantile Estimation through Copula-Based Multivariate Models

NOH, H., EL GHOUCH, A. and I. VAN KEILEGOM

Semiparametric Conditional Quantile Estimation through Copula-Based Multivariate Models

Hohsuk Noh
Université catholique de Louvain

Anouar El Ghouch
Université catholique de Louvain

Ingrid Van Keilegom
Université catholique de Louvain

June 14, 2013

Abstract

We consider a new approach in quantile regression modeling based on the copula function that defines the dependence structure between the variables of interest. The key idea of this approach is to rewrite the characterization of a regression quantile in terms of a copula and marginal distributions. After the copula and the marginal distributions are estimated, the new estimator is obtained as the weighted quantile of the response variable in the model. The proposed conditional estimator has three main advantages: it applies to both *iid* and time series data, it is automatically monotonic across quantiles and it can easily consider the case of multiple covariates in the estimation without introducing any extra complication. We show the asymptotic properties of our estimator when the copula is estimated by maximizing the pseudo log-likelihood and the margins are estimated nonparametrically including the case where the copula family is misspecified. We also present the finite sample performance of the estimator and illustrate the usefulness of our proposal by an application to the historical volatilities of Google and Yahoo companies.

Key words: Dependence Modeling; Check function; Markov Process; Pseudo Log-likelihood; Vine copula.

1 Introduction

Appropriate understanding and modeling of the dependence structure between financial assets is an important task. Especially, the characterization of the conditional dependence between random variables at a give quantile constitutes an important ingredient in modern risk management. As a copula has emerged as an effective tool to model dependence between non-elliptic and fat-tailed random variables, several authors attempted to propose conditional quantile estimation methods which are able to make use of the advantages of copulas in dependence modeling. Their common starting point is that since the copula function holds all information on the different forms of dependence between random variables, the form of the conditional quantile relationship is implied by the copula joining those random variables.

Some examples of such work include Bouyé and Salmon (2009), Chen and Fan (2006) and Chen et al. (2009). In an earlier version of Bouyé and Salmon (2009), the authors explicitly showed the link between the form of the conditional quantile relationship and the copula function for several well-known copula families such as elliptical copulas and Archimedean copulas. Further, they illustrated how such link can be used in conditional quantile estimation both when modeling the interdependence between random variables and when modeling the temporal dependence between them. Focusing more on the latter case, Chen and Fan (2006) studied a class of univariate copula-based stationary Markov models. Under the assumption of correct specification of the parametric copula, Chen and Fan (2006) established asymptotic properties of their quantile regression estimator when the copula is estimated by maximizing the pseudo log-likelihood and the marginals are estimated nonparametrically. Additionally, also in the time series context, Chen et al. (2009) employed parametric copula models to propose several distinct nonlinear quantile autoregression models and investigated the asymptotic properties of their estimator when both the copulas and the marginals could be globally misspecified but assuming the correct specification of a conditional quantile function at a particular quantile.

However, all these works consider a conditional quantile given just one covariate such as the conditional quantile of Y given X or the conditional quantile of Y_t given its lagged observation Y_{t-1}. Nevertheless, often it is necessary to consider multivariate quantile regression conditioning on more than one covariate. Apart from examples in the *iid* setup, there are many such examples in the time series setting where the copula-based quantile estimation methods should be extended. One example is a copula-based Markov process of higher order, for which Ibragimov (2009) studied how a copula characterizes the statistical properties of the corresponding Markov process. However, they did not investigate the issue of the conditional quantile estimation there. Another example is copula-based multivariate time series models, where for instance the dependence between two Markovian (stationary) time series X_t and Y_t is modeled via a copula which characterizes the dependence between $X_{t-1},\ Y_{t-1},\ X_t$ and Y_t, in other words, serial dependence and interdependence between two time series. Rémillard et al. (2012) discussed parameter estimation and goodness-of-fit testing for this model but did not address the issue of quantile estimation such as the conditional quantile of Y_t given $X_{t-1},\ Y_{t-1}$ and X_t.

Based on this observation, we are motivated to develop an extended version of the previous copula-based quantile regression methods to handle multiple covariates. The key idea of the previous methods is to express the conditional distribution function in terms of a certain partial derivative of the copula function and the marginal distributions, and obtain the conditional quantile through it. Although it is possible to consider an extension based on this idea, we find it better for convenient computation and concise theoretical development to estimate the conditional quantile function directly using the so-called 'check' function in Koenker and Bassett (1978) without going through the conditional distribution. The main idea of our new approach is to rewrite the check-function based characterization of a regression quantile in terms of a copula function and marginal distributions. Actually, our proposal is an extension of the recent work of Noh et al. (2013) from mean regression to quantile regression. However, non-differentiability of the check loss function in quantile regression makes the extension nontrivial, which needs a separate treatment. Additionally, to broaden the area of application, we derive the asymptotic properties of the estimator under general conditions where both the *iid* setting and the time series setting can be considered.

The rest of this paper is organized as follows. In Section 2, we introduce our conditional quantile estimation method. We present the asymptotic properties of the proposed estimator in Section 3 and present the finite sample performance of our estimator via some numerical simulations both in the *iid* and time series setting in Section 4. Finally, we analyze the daily log returns data of Google and Yahoo companies in Section 5 to illustrate the usefulness of our proposal. All technical details are deferred to the Appendix.

2 Copula-based Quantile Regression Estimator

Let $\boldsymbol{X} = (X_1, \ldots, X_d)$ be a covariate vector of dimension $d \geq 1$ and Y be a response variable with continuous cumulative distribution functions (c.d.f.s) $F_1, \ldots, F_d$ and F_0, respectively. We denote the density of X_j and Y by f_j and f_0, respectively. For a given $\boldsymbol{x} = (x_1, \ldots, x_d)^\top$, from the seminal work of Sklar (1959), the c.d.f. of $(Y, \boldsymbol{X})$ evaluated at $(y, \boldsymbol{x})$ can be expressed as $C(F_0(y), \boldsymbol{F}(\boldsymbol{x}))$. Here, $\boldsymbol{F}(\boldsymbol{x}) = (F_1(x_1), \ldots, F_d(x_d))^\top$ and C is the copula distribution of $(Y, \boldsymbol{X})$ defined by $C(u_0, u_1, \ldots, u_d) = P(U_0 \leq u_0, U_1 \leq u_1, \ldots, U_d \leq u_d)$, where $U_0 = F_0(Y)$ and $U_j = F_j(X_j),\ j = 1, \ldots, d$. The copula C is considered to hold all information on the dependence of $(Y, \boldsymbol{X})$ since it joins the marginals together to give the joint distribution. Naturally, it is expected that a given copula function implies a certain form of the conditional quantile relationship.

More precisely, the following link holds between the copula and the conditional distribution when the dimension of $\boldsymbol{X}$ is one:

$$\frac{\partial C(u_0, u_1)}{\partial u_1} = F_{Y|X_1}(F_0^{-1}(u_0)|F_1^{-1}(u_1)),$$

where $F_{Y|X_1}$ is the conditional distribution of Y given X_1. From this link, the conditional quantile function $m_\tau(x_1)$ of Y given $X_1 = x_1$ is derived in terms of the copula function and the marginals:

$$m_\tau(x_1) = F_0^{-1}(Q_{U_0|U_1}(\tau|F_1(x_1))), \tag{1}$$

where $Q_{U_0|U_1}(\tau|u_1)$ is the conditional τ-quantile function of U_0 given $U_1 = u_1$, which is the inverse function of $\partial C(u_0, u_1)/\partial u_1$ with respect to u_0. The expression (1) is the key idea underlying the previous works (Chen and Fan, 2006; Bouyé and Salmon, 2009; Chen et al., 2009). Although it is possible to consider the extension of the relation (1) to multiple covariate case ($d \geq 2$), we use another link between the conditional quantile and the copula via the check function to propose an extension which has computational convenience and for which concise asymptotic theory can be easily developed.

For that purpose, we note that from the definition of copula function, the conditional density of Y given $\boldsymbol{X} = \boldsymbol{x}$ is expressed as

$$f_0(y)\frac{c(F_0(y), \boldsymbol{F}(\boldsymbol{x}))}{c_{\boldsymbol{X}}(\boldsymbol{F}(\boldsymbol{x}))}, \tag{2}$$

where $c(u_0, \boldsymbol{u}) \equiv c(u_0, u_1, \ldots, u_d) = \partial^{d+1} C(u_0, u_1, \ldots, u_d)/\partial u_0 \partial u_1 \ldots \partial u_d$ is the copula density corresponding to C and $c_{\boldsymbol{X}}(\boldsymbol{u}) \equiv c_{\boldsymbol{X}}(u_1, \ldots, u_d) = \partial^d C(1, u_1, \ldots, u_d)/\partial u_1 \ldots \partial u_d$ is the copula density of $\boldsymbol{X}$. Interestingly, thanks to the expression (2), the τ-conditional quantile $m_\tau(\boldsymbol{x})$ of Y given $\boldsymbol{X} = \boldsymbol{x}$

can be written in terms of the copula and the marginals as follows:

$$\begin{aligned} m_\tau(\boldsymbol{x}) &= \arg\min_a \mathbb{E}\left[\rho(Y-a)|\boldsymbol{X}=\boldsymbol{x}\right] \\ &= \arg\min_a \mathbb{E}\left[\rho(Y-a)\ c(F_0(Y),\boldsymbol{F}(\boldsymbol{x}))\right], \end{aligned} \tag{3}$$

where $\rho(y) \equiv \rho_\tau(y) = y(\tau - I(y < 0))$ is the well known check function. Note that different from (1), the expression (3) is not affected by the dimension of the covariate vector $\boldsymbol{X}$.

If $\hat{c}$, $\hat{F}_0$ and $\hat{F}_j$ are any given estimators for c, F_0 and F_j, $j = 1, \ldots, d$, respectively, then $m_\tau(\cdot)$ can be estimated by

$$\hat{m}_\tau(\boldsymbol{x}) = \arg\min_a \sum_{i=1}^n \rho(Y_i - a)\, \hat{c}(\hat{F}_0(Y_i), \hat{\boldsymbol{F}}(\boldsymbol{x})), \tag{4}$$

where $\hat{\boldsymbol{F}}(\boldsymbol{x}) = (\hat{F}_1(x_1), \ldots, \hat{F}_d(x_d))^\top$. Since there are many different methods available in the literature for estimating a copula and a c.d.f., $\hat{m}(\boldsymbol{x})$ can be a nonparametric or a semiparametric or a fully parametric estimator depending on the method of estimating the components in (4). In this paper, we consider a semiparametric approach where the copula is parametrized but the marginal distributions are left unspecified as in Noh et al. (2013). Specifically, we assume a certain parametric family of copula densities, $\mathcal{C} = \{c(u_0, \boldsymbol{u}; \boldsymbol{\theta}),\ \boldsymbol{\theta} \in \Theta\}$, where Θ is a compact subset of $\mathbb{R}^p$, to which the true copula density belongs or by which it is well approximated.

3 Asymptotic Properties of the Proposed Estimator

In this section, we first provide general assumptions about the estimator, which will allow us to investigate its asymptotic properties both in the *iid* and dependent settings. Then, we will present the asymptotic representation of the estimator derived from the assumptions.

3.1 Assumptions

Before stating the assumptions, we introduce some notations, which will be used throughout the asymptotic analysis of our estimator. As mentioned in the previous section, we consider a certain family of copula densities, $\mathcal{C} = \{c(u_0, \boldsymbol{u}; \boldsymbol{\theta}),\ \boldsymbol{\theta} \in \Theta\}$, for estimating the true density $c(u_0, \boldsymbol{u})$. Define

$\boldsymbol{\theta}^*$ to be the (unique) pseudo-true copula parameter which lies in the interior of Θ and minimizes

$$I(\boldsymbol{\theta}) = \int_{[0,1]^{d+1}} \ln\left(\frac{c(u_0, \boldsymbol{u})}{c(u_0, \boldsymbol{u}; \boldsymbol{\theta})}\right) dC(u_0, \boldsymbol{u}).$$

Here, $I(\boldsymbol{\theta})$ is the classical Kullback-Leibler information criterion expressed in terms of copula densities instead of the traditional densities. It is clear that when the true copula density belongs to the given family, i.e. $c(.) = c(.; \boldsymbol{\theta}_0)$ for a certain $\boldsymbol{\theta}_0$, then we have $\boldsymbol{\theta}^* = \boldsymbol{\theta}_0$. Additionally, we let $c_{\boldsymbol{X}}(\boldsymbol{u}; \boldsymbol{\theta}) = \int c(u_0, \boldsymbol{u}; \boldsymbol{\theta}) du_0$. Also define

$$f_{\boldsymbol{\theta}}(y|\boldsymbol{x}) = f_0(y)\frac{c(F_0(y), \boldsymbol{F}(\boldsymbol{x}); \boldsymbol{\theta})}{c_{\boldsymbol{X}}(\boldsymbol{F}(\boldsymbol{x}); \boldsymbol{\theta})} \quad \text{and} \quad m_\tau(x; \boldsymbol{\theta}) = \arg\min_y \mathbb{E}\left[\rho_\tau(Y - y)c(F_0(Y), \boldsymbol{F}(\boldsymbol{x}); \boldsymbol{\theta})\right].$$

Concerning the partial derivatives of the copula density, we define

$$D_j c = \frac{\partial c}{\partial u_j}, \; j = 0, \ldots, d, \; \boldsymbol{c}' = (D_1 c, \ldots, D_d c)^\top \text{ and } \dot{\boldsymbol{c}} = \left(\frac{\partial c}{\partial \theta_1}, \ldots, \frac{\partial c}{\partial \theta_p}\right)^\top.$$

Here are the assumptions for our estimator.

(C0) $\{Y_i\}_{i\geq 1}$ is a strictly stationary process with β-mixing coefficient $\beta(i)$ satisfying $\beta(i) = O(i^{-\nu})$, as $i \to \infty$, for some $\nu > 1$.

(C1) $\hat{\boldsymbol{F}}(\boldsymbol{x}) - \boldsymbol{F}(\boldsymbol{x}) = O_p(n^{-1/2})$, where $\hat{\boldsymbol{F}}(\boldsymbol{x}) = (\hat{F}_1(x_1), \ldots, \hat{F}_d(x_d))^\top$ and $\hat{F}_j(\cdot)$ is an estimator for $F_j(\cdot)$.

(C2) $\hat{\boldsymbol{\theta}} - \boldsymbol{\theta}^* = O_p(n^{-1/2})$, where $\hat{\boldsymbol{\theta}}$ is an estimator of $\boldsymbol{\theta}^*$.

(C3) Let g denote either $\dot{\boldsymbol{c}}$ or $D_j c$, $j = 0, \ldots, d$ and $\boldsymbol{x} \in \mathbb{R}^d$ be a given point of interest such that $\boldsymbol{F}(\boldsymbol{x}) \in (0, 1)^d$.

  (i) $c(1, \boldsymbol{F}(\boldsymbol{x}); \boldsymbol{\theta}^*) < \infty$, $c(0, \boldsymbol{F}(\boldsymbol{x}); \boldsymbol{\theta}^*) < \infty$ and $0 < c_{\boldsymbol{X}}(\boldsymbol{F}(\boldsymbol{x}); \boldsymbol{\theta}^*) < \infty$;

  (ii) $(\boldsymbol{u}, \boldsymbol{\theta}) \mapsto g(u_0, \boldsymbol{u}; \boldsymbol{\theta})$ is continuous in $(\boldsymbol{u}, \boldsymbol{\theta})$ at $(\boldsymbol{F}(\boldsymbol{x}), \boldsymbol{\theta}^*)$ uniformly in $u_0 \in [0, 1]$;

  (iii) $u_0 \mapsto g(u_0, \boldsymbol{F}(\boldsymbol{x}); \boldsymbol{\theta}^*)$ is continuous on $[0, 1]$.

(C4) f_0 is continuous at $m_\tau(\boldsymbol{x}; \boldsymbol{\theta}^*)$.

(C5) $y \mapsto f_{\boldsymbol{\theta}^*}(y|\boldsymbol{x})$ is continuous at $m_\tau(\boldsymbol{x}; \boldsymbol{\theta}^*)$ and $f_{\boldsymbol{\theta}^*}(m_\tau(\boldsymbol{x}; \boldsymbol{\theta}^*)|\boldsymbol{x}) > 0$.

Assumption (C3) is satisfied for many popular copula families. (C4) and (C5) are typically assumed in quantile regression. Hence, in the following we will give some examples where (C0), (C1) and (C2) are satisfied in the *iid* setting and the time series setting.

3.1.1 IID setting

Suppose that we have $(Y_i, \boldsymbol{X}_i)$, $i = 1, \ldots, n$, an independent and identically distributed (*iid*) sample of n observations generated from the distribution of $(Y, \boldsymbol{X})$. In this case, (C0) is trivially satisfied. Concerning (C1), it is satisfied with the empirical distribution of X_j and its rescaled version which is popular in copula estimation context and is defined by

$$\hat{F}_j^s(x_j) = \frac{1}{n+1} \sum_{i=1}^{n} I(X_{j,i} \leq x_j).$$

Additionally, we can also use kernel smoothing method for estimating $F_j(\cdot)$, $j = 1, \ldots, d$. Let $k(\cdot)$ be a function which is a symmetric probability density function and $h \equiv h_n \to 0$ be a bandwidth parameter. Then, a kernel smoothing estimator $\tilde{F}_j$ is given by

$$\tilde{F}_j(x_j) = \frac{1}{n} \sum_{i=1}^{n} K\left(\frac{x_j - X_{j,i}}{h}\right),$$

where $K(x) = \int_{-\infty}^{x} k(t)dt$. If $nh^4 \to 0$ holds for the bandwidth h, then for $\hat{F}_j = \tilde{F}_j$, the following condition is satisfied:

(C1') $\hat{F}_j(x_j) = n^{-1} \sum_{i=1}^{n} I(X_{j,i} \leq x_j) + o_p(n^{-1/2})$,

from which (C1) follows. One advantage of using $\tilde{F}_j$ is that it results in a smooth estimate $\hat{m}_\tau(\boldsymbol{x})$, whereas the empirical distribution or its rescaled version does not. As for (C2), one example of the estimator $\hat{\theta}$ that satisfies (C2) in the literature is the maximum pseudo-likelihood (PL) estimator $\hat{\boldsymbol{\theta}}_{PL}$, which maximizes

$$L(\boldsymbol{\theta}) = \sum_{i=1}^{n} \log c\left(\hat{F}_0^s(Y_i), \hat{\boldsymbol{F}}^s(\boldsymbol{X}_i); \boldsymbol{\theta}\right), \quad (5)$$

where $\hat{\boldsymbol{F}}^s(\boldsymbol{X}_i) = (\hat{F}_1^s(X_{1,i}), \ldots, \hat{F}_d^s(X_{d,i}))^\top$. The estimator $\hat{\boldsymbol{\theta}}_{PL}$ was studied by several authors including Genest et al. (1995), Klaassen and Wellner (1997), Silvapulle et al. (2004), Tsukahara (2005) and Kojadinovic and Yan (2011), etc. If the score function of $c(u_0, \boldsymbol{u}; \boldsymbol{\theta})$ satisfies the assumptions (A.1)-(A.5) given in Tsukahara (2005), the PL estimator satisfies (C2) even when the copula family is misspecified as checked in Noh et al. (2013).

3.1.2 Time series setting

Assumptions under which a copula-based Markov process satisfies α-mixing or β-mixing have been studied by many authors. For example, if $\{Y_i\}_{i\geq 1}$ is a (stationary) univariate first-order Markov process and the copula of $(Y_i, Y_{i-1}) \equiv (Y_i, X_{i,1})$ satisfies certain conditions (see Proposition 2.1 of Chen and Fan (2006)), then Assumption (C0) holds and hence (C1) also holds with any estimator $\hat{F}_j(\cdot)$ satisfying (C1') (see Rio (2000)). Following similar ideas, Rémillard et al. (2012) extended this result to the case of copula-based multivariate first-order Markov process. As for copula-based Markov processes of higher order, unfortunately such results are rare. The only related work that we have found is Ibragimov (2009), who obtained a characterization of higher-order Markov processes in terms of copulas, but he did not discuss the mixing properties of the resulting process.

Concerning Assumption (C2), if we consider an extension of the maximum pseudo likelihood estimator studied by Genest et al. (1995) to the Markovian case, the resulting estimator satisfies (C2). For example, suppose that we have a sample $\{(Y_i, X_i) : i = 1, \ldots, n\}$ of a multivariate first-order Markov process generated from $(F_0(\cdot), F_1(\cdot), c(\cdot,\cdot,\cdot,\cdot;\boldsymbol{\theta}^*))$, where $c(\cdot,\cdot,\cdot,\cdot;\boldsymbol{\theta}^*)$ is the true parametric copula density associated with $(\mathbf{Z}_i, \mathbf{Z}_{i-1})$ with $\mathbf{Z}_i = (Y_i, X_i)^\top$ up to the unknown value θ^*. If we consider the following estimator for $\boldsymbol{\theta}^*$,

$$\hat{\boldsymbol{\theta}}_{PL}^{dep} = \arg\max_{\boldsymbol{\theta}\in\Theta} \sum_{i=2}^{n} \log\left\{\frac{c(\hat{\mathbf{U}}_i, \hat{\mathbf{U}}_{i-1};\boldsymbol{\theta})}{q(\hat{\mathbf{U}}_{i-1};\boldsymbol{\theta})}\right\}, \tag{6}$$

where

$$\hat{F}_0^s(\cdot) = \frac{1}{n+1}\sum_{i=1}^{n} I(Y_i \leq \cdot),\ \hat{F}_1^s(\cdot) = \frac{1}{n+1}\sum_{i=1}^{n} I(X_i \leq \cdot),\ \hat{\mathbf{U}}_i = (\hat{F}_0^s(Y_i), \hat{F}_1^s(X_i))^\top,$$
$$q(u_2, u_3;\boldsymbol{\theta}) = \int_0^1\int_0^1 c(u_0, u_1, u_2, u_3;\boldsymbol{\theta})\, du_1 du_0,$$

then (C2) holds according to Theorem 1 in Rémillard et al. (2012) under the assumptions (A1)-(A4) that they provided. For univariate copula-based first-order Markov models, a similar result can be found in Chen and Fan (2006).

3.2 Asymptotic representation of the estimator $\hat{m}_\tau(\boldsymbol{x})$

To realize the theoretical analysis of our estimator, we begin by introducing a few more notations:

- $\hat{F}_0(y) = n^{-1}\sum_{i=1}^n I(Y_i \le y)$.
- $\epsilon_i \equiv \epsilon_i(\boldsymbol{x};\boldsymbol{\theta}^*) = Y_i - m_\tau(\boldsymbol{x};\boldsymbol{\theta}^*)$.
- $\boldsymbol{e}'(\boldsymbol{x}) = \mathbb{E}\left[\psi_\tau(\epsilon_i)\boldsymbol{c}'(F_0(Y),\boldsymbol{F}(\boldsymbol{x});\boldsymbol{\theta}^*)\right]$ and $\dot{\boldsymbol{e}}(\boldsymbol{x}) = \mathbb{E}\left[\psi_\tau(\epsilon_i)\dot{\boldsymbol{c}}(F_0(Y),\boldsymbol{F}(\boldsymbol{x});\boldsymbol{\theta}^*)\right]$,

  where $\psi_\tau(y) = \tau - I(y \le 0)$.

Now we are ready to present the asymptotic representation of the proposed estimator.

Theorem 3.1 *Suppose that Assumptions (C0)-(C5) hold. Then, we have*

$$\sqrt{n}(\hat{m}_\tau(\boldsymbol{x}) - m_\tau(\boldsymbol{x};\boldsymbol{\theta}^*)) = \frac{1}{f_0(m_\tau(\boldsymbol{x};\boldsymbol{\theta}^*))c(F_0(m_\tau(\boldsymbol{x};\boldsymbol{\theta}^*)),\boldsymbol{F}(\boldsymbol{x});\boldsymbol{\theta}^*)}U_n + o_p(1),$$

where

$$\begin{aligned} U_n &= \sqrt{n}(\hat{\boldsymbol{F}}(\boldsymbol{x}) - \boldsymbol{F}(\boldsymbol{x}))^\top \boldsymbol{e}'(\boldsymbol{x}) + \sqrt{n}(\hat{\boldsymbol{\theta}} - \boldsymbol{\theta}^*)^\top \dot{\boldsymbol{e}}(\boldsymbol{x}) \\ &\quad -\sqrt{n}\left(\hat{F}_0(m_\tau(\boldsymbol{x};\boldsymbol{\theta}^*)) - F_0(m_\tau(\boldsymbol{x};\boldsymbol{\theta}^*))\right) c(F_0(m_\tau(\boldsymbol{x};\boldsymbol{\theta}^*)),\boldsymbol{F}(\boldsymbol{x});\boldsymbol{\theta}^*). \end{aligned}$$

Theorem 3.1 implies that the estimator $\hat{m}_\tau(\boldsymbol{x})$ converges in probability to $m_\tau(\boldsymbol{x};\boldsymbol{\theta}^*)$ as $n \to \infty$. Hence, when the copula family is misspecified, the estimator $\hat{m}_\tau(\boldsymbol{x})$ is no more consistent. In such situation, since $c(\cdot;\boldsymbol{\theta}^*)$ is just the best approximation to the true copula density $c(\cdot)$, we have a bias in the estimation of the true conditional quantile function $\hat{m}_\tau(\boldsymbol{x})$, which is (asymptotically) the difference between $m_\tau(\boldsymbol{x})$ and its best approximation $m_\tau(\boldsymbol{x};\boldsymbol{\theta}^*)$ among the function class $\{m(\boldsymbol{x};\boldsymbol{\theta}) : \boldsymbol{\theta} \in \Theta\}$.

As an application of Theorem 3.1, we consider the asymptotic normality of the estimator, for which we have to make a stronger assumption than (C2):

(C2') $\hat{\boldsymbol{\theta}} - \boldsymbol{\theta}^* = n^{-1}\sum_{i=1}^n \boldsymbol{\eta}_i + o_p(n^{-1/2})$, where $\boldsymbol{\eta}_i = \boldsymbol{\eta}(U_{0,i},\boldsymbol{U}_i;\boldsymbol{\theta}^*)$ is a p-dimensional random vector such that $\mathbb{E}\boldsymbol{\eta} = \boldsymbol{0}$ and $\mathbb{E}\,\boldsymbol{\eta}^\top\boldsymbol{\eta} < \infty$ and $\boldsymbol{U}_i = (U_{1,i},\ldots,U_{d,i})^\top$.

This stronger assumption also holds for $\hat{\boldsymbol{\theta}}_{PL}$ in the *iid* setting and $\hat{\boldsymbol{\theta}}_{PL}^{dep}$ in the time series setting with the same conditions for (C2). For the *iid* case, the function η is given by

$$\eta(U_0,\boldsymbol{U};\boldsymbol{\theta}) = J^{-1}(\boldsymbol{\theta}) \times K(U_0,\boldsymbol{U};\boldsymbol{\theta}), \tag{7}$$

where

$$J(\boldsymbol{\theta}) = \int_{[0,1]^{d+1}} \left(-\frac{\partial^2}{\partial \boldsymbol{\theta} \partial \boldsymbol{\theta}^\top} \log c(u_0, \boldsymbol{u}; \boldsymbol{\theta}) \right) dC(u_0, \boldsymbol{u})$$

and $K(U_0, \boldsymbol{U}; \boldsymbol{\theta})$ is a p-dimensional vector whose k-th element is

$$\frac{\partial}{\partial \theta_k} \log c(U_0, \boldsymbol{U}; \boldsymbol{\theta}) + \sum_{j=0}^{d} \int_{[0,1]^{d+1}} (I(U_j \le u_j) - u_j) \left(\frac{\partial^2}{\partial \theta_k \partial u_j} \log c(u_0, \boldsymbol{u}; \boldsymbol{\theta}) \right) dC(u_0, \boldsymbol{u}).$$

Concerning the time series case, see Chen and Fan (2006) for the univariate case and Rémillard et al. (2012) for the multivariate case. Replacing (C2) with (C2') in the previous assumptions, we have an asymptotic linear representation of the conditional quantile estimator $\hat{m}_\tau(\boldsymbol{x})$, which implies the asymptotic normality of $\hat{m}_\tau(\boldsymbol{x})$.

Corollary 3.2 *Suppose that Assumptions (C0)-(C1), (C1'), (C2'), (C3)-(C5) hold. Then, we have*

$$\begin{aligned}
&\sqrt{n}(\hat{m}_\tau(\boldsymbol{x}) - m_\tau(\boldsymbol{x}; \boldsymbol{\theta}^*)) \\
&= \frac{1}{\sqrt{n}} \sum_{i=1}^{n} \Big[\sum_{j=1}^{d} \{I(X_{j,i} \le x_j) - F_j(x_j)\}\ e_j(\boldsymbol{x}) + \boldsymbol{\eta}_i^\top \dot{\boldsymbol{e}}(\boldsymbol{x}) - \{I(Y_i \le m_\tau(\boldsymbol{x}; \boldsymbol{\theta}^*)) - F_0(m_\tau(\boldsymbol{x}; \boldsymbol{\theta}^*))\} \\
&\quad \times c(F_0(m_\tau(\boldsymbol{x}; \boldsymbol{\theta}^*)), \boldsymbol{F}(\boldsymbol{x}); \boldsymbol{\theta}^*) \Big] [f_0(m_\tau(\boldsymbol{x}; \boldsymbol{\theta}^*)) c(F_0(m_\tau(\boldsymbol{x}; \boldsymbol{\theta}^*)), \boldsymbol{F}(\boldsymbol{x}); \boldsymbol{\theta}^*)]^{-1} + o_p(1) \qquad (8)
\end{aligned}$$

where $\boldsymbol{e}'(\boldsymbol{x}) = (e_1(\boldsymbol{x}), \dots, e_d(\boldsymbol{x}))^\top$.

Specifically, Corollary 3.2 implies that $\sqrt{n}(\hat{m}_\tau(\boldsymbol{x}) - m_\tau(\boldsymbol{x}; \boldsymbol{\theta}^*))$ follows asymptotically a normal distribution with mean 0 and variance $\sigma^2 = \mathbb{V}ar(E_1) + 2\sum_{j=1}^{\infty} \mathbb{C}ov(E_{j+1}, E_1)$, where E_i denotes each summand in the summation of (8) (see Theorem 4.2 in Rio (2000)). Especially, since $\sigma^2 = \mathbb{V}ar(E_1)$ when the data are *iid*, we can estimate σ^2 by an estimator $\hat{\sigma}^2 = n^{-1} \sum_{i=1}^{n} \left\{ \widehat{E_i} - n^{-1} \sum_{i=1}^{n} \widehat{E_i} \right\}^2$, where $\widehat{E_i}$ is an estimator of E_i obtained by replacing all the unknown quantities in E_i by their corresponding estimates, for example, $\boldsymbol{\theta}^*$ by $\hat{\boldsymbol{\theta}}$ and F_0 by $\hat{F}_0$. Thanks to this estimator $\hat{\sigma}$, we can easily calculate the confidence interval for $m_\tau(\boldsymbol{x}; \boldsymbol{\theta}^*)$ using Corollary 3.2.

However, according to our simulation studies (see Section 4), the accuracy of the coverage of this confidence interval seems to sensitively depend on the accuracy of the estimation of the unknown quantities involved in E_i. Due to this, we propose to use a bootstrap method to approximate the asymptotic variance of the estimator $\hat{m}_\tau(\boldsymbol{x})$. The bootstrap that we use for the *iid* data in our simulations is outlined below:

Step 0: Obtain the copula parameter estimate $\hat{\boldsymbol{\theta}}$ from (5) and the marginal distribution estimates $\tilde{F}_0$ and $\tilde{F}_j$ using the kernel smoothing method with appropriate bandwidths $h_{0,n}$ and $h_{j,n},\ j = 1, \ldots, d$.

Step 1: Generate n independent random vectors $\mathbf{U}_i = (U^b_{0,i}, \ldots, U^b_{d,i})^\top,\ i = 1, \ldots, n$ from the estimated copula $c(u_0, \ldots, u_d; \hat{\boldsymbol{\theta}})$, where $\mathbf{U}_i = (U_{0,i}, \ldots, U_{d,i})^\top$.

Step 2: Let $Y^b_i = \tilde{F}_0^{-1}(U_{0,i})$ and $X^b_{j,i} = \tilde{F}_j^{-1}(U_{j,i}),\ j = 1, \ldots, d$.

Step 3: Repeating Steps 1 and 2 a large number of times, compute the bootstrap values $\hat{m}^b_\tau(\boldsymbol{x}),\ b = 1, \ldots, B$ of the estimator $\hat{m}_\tau(\boldsymbol{x})$ and then calculate the estimate of the asymptotic variance using them:

$$\hat{\sigma}^2_{boot} = \frac{1}{B}\sum_{b=1}^{B}(\hat{m}^b_\tau(\boldsymbol{x}))^2 - \left(\frac{1}{B}\sum_{b=1}^{B}\hat{m}^b_\tau(\boldsymbol{x})\right)^2.$$

Following similar ideas but modifying Steps 0 and 1 (see Section 4.3 in Chen and Fan (2006) and Section 2 in Rémillard et al. (2012)), we can estimate the asymptotic variance σ^2 in the time series setting using a bootstrap procedure. In Section 4, we investigate how the proposed bootstrap procedures perform by checking the coverage probabilities of the confidence interval for $\hat{m}_\tau(\boldsymbol{x})$ based on this procedure. In our simulations, we observe that it is important for satisfactory accuracy of the coverage of the confidence interval to use the kernel smoothing estimates of the marginal distributions following the concept of the smoothed bootstrap (Silverman and Young, 1987) in both the *iid* and time series setting. If we use the empirical distribution or its rescaled version, the coverage probability of the confidence interval does not approach the nominal confidence level at all.

4 Numerical Results

In this section, we firstly check whether the asymptotic theory for $\hat{m}_\tau(\boldsymbol{x})$ works both in the *iid* setting and in the time series setting. Secondly, we compare our semiparametric estimator with some competitors. For this purpose, we consider the following data generating processes (DGP):

- **DGP A** $(F_0(Y), F_1(X_1)) \sim$ Clayton copula with paramter α

– The resulting quantile regression function is

$$m_\tau(x_1) = F_0^{-1}((1 + F_1(x_1)^{-\alpha}(\tau^{-\alpha/(1+\alpha)} - 1))^{-1/\alpha}).$$

- **DGP B** $(F_0(Y), F_1(X_1), \ldots, F_d(X_d)) \sim$ Gaussian copula with correlation matrix $\Sigma = \begin{bmatrix} 1 & \boldsymbol{\rho}^\top \\ \boldsymbol{\rho} & \Sigma_{\boldsymbol{X}} \end{bmatrix}$

  – The resulting quantile regression function is

$$m_\tau(\boldsymbol{x}) = F_0^{-1}\left(\Phi\left(\sum_{j=1}^{d} a_j \Phi^{-1}(F_j(x_j)) + \sqrt{1 - \boldsymbol{\rho}^\top \boldsymbol{a}}\Phi^{-1}(\tau)\right)\right)$$

  where $\boldsymbol{a} = (a_1, \ldots, a_d)^\top = \Sigma_{\boldsymbol{X}}^{-1}\boldsymbol{\rho}$ and $\Phi(\cdot)$ is the c.d.f. of a standard normal distribution.

Although we describe each DGP with the focus on the *iid* setting, it can be also described for the time series setting. For example, using DGP A with $Y = Y_i$, $X_1 = Y_{i-1}$ and $F_0 = F_1 = F$, we can generate a sample $\{Y_i\}_{i=1}^n$ from a univariate first-order Markov model. Then, the conditional quantile function of Y_i given $Y_{i-1} = y$ is given by $m_\tau(y) = F^{-1}((1 + F(y)^{-\alpha}(\tau^{-\alpha/(1+\alpha)} - 1))^{-1/\alpha})$. Table 1 shows the parameters of the copula and the marginal distributions of each DGP subspecialized from DGPs A and B. All computations are done with R (R Development Core Team, 2011)

4.1 Verifying the asymptotic results about $\hat{m}_\tau(\boldsymbol{x})$

In this section, to verify the established asymptotic results, we compute a confidence interval for $m_\tau(\boldsymbol{x})$ either by the asymptotic representation or the bootstrap proposed in Section 3. By verifying whether the empirical coverage probabilities (ECP) of the $(1-\alpha)$-confidence interval for $m_\tau(\boldsymbol{x})$ are close to the nominal confidence level $(1 - \alpha)$, we indirectly check whether the estimator $\hat{m}_\tau(\boldsymbol{x})$ is asymptotically normal.

First, concerning the *iid* setting, we generate 500 random samples of size $n = 50, 100$ and $n = 200$ from DGPs A.1, B.1 and B.2 and compute a confidence interval for $m_\tau(\boldsymbol{x})$ using $\hat{\sigma}^2$ and Corollary 3.2. Table 2 shows the ECP of the $(1 - \alpha)$-confidence interval for $m_\tau(\boldsymbol{x})$ with $\alpha = 0.05$ and 0.1. We observe that depending on the location of the covariates and the quantile level, sometimes the ECP has a quite different value from its nominal confidence level. As mentioned before, the main reason for this is the inaccuracy of the estimation of the unknown quantities involved in the asymptotic

Table 1: The copula parameters and the marginal distributions for each subspecialized DGP. $\Phi_\nu(\cdot)$ is the c.d.f. of a random variable $t(\nu)$.

DGP	copula parameter	marginal distribution	m_τ
A.1 $(Y_i, X_{1,i})$	$\alpha = 1$	$Y_i \sim N(0,1),\ X_{1,i} \sim N(0,1)$	$\Phi^{-1}((1+\Phi(X_{1,i})^{-1}(\tau^{-1/2}-1))^{-1})$
A.2 (Y_i, Y_{i-1})	$\alpha = 0.5$	$Y_i,\ Y_{i-1} \sim N(0,1)$	$\Phi^{-1}((1+\Phi(Y_{i-1})^{-1/2}(\tau^{-1/3}-1))^{-2})$
B.1 $(Y_i, X_{1,i})$	$\rho = 0.6$	$Y_i \sim N(1,1),\ X_{1,i} \sim N(0,1)$	$1 + 0.8\Phi^{-1}(\tau) + 0.6X_{1,i}$
B.2 $(Y_i, X_{1,i}, X_{2,i})$	$\Sigma = \begin{pmatrix} 1 & -0.5 & 0.9 \\ -0.5 & 1 & -0.4 \\ 0.9 & -0.4 & 1 \end{pmatrix}$	$Y_i \sim U[0,1],\ X_{1,i} \sim N(0,1),$ $X_{2,i} \sim N(0,1)$	$\Phi(-0.17X_{1,i} + 0.83X_{2,i} + 0.41\Phi^{-1}(\tau))$
B.3 $(Y_i, X_{1,i}, X_{2,i}, X_{3,i})$	$\Sigma = \begin{pmatrix} 1 & 0.3 & 0.9 & 0.7 \\ 0.3 & 1 & 0.5 & 0.25 \\ 0.9 & 0.5 & 1 & 0.5 \\ 0.7 & 0.25 & 0.5 & 1 \end{pmatrix}$	$Y_i \sim U[0,1],\ X_{1,i} \sim N(0,1),$ $X_{2,i} \sim N(0,1),\ X_{3,i} \sim N(0,1)$	$\Phi(-0.20X_{1,i} + 0.83X_{2,i} + 0.33X_{3,i} + 0.27\Phi^{-1}(\tau))$
B.4 $(Y_i, X_i, Y_{i-1}, X_{i-1})$	$\Sigma = \begin{pmatrix} 1 & 0.6 & 0.3 & 0.4 \\ 0.6 & 1 & 0.5 & 0.2 \\ 0.3 & 0.5 & 1 & 0.6 \\ 0.4 & 0.2 & 0.6 & 1 \end{pmatrix}$	$Y_i,\ Y_{i-1} \sim N(0,1),$ $X_i,\ X_{i-1} \sim t(4)$	$0.65\Phi^{-1}(\Phi_4(X_i)) - 0.30Y_{i-1} + 0.44\Phi^{-1}(\Phi_4(X_{i-1})) + 0.72\Phi^{-1}(\tau)$

representation. As is typical in quantile regression, the asymptotic representation for $\hat{m}_\tau(\boldsymbol{x})$ involves the conditional density $f_{Y|\boldsymbol{X}}(m_\tau(\boldsymbol{x})|\boldsymbol{x})$, which is equal to $f_0(m_\tau(\boldsymbol{x};\boldsymbol{\theta}^*))c(F_0(m_\tau(\boldsymbol{x};\boldsymbol{\theta}^*)), \boldsymbol{F}(\boldsymbol{x});\boldsymbol{\theta}^*)$ when the copula family is correct. Since it controls the scale of the estimated asymptotic variance of $\hat{m}_\tau(\boldsymbol{x})$, the estimation accuracy of it seems to affect the ECP a lot. To confirm our claim, we evaluate the asymptotic representation using the true values of all involved quantities and compute the ECP. As was expected, we observe in Table 3 that the recalculated ECP is close to the nominal confidence level as the sample size increases. Finally, we compute the confidence interval and its ECP using the bootstrap ($B = 200$) proposed in Section 3. Table 3 suggests that the bootstrap method seems to solve the problem more or less.

To verify the asymptotic behavior of our estimator under misspecification, we generate data from a Clayton copula according to DGP A.1 but in the estimation procedure we use a Gaussian copula. The 'pseudo'-true quantile regression function is $m_\tau(x_1;\rho^*) = \Phi^{-1}(\tau)\sqrt{1-\rho^{*2}} + \rho^* x_1$ with $\rho^* = 0.503$ for the Clayton copula with $\alpha = 1$. Figure 1 shows the boxplots of the estimators $\hat{m}_\tau(x_1)$ obtained from 500 random samples of size 200. We see that the observed values are symmetrically distributed around the pseudo-true parameter $m_\tau(x_1;\rho^*)$ instead of the true parameter $m_\tau(x_1)$ as expected according to Theorem 3.1. The difference between these two quantities corresponds exactly to the asymptotic bias.

As for the time series setting, we generate 500 random samples of size $n = 100$ and 200 from DGPs A.2 and B.4. To compute a confidence interval for $m_\tau(\boldsymbol{x})$, we estimate σ^2 using the bootstrap described in Section 3 with $B = 200$. We observe that the bootstrap seems to work reasonably well in terms of the ECP as shown in Table 4. The ECP when $\alpha = 0.1$ seems to be somewhat higher than the nominal confidence level but gets closer to it as the sample size grows.

4.2 Comparison with other methods

In this subsection we compare our semiparametric estimator both with semiparametric and nonparametric competitors. We consider four estimators for comparison.

- $\hat{m}_{tc}$: our estimator when the true copula family is used.
- $\hat{m}_{uc}$: our estimator when the copula density family is adaptively selected using the data (see the explanation below).
- $\hat{m}_{ll}$: local linear estimator with the bandwidth selected by cross-validation based on the check-function.

- $\hat{m}_{si}$: single index regression estimator based on a two stage estimation method; the single-index coefficients are first estimated by the method of Zhu et al. (2012), and then the link function is estimated in the same way as for $\hat{m}_{ll}$.

In addition to this, we consider the nonlinear quantile regression estimator $\hat{m}_{nl}$, which exploits the true link function as a reference case. We use the R package *quantreg* to calculate $\hat{m}_{nl}$. Concerning the estimator $\hat{m}_{uc}$, we use the simplified pair-copula decomposition of the copula density (R-vine) as in Noh et al. (2013). The main idea of it is to decompose a multivariate copula to a cascade of bivariate copulas so that we can take advantage of the relative simplicity of bivariate copula selection and estimation. For details, we refer to Aas et al. (2009), Brechmann (2010), Noh et al. (2013) and references therein. Specifically, we choose one decomposition of the copula density (among many R-vine structures) for the data, and then choose the pair-copulas independently among ten candidate copulas: two are elliptical (Gaussian and Student t) and eight are Archimedean (Clayton, Gumbel, Frank, Joe, Clayton-Gumbel, Joe-Gumbel, Joe-Clayton and Joe-Frank) using the R package *VineCopula*. As a selection criterion for bivariate copulas, we use the Akaike information criterion (AIC), which is shown to work in this context (see Dißmann et al. (2013)).

For comparison with other methods, we consider DGPs B.2 and B.3 to generate data. For performance evaluation of each method, we consider the empirical integrated mean squared error (IMSE), which is defined by

$$\text{IMSE} = \frac{1}{N}\sum_{l=1}^{N} \text{ISE}(\hat{m}_\tau^{(l)}) = \frac{1}{N}\sum_{l=1}^{N}\left[\frac{1}{I}\sum_{i=1}^{I}\left(\hat{m}_\tau^{(l)}(\boldsymbol{x}_i) - m_\tau(\boldsymbol{x}_i)\right)^2\right],$$

where $\{\boldsymbol{x}_i, i = 1, \ldots, I\}$ is a fixed evaluation set which corresponds to a random sample of size $I = 500$ generated from the distribution of $\boldsymbol{X}$, $\hat{m}^{(l)}(\cdot)$ is the estimated regression function from the l-th data sample. As expected, the estimator $\hat{m}_{nl}$ performs best in both DGPs. Our estimator $\hat{m}_{tc}$, which uses the information about the copula family, ranks the second. Additionally, even the estimator $\hat{m}_{uc}$ is a bit behind $\hat{m}_{tc}$ in performance due to the pair-copula selection step before the estimation, but it is still advantageous over the other semiparametric estimator $\hat{m}_{si}$ and the nonparametric estimator $\hat{m}_{np}$. From this observation, we see that when the true DGP can be described using a certain copula which belongs to the copula family under consideration, which is the case here, our proposed methods can be a good choice in quantile regression. However, the performance of our estimators may depend on whether the true copula density fits into the copula family under our consideration or not. To see the impact of it, we consider an additional DGP.

- **DGP C** $Y = m(X_1, X_2, X_3) + \sigma\epsilon$, where $\epsilon \sim N(0, 1)$ independent of $\mathbf{X}$.
  - $m(X_1, X_2, X_3) = \Psi(-0.3X_1 + 0.9X_2 + 0.3X_3)$ where Ψ is the c.d.f. of the standard Cauchy distribution and $\sigma = 0.1$
  - The resulting quantile regression function is $\Psi(-0.3X_1 + 0.9X_2 + 0.3X_3) + 0.1\Phi^{-1}(\tau)$.
  - $\boldsymbol{X} = (X_1, X_2, X_3)^\top$ is multivariate normal with mean $\mathbf{0}$ and $cov(X_{j_1}, X_{j_2}) = 0.5^{|j_1 - j_2|}$.

Table 7 shows the performance of each method. Note that since we have no knowledge about the true copula, the estimator $\hat{m}_{tc}$ is not available. In this case, as before the estimator $\hat{m}_{nl}$ performs best but the single-index estimator ranks the second in most cases. However, our estimator $\hat{m}_{uc}$ still shows comparable performance to $\hat{m}_{si}$ and performs better than the nonparametric estimator $\hat{m}_{np}$. This suggests that the copula family under consideration is flexible enough to approximate the true copula density in a certain degree although it does not include the density. Additionally, it also implies that our method has the advantage over the classical single-index model and that it is more flexible and adapts better to different settings.

5 Empirical Application

To illustrate the usefulness of our method, we analyze the historical volatilities of Yahoo($\{Y_i\}$) and Google($\{X_i\}$) companies over a nine-year period (2004-2013, 2160 trading days). Every 5 trading days we compute the standard deviation of the log returns of each company and consider it as the historical volatility of the period. The volatilities of both companies over the whole period are plotted in Figure 2 (432 observations for each time series). When the volatility data for both companies in a certain length of period until a particular time point is given ($\{(Y_i, X_i), i = 1, \ldots, n\}$), we consider a problem of predicting the volatility of the Yahoo company for the following period consisting of 5 trading days (Y_{n+1}) using various copula-based estimation models considered in this work. For prediction, we will use the conditional median estimate from each model. Here is the description of each model (M1-M6):

- $(Y_i, Y_{i-1}) \sim C(\cdot, \cdot; \theta) \Rightarrow Y_i | Y_{i-1}$

  M1 - C: Gaussian copula, M2 - C: Student t copula

- $(Y_i, X_i) \sim C(\cdot, \cdot; \theta) \Rightarrow Y_i | X_i$

  M3 - C: Gaussian copula, M4 - C: Student t copula

- $(Y_i, X_i, Y_{i-1}, X_{i-1}) \sim C(\cdot,\cdot,\cdot,\cdot;\boldsymbol{\theta}) \Rightarrow Y_i|X_i, Y_{i-1}, X_{i-1}$

  M5 - C: Gaussian copula, M6 - C: Student t copula

M1 and M2 only consider temporal dependence between the returns of the Yahoo company for prediction, whereas M3 and M4 consider both interdependence between the returns of the two companies and temporal dependence in each company's returns. Different from these models, M5 and M6 ignore temporal dependence and only focus on interdependence for prediction. To evaluate the performance of each model, we calculate the predicted value of Y_{n+1} repeatedly as we slide the time window of size $n = 50$ (250 trading days = 1 year) by one week (5 trading days), and compare the predicted values with the observed ones. From Figure 2, since it is clear that there exist both temporal dependence and interdependence, we expect that Models M3 and M4 considering both types of dependence will be better in prediction than the models considering just one of them. Before fitting the models, we checked whether the data of each company in each window satisfy at least stationarity using the Phillips-Perron unit root test (Phillips and Perron, 1988). The tests never reject the stationarity assumption for both time series.

Table 8 shows the prediction performance of each model measured by the criterion (PRED = $\sum_{k=1}^{382}(\hat{Y}_k - Y_k)^2$, k is an index for denoting evaluation points). As was expected, we observe that considering both dependence is better for prediction than only considering one kind of dependence regardless of the kind of copula used. This finding suggests that our extension to the multiple covariates case seems to be a useful contribution to the implementation of such idea. Additionally, from the fact that M5 and M6 are comparable with M3 and M4 in performance, we see that for these data the inter-correlation between two time series is a more important factor for prediction than the auto-correlation in the time series but this might not be the case in other data. Finally, one might think that since the current information (X_{n+1}) of the other company (Google), which might have some link with the company of interest (Yahoo), is not always available for the prediction (of Y_{n+1}), the model has some limitation in practice. However, considering stocks of a company which has many branches overseas, such current information is available due to time difference.

6 Concluding Remarks

We proposed a new semiparametric conditional quantile estimation method using copula-based multivariate models, especially with the focus on the extension to the case of multiple covariates. We established the asymptotic properties of our estimator under general assumptions, which cover both the *iid* and the dependent case taking misspecification into account. Although we present some examples which fit into our theoretical framework, there may be other interesting examples which can be treated in the framework. One example is a copula-based Markov process of higher order. Because the convergence of the copula parameter estimator is a key part in our assumptions, the study of copula parameter estimation in such models is not only a good future research topic, but also an important step to broaden the application of our work.

Acknowledgements

All the authors acknowledge financial support from IAP research network P7/06 of the Belgian Government (Belgian Science Policy). Additionally, H. Noh and I. Van Keilegom acknowledge financial support from the European Research Council under the European Community's Seventh Framework Programme (FP7/2007-2013) / ERC Grant agreement No. 203650, and A. El Ghouch and I. Van Keilegom acknowledge the support from the contract 'Projet d'Actions de Recherche Concertées' (ARC) 11/16-039 of the 'Communauté française de Belgique', granted by the 'Académie universitaire Louvain'. Finally, the authors would like to thank Fabian Y.R.P. Bocart for giving much help and sharing his insight concerning the application of our method to financial data.

Appendix

In this appendix, we first prove a technical lemma and then present the proof of Theorem 3.1.

Lemma 6.1 *Define* $A_n(t) = \sum_i \left(\rho_\tau(\epsilon_i - t/\sqrt{n}) - \rho_\tau(\epsilon_i)\right) c(\hat{F}_0(Y_i), \hat{\boldsymbol{F}}(\boldsymbol{x}); \hat{\boldsymbol{\theta}})$. *If the assumptions (C0)-(C5) hold, then we have*

$$A_n(t) = -tU_n + \frac{1}{2}t^2 f_0(m_\tau(\boldsymbol{x};\boldsymbol{\theta}^*))c(F_0(m_\tau(\boldsymbol{x};\boldsymbol{\theta}^*)), \boldsymbol{F}(\boldsymbol{x}); \boldsymbol{\theta}^*) + o_p(1),$$

where

$$\begin{aligned} U_n &= \sqrt{n}(\hat{\boldsymbol{F}}(\boldsymbol{x}) - \boldsymbol{F}(\boldsymbol{x}))^\top \boldsymbol{e}'(\boldsymbol{x}) + \sqrt{n}(\hat{\boldsymbol{\theta}} - \boldsymbol{\theta}^*)^\top \dot{\boldsymbol{e}}(\boldsymbol{x}) \\ &\quad -\sqrt{n}\left(\hat{F}_0(m_\tau(\boldsymbol{x};\boldsymbol{\theta}^*)) - F_0(m_\tau(\boldsymbol{x};\boldsymbol{\theta}^*))\right) c(F_0(m_\tau(\boldsymbol{x};\boldsymbol{\theta}^*)), \boldsymbol{F}(\boldsymbol{x}); \boldsymbol{\theta}^*). \end{aligned}$$

Proof.

Using Knight's (1998) identity, $\rho_\tau(u-v)-\rho_\tau(u) = -v\psi_\tau(u)+r(u,v)$, with $r(u,v) = \int_0^v (I(u \leq s) - I(u \leq 0))\, ds$, $A_n(t)$ can be written as $A_n(t) = -tA_{1,n} + A_{2,n}(t)$, where

$$A_{1,n} = \frac{1}{\sqrt{n}} \sum_i \psi_\tau(\epsilon_i) c(\hat{F}_0(Y_i), \hat{\boldsymbol{F}}(\boldsymbol{x}); \hat{\boldsymbol{\theta}}) \text{ and } A_{2,n}(t) = \sum_i r(\epsilon_i, t/\sqrt{n}) c(\hat{F}_0(Y_i), \hat{\boldsymbol{F}}(\boldsymbol{x}); \hat{\boldsymbol{\theta}}).$$

Using a first-order Taylor expansion, we have

$$A_{1,n} = n^{-1/2} \sum_{i=1}^n \psi_\tau(\epsilon_i) c(F_0(Y_i), \boldsymbol{F}(\boldsymbol{x}); \boldsymbol{\theta}^*) + A_{11,n} + A_{12,n} + A_{13,n}, \tag{9}$$

where

$$\begin{aligned}
A_{11,n} = &\quad n^{-1/2} \sum_{i=1}^n \psi_\tau(\epsilon_i)(\hat{F}_0(Y_i) - F_0(Y_i))\, D_0 c(\tilde{U}_{0,i}, \tilde{\boldsymbol{U}}_i; \tilde{\boldsymbol{\theta}}_i), \\
A_{12,n} = &\quad n^{-1/2} \sum_{i=1}^n \psi_\tau(\epsilon_i)(\hat{\boldsymbol{F}}(\boldsymbol{x}) - \boldsymbol{F}(\boldsymbol{x}))^\top \boldsymbol{c}'(\tilde{U}_{0,i}, \tilde{\boldsymbol{U}}_i; \tilde{\boldsymbol{\theta}}_i), \\
A_{13,n} = &\quad n^{-1/2} \sum_{i=1}^n \psi_\tau(\epsilon_i)(\hat{\boldsymbol{\theta}} - \boldsymbol{\theta}^*)^\top\, \dot{\boldsymbol{c}}(\tilde{U}_{0,i}, \tilde{\boldsymbol{U}}_i; \tilde{\boldsymbol{\theta}}_i),
\end{aligned}$$

with $\tilde{U}_{0,i} = F_0(Y_i) + t_{i,n}(\hat{F}_0(Y_i) - F_0(Y_i))$, $\tilde{\boldsymbol{U}}_i = \boldsymbol{F}(\boldsymbol{x}) + t_{i,n}(\hat{\boldsymbol{F}}(\boldsymbol{x}) - \boldsymbol{F}(\boldsymbol{x}))$ and $\tilde{\boldsymbol{\theta}}_i = \boldsymbol{\theta}^* + t_{i,n}(\hat{\boldsymbol{\theta}} - \boldsymbol{\theta}^*)$ for some random quantity $t_{i,n} \in [0,1]$. By adding and subtracting $D_0 c(F_0(Y_i), \boldsymbol{F}(\boldsymbol{x}); \boldsymbol{\theta}^*)$ in the sum, decompose further the term $A_{11,n}$ as

$$A_{11,n} = n^{-1/2} \sum_{i=1}^n \psi_\tau(\epsilon_i)(\hat{F}_0(Y_i) - F_0(Y_i))\, D_0 c(F_0(Y_i), \boldsymbol{F}(\boldsymbol{x}); \boldsymbol{\theta}^*) + R_{1,n},$$

where

$$R_{1,n} = n^{-1/2} \sum_{i=1}^n \psi_\tau(\epsilon_i)(\hat{F}_0(Y_i) - F_0(Y_i)) \left[D_0 c(\tilde{U}_{0,i}, \tilde{\boldsymbol{U}}_i; \tilde{\boldsymbol{\theta}}_i) - D_0 c(F_0(Y_i), \boldsymbol{F}(\boldsymbol{x}); \boldsymbol{\theta}^*) \right].$$

By Assumption (C3), $\max_{1 \leq i \leq n} \left| D_0 c(\tilde{U}_{0,i}, \tilde{\boldsymbol{U}}_i; \tilde{\boldsymbol{\theta}}_i) - D_0 c(F_0(Y_i), \boldsymbol{F}(\boldsymbol{x}); \boldsymbol{\theta}^*) \right| = o_p(1)$. Moreover, by Assumption (C0) and Donsker's Theorem, see Theorem 7.2 in Rio (2000), $\sup_y |\hat{F}_0(y) - F_0(y)| =$

$O_p(n^{-1/2})$. So $R_{1,n} = o_p(1)$. Thus,

$$A_{11,n} = n^{-1/2} \sum_{i=1}^{n} \psi_\tau(\epsilon_i)(\hat{F}_0(Y_i) - F_0(Y_i))\, D_0 c(F_0(Y_i), \boldsymbol{F}(\boldsymbol{x}); \boldsymbol{\theta}^*) + o_p(1). \tag{10}$$

Now, we turn to the second term $A_{12,n}$. Following the same arguments as for $A_{11,n}$, by Assumptions (C1) and (C3), we have

$$\begin{aligned} A_{12,n} &= n^{-1/2} \sum_{i=1}^{n} \psi_\tau(\epsilon_i)(\hat{\boldsymbol{F}}(\boldsymbol{x}) - \boldsymbol{F}(\boldsymbol{x}))^\top \boldsymbol{c}'(F_0(Y_i), \boldsymbol{F}(\boldsymbol{x}); \boldsymbol{\theta}^*) + o_p(1) \\ &= \sqrt{n}(\hat{\boldsymbol{F}}(\boldsymbol{x}) - \boldsymbol{F}(\boldsymbol{x}))^\top \boldsymbol{e}'(\boldsymbol{x}) + o_p(1), \end{aligned} \tag{11}$$

where, in the last equality, we used the weak law of large numbers and Assumption (C1). Similarly, by Assumptions (C2) and (C3), the last term $A_{13,n}$ can be expressed as

$$A_{13,n} = \sqrt{n}(\hat{\boldsymbol{\theta}}_n - \boldsymbol{\theta}^*)^\top \dot{\boldsymbol{e}}(\boldsymbol{x}) + o_p(1). \tag{12}$$

Recollecting the elements (10), (11), (12) and (9) gives

$$A_{1,n} = n^{-1/2} \sum_{i=1}^{n} \psi_\tau(\epsilon_i) c(F_0(Y_i), \boldsymbol{F}(\boldsymbol{x}); \boldsymbol{\theta}^*) + \sqrt{n} V_n + \sqrt{n}(\hat{\boldsymbol{F}}(\boldsymbol{x}) - \boldsymbol{F}(\boldsymbol{x}))^\top \boldsymbol{e}'(\boldsymbol{x}) + \sqrt{n}(\hat{\boldsymbol{\theta}}_n - \boldsymbol{\theta}^*)^\top \dot{\boldsymbol{e}}(\boldsymbol{x}) + o_p(1),$$

where $V_n = n^{-2} \sum_{i,j} h(Y_i, Y_j)$ is a V-statistic, with

$$\begin{aligned} h(Y_i, Y_j) \quad = \quad & \frac{1}{2} \left[\psi_\tau(Y_i - m_\tau(\boldsymbol{x}; \boldsymbol{\theta}^*))(I(Y_j \le Y_i) - F_0(Y_i)) D_0 c(F_0(Y_i), \boldsymbol{F}(\boldsymbol{x}); \boldsymbol{\theta}^*)\right. \\ & \left. + \psi_\tau(Y_j - m_\tau(\boldsymbol{x}; \boldsymbol{\theta}^*))(I(Y_i \le Y_j) - F_0(Y_j)) D_0 c(F_0(Y_j), \boldsymbol{F}(\boldsymbol{x}); \boldsymbol{\theta}^*)\right]. \end{aligned}$$

By Assumption (C0), using Hoeffding's projection method and applying Proposition 2 in Denker and Keller (1983), we get that

$$V_n = n^{-1} \sum_{i=1}^{n} \lambda(Y_i) + o_p(n^{-1/2}), \tag{13}$$

where

$$\lambda(y) = \mathbb{E}\left[\psi_\tau(Y - m_\tau(\boldsymbol{x}; \boldsymbol{\theta}^*))(I(y \le Y) - F_0(Y)) D_0 c(F_0(Y), \boldsymbol{F}(\boldsymbol{x}); \boldsymbol{\theta}^*)\right].$$

Using Assumption (C3)-(i), some easy calculations show that,

$$\begin{aligned}\lambda(y) \;=\; & -\psi_\tau(y - m_\tau(\boldsymbol{x};\boldsymbol{\theta}^*))c(F_0(y), \boldsymbol{F}(\boldsymbol{x}); \boldsymbol{\theta}^*) \\ & - \{I(y \leq m_\tau(\boldsymbol{x};\boldsymbol{\theta}^*)) - F_0(m_\tau(\boldsymbol{x};\boldsymbol{\theta}^*))\}\ c(F_0(m_\tau(\boldsymbol{x};\boldsymbol{\theta}^*)), \boldsymbol{F}(\boldsymbol{x}); \boldsymbol{\theta}^*).\end{aligned}$$

We conclude that $A_{1,n} = n^{-1/2}U_n + o_p(1)$, where U_n is defined in the statement of the lemma.

We now turn to $A_{2,n}(t)$ which can be written as,

$$A_{2,n}(t) = \sum_{i=1}^{n} r(\epsilon_i, t/\sqrt{n})c(F_0(Y_i), \boldsymbol{F}(\boldsymbol{x}); \boldsymbol{\theta}^*) + R_{2,n}(t),$$

where $R_{2,n}(t) = \sum_i r(\epsilon_i, t/\sqrt{n})(c(\hat{F}_0(Y_i), \hat{\boldsymbol{F}}(\boldsymbol{x}); \hat{\boldsymbol{\theta}}) - c(F_0(Y_i), \boldsymbol{F}(\boldsymbol{x}); \boldsymbol{\theta}^*))$. First we show that $R_{2,n}(t) = o_p(1)$. Since, by Assumption (C3), $\max_{1\leq i\leq n} \left| c(\tilde{U}_{0,i}, \tilde{\boldsymbol{U}}_i; \tilde{\boldsymbol{\theta}}_i) - c(F_0(Y_i), \boldsymbol{F}(\boldsymbol{x}); \boldsymbol{\theta}^*) \right| = o_p(1)$, it suffices to prove that $\sum_{i=1}^n r(\epsilon_i, t/\sqrt{n}) = O_p(1)$.

$$\begin{aligned}\mathbb{E}(r(\epsilon_i, t/\sqrt{n})) &= \int_0^{t/\sqrt{n}} [F_0(s + m_\tau(\boldsymbol{x};\boldsymbol{\theta}^*)) - F_0(m_\tau(\boldsymbol{x};\boldsymbol{\theta}^*))]ds \\ &= \int_0^{t/\sqrt{n}} s f_0(m_\tau(\boldsymbol{x};\boldsymbol{\theta}^*) + zs)ds, \;\; \text{for some } z \in [0,1] \\ &= \frac{t^2}{2n} f_0(m_\tau(\boldsymbol{x};\boldsymbol{\theta}^*)) + \int_0^{t/\sqrt{n}} s[f_0(m_\tau(\boldsymbol{x};\boldsymbol{\theta}^*) + zs) - f_0(m_\tau(\boldsymbol{x};\boldsymbol{\theta}^*))]ds \\ &= \frac{t^2}{2n} f_0(m_\tau(\boldsymbol{x};\boldsymbol{\theta}^*)) + o(n^{-1}),\end{aligned}$$

where, in the last equality, we used Assumption (C4). For the variance, observe that,

$$\mathbb{V}ar\left[\sum_{i=1}^n r(\epsilon_i, t/\sqrt{n})\right] \leq \sum_{i=1}^n \mathbb{E}\left[r^2(\epsilon_i, t/\sqrt{n})\right] + 2n\sum_{i=1}^{n-1} |\mathbb{C}ov(r(\epsilon_1, t/\sqrt{n}), r(\epsilon_{i+1}, t/\sqrt{n}))|.$$

Since $r^2(\epsilon_i, t/\sqrt{n}) \leq \frac{|t|}{\sqrt{n}} r(\epsilon_i, t/\sqrt{n})$, $\sum_{i=1}^n \mathbb{E}(r^2(\epsilon_i, t/\sqrt{n})) = O(n^{-1/2})$. Also, by the Cauchy-Schwarz's inequality, we deduce that if $n - 1 \geq k_n$,

$$n\sum_{i=1}^{k_n} |\mathbb{C}ov(r(\epsilon_1, t/\sqrt{n}), r(\epsilon_{i+1}, t/\sqrt{n}))| \leq n\sum_{i=1}^{k_n} \mathbb{E}(r^2(\epsilon_i, t/\sqrt{n})) = O(k_n/\sqrt{n}),$$

for any integer $k_n \to \infty$. On the other hand, by Assumption (C0), using Billingsley's inequality, see e.g. Lemma 3 in Doukhan (1994), we also have that, for sufficiently large n,

$$n \sum_{i=k_n+1}^{n-1} |\mathbb{C}ov(r(\epsilon_1, t/\sqrt{n}), r(\epsilon_{i+1}, t/\sqrt{n}))| \le 16t^2 \sum_{i>k_n} \beta(i) = O(1) \sum_{i>k_n} i^{-\nu} = O(k_n^{1-\nu}) = o(1),$$

since $\nu > 1$. So, taking $k_n = \lfloor n^\alpha \rfloor$, for some $0 < \alpha < 1/2$, yields $\mathbb{V}ar(\sum_{i=1}^n r(\epsilon_i, t/\sqrt{n})) = o(1)$. We conclude that, $\sum_{i=1}^n r(\epsilon_i, t/\sqrt{n}) = O_p(1)$.

By similar arguments, using Assumption (C0), (C3) and (C5), one can also show that

$$\mathbb{E}\left[\sum_{i=1}^n r(\epsilon_i, t/\sqrt{n}) c(F_0(Y_i), \boldsymbol{F}(\boldsymbol{x}); \boldsymbol{\theta}^*)\right] = \frac{t^2}{2} f_{\boldsymbol{\theta}^*}(m_\tau(\boldsymbol{x}; \boldsymbol{\theta}^*)|\boldsymbol{x}) c_{\boldsymbol{X}}(\boldsymbol{F}(\boldsymbol{x}); \boldsymbol{\theta}^*) + o(1), \text{ and}$$
$$\mathbb{V}ar\left[\sum_{i=1}^n r(\epsilon_i, t/\sqrt{n}) c(F_0(Y_i), \boldsymbol{F}(\boldsymbol{x}); \boldsymbol{\theta}^*)\right] = o(1).$$

This implies that

$$\begin{aligned} A_{2,n}(t) &= \frac{1}{2} t^2 f_{\boldsymbol{\theta}^*}(m_\tau(\boldsymbol{x}; \boldsymbol{\theta}^*)|\boldsymbol{x}) c_{\boldsymbol{X}}(\boldsymbol{F}(\boldsymbol{x}); \boldsymbol{\theta}^*) + o_p(1) \\ &= \frac{1}{2} t^2 f_0(m_\tau(\boldsymbol{x}; \boldsymbol{\theta}^*)) c(F_0(m_\tau(\boldsymbol{x}; \boldsymbol{\theta}^*)), \boldsymbol{F}(\boldsymbol{x}); \boldsymbol{\theta}^*) + o_p(1), \end{aligned}$$

which concludes the proof of Lemma 6.1. □

Proof of Theorem 3.1.

First, observe that, by definition,

$$\begin{aligned} \arg\min_t A_n(t) &= \arg\min_t \left[\sum_{i=1}^n \rho_\tau(Y_i - (m_\tau(\boldsymbol{x}; \boldsymbol{\theta}^*) + t/\sqrt{n})) c(\hat{F}_0(Y_i), \hat{\boldsymbol{F}}(\boldsymbol{x}); \hat{\boldsymbol{\theta}})\right] \\ &= \sqrt{n}(\hat{m}_\tau(\boldsymbol{x}) - m_\tau(\boldsymbol{x}; \boldsymbol{\theta}^*)). \end{aligned}$$

Also, since ρ_τ is a convex function and $c(\hat{F}_0(Y_i), \hat{\boldsymbol{F}}(\boldsymbol{x}); \hat{\boldsymbol{\theta}}) \ge 0$, A_n is a convex function of t. Thanks to Lemma 6.1 and the quadratic approximation lemma (Basic Corollary in Hjort and Pollard (1993)) with $U_n = O_p(1)$, we have

$$\sqrt{n}(\hat{m}_\tau(\boldsymbol{x}) - m_\tau(\boldsymbol{x};\boldsymbol{\theta}^*)) = \frac{1}{f_0(m_\tau(\boldsymbol{x};\boldsymbol{\theta}^*))c(F_0(m_\tau(\boldsymbol{x};\boldsymbol{\theta}^*)),\boldsymbol{F}(\boldsymbol{x});\boldsymbol{\theta}^*)} U_n + o_p(1),$$

which is the desired result. □

References

K. Aas, C. Czado, A. Frigessi, and H. Bakken. Pair-copula constructions of multiple dependence. *Insurance: Mathematics and Economics*, 44:182–198, 2009.

E. Bouyé and M. Salmon. Dynamic copula quantile regressions and tail area dynamic dependence in forex markets. *The European Journal of Finance*, 15:721–750, 2009.

E. C. Brechmann. *Truncated and simplified regular vines and their applications*. PhD thesis, Technische Universität München, 2010.

X. Chen and Y. Fan. Estimation of copula-based semiparametric time series models. *Journal of Econometrics*, 130:307–335, 2006.

X. Chen, R. Koenker, and Z. Xiao. Copula-based nonlinear quantile autoregression. *Econometrics Journal*, 12, 2009.

M. Denker and G. Keller. On U-statistics and v. Mises' statistics for weakly dependent processes. *Z. Wahrsch. Verw. Gebiete*, 64:505–522, 1983.

J. Dißmann, E. C. Brechmann, C. Czado, and D. Kurowicka. Selecting and estimating regular vine copulae and application to financial returns. *Computational Statistics and Data Analysis*, 59:52–69, 2013.

P. Doukhan. *Mixing*. Lecture Notes in Statistics. Springer-Verlag, New York, 1994.

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

N. L. Hjort and D. Pollard. Asymptotics for minimisers of convex process. Technical report, Yale University, 1993.

R. Ibragimov. Copula-based characterizations for higher order Markov processes. *Econometric Theory*, 25:819–846, 2009.

C. A. J. Klaassen and J. A. Wellner. Efficient estimation in the bivariate normal copula model: normal margins are least favourable. *Bernoulli*, 3:55–77, 1997.

R. Koenker and G. Bassett. Regression quantiles. *Econometrica*, 46:33–50, 1978.

I. Kojadinovic and J. Yan. A goodness-of-fit test for multivariate multiparameter copulas based on multiplier central limit theorems. *Statistics and Computing*, 21:17–30, 2011.

H. Noh, A. El Ghouch, and T. Bouezmarni. Copula-based regression estimation and inference. *Journal of the American Statistical Association*, in press, 2013.

P. Phillips and P. Perron. Testing for a unit root in time series regression. *Biometrika*, 75:335–346, 1988.

R Development Core Team. *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria, 2011. URL `http://www.R-project.org/`.

B. Rémillard, N. Papageogiou, and F. Soustra. Copula-based semiparametric models for multivariate time series. *Journal of Multivariate Analysis*, 110:30–42, 2012.

E. Rio. *Théorie asymptotique des processus aléatoires faiblement dépendants*. Springer-Verlag, 2000.

P. Silvapulle, G. Kim, and M. J. Silvapulle. Robustness of a semiparametric estimator of a copula. Econometric Society 2004 Australasian Meetings 317, Econometric Society, 2004.

B. W. Silverman and G. A. Young. The bootstrap: To smooth or not to smooth? *Biometrika*, 74(3): 469–479, 1987.

M. Sklar. Fonctions de répartition à n dimensions et leurs marges. *Publ. Inst. Statist. Univ. Paris*, 8: 229–231, 1959.

H. Tsukahara. Semiparametric estimation in copula models. *Canadian Journal of Statistics*, 33: 357–375, 2005.

L. Zhu, M. Huang, and R. Li. Semiparametric quantile regression with high-dimensional covariates. *Statistica Sinica*, 22:1379–1401, 2012.

DGP	$\boldsymbol{x}$	τ	$\alpha = 0.05$			$\alpha = 0.1$		
			$n = 50$	$n = 100$	$n = 200$	$n = 50$	$n = 100$	$n = 200$
A.1	0.00	0.2	0.946	0.941	0.946	0.905	0.901	0.953
		0.5	0.973	0.964	0.964	0.931	0.926	0.926
		0.8	0.946	0.950	0.942	0.907	0.896	0.882
	-1.64	0.2	0.960	0.950	0.953	0.930	0.906	0.911
		0.5	0.959	0.959	0.960	0.915	0.923	0.918
		0.8	0.924	0.938	0.938	0.881	0.894	0.892
B.1	0.00	0.2	0.947	0.951	0.948	0.898	0.895	0.901
		0.5	0.956	0.946	0.958	0.913	0.899	0.911
		0.8	0.952	0.948	0.962	0.912	0.888	0.918
	-1.64	0.2	0.846	0.879	0.891	0.792	0.825	0.829
		0.5	0.931	0.957	0.953	0.887	0.903	0.898
		0.8	0.961	0.958	0.959	0.932	0.906	0.914
B.2	(0.00,0.00)	0.2	0.892	0.904	0.924	0.869	0.837	0.885
		0.5	0.898	0.917	0.922	0.828	0.860	0.865
		0.8	0.886	0.918	0.932	0.842	0.868	0.876
	(1.53,1.53)	0.2	0.928	0.930	0.930	0.909	0.884	0.886
		0.5	0.943	0.940	0.927	0.902	0.887	0.879
		0.8	0.945	0.939	0.941	0.915	0.905	0.882

Table 2: ECPs of the confidence interval for $m_\tau(\boldsymbol{x})$ in the *iid* setting based on the asymptotic representation in Corollary 3.2.

DGP	Method	$\boldsymbol{x}$	τ	$\alpha = 0.05$			$\alpha = 0.1$		
				$n = 50$	$n = 100$	$n = 200$	$n = 50$	$n = 100$	$n = 200$
B.2	TRUE	(0.00,0.00)	0.2	0.947	0.935	0.959	0.902	0.883	0.893
			0.5	0.936	0.925	0.947	0.867	0.891	0.885
			0.8	0.933	0.954	0.943	0.885	0.900	0.890
	BT	(0.00,0.00)	0.2	0.950	0.946	0.948	0.896	0.906	0.906
			0.5	0.950	0.948	0.952	0.884	0.890	0.898
			0.8	0.932	0.940	0.960	0.876	0.886	0.878

Table 3: ECPs of the confidence interval for $m_\tau(\boldsymbol{x})$ based on the true asymptotic representation and the bootstrap approach.

DGP	$\boldsymbol{x}$	τ	$\alpha = 0.05$		$\alpha = 0.1$	
			$n = 100$	$n = 200$	$n = 100$	$n = 200$
A.2	0.00	0.2	0.948	0.958	0.926	0.920
		0.5	0.962	0.954	0.934	0.924
		0.8	0.948	0.956	0.914	0.906
B.4	(0.56,0.00,-0.27)	0.2	0.952	0.952	0.924	0.902
		0.5	0.954	0.952	0.920	0.898
		0.8	0.956	0.952	0.920	0.908

Table 4: ECPs of the confidence interval for $m_\tau(\boldsymbol{x})$ in the time series setting based on the bootstrap approach.

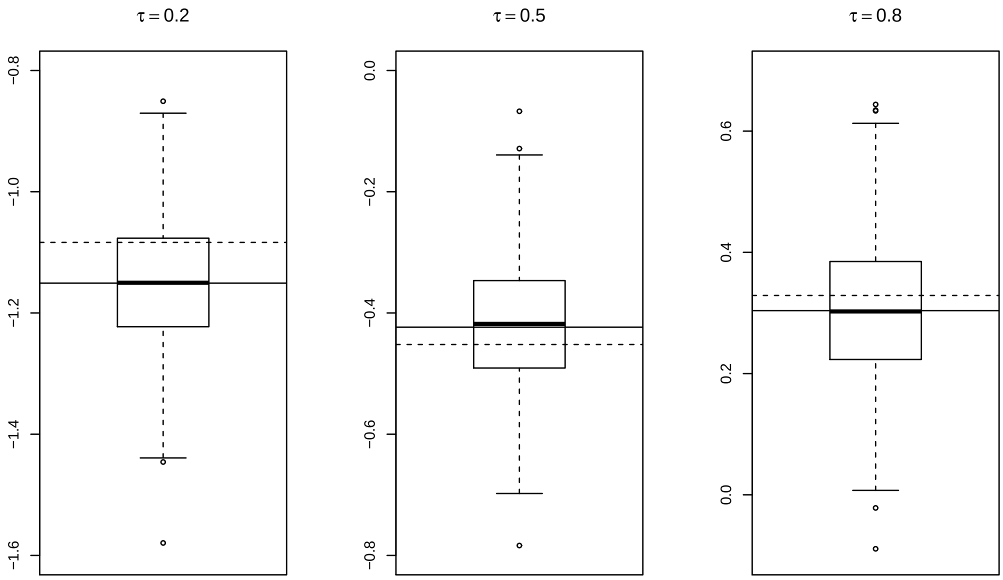

Figure 1: Boxplots of $\hat{m}_\tau(x_1)$ at $x_1 = F_1^{-1}(0.2) = -0.8416$ for different quantile levels ($\tau = 0.2, 0.5$ and 0.8). The horizontal solid line represents $m_\tau(x_1; \rho^*)$ and the dotted line represents $m_\tau(x_1)$.

N	τ	$\hat{m}_{tc}$	$\hat{m}_{uc}$	$\hat{m}_{np}$	$\hat{m}_{si}$	$\hat{m}_{nl}$
50	0.2	2.847	3.377	5.987	4.366	1.870
	0.5	2.561	2.941	6.265	3.773	1.406
	0.8	3.052	3.284	5.517	4.499	1.685
100	0.2	1.280	1.569	3.244	2.252	0.864
	0.5	1.136	1.384	2.829	1.701	0.701
	0.8	1.356	1.577	3.176	2.259	0.806
200	0.2	0.634	0.796	1.778	1.029	0.370
	0.5	0.559	0.709	1.540	1.005	0.307
	0.8	0.660	0.795	1.765	1.211	0.428

Table 5: $1000 \times IMSE$ for DGP B.2

N	τ	$\hat{m}_{tc}$	$\hat{m}_{uc}$	$\hat{m}_{np}$	$\hat{m}_{si}$	$\hat{m}_{nl}$
50	0.2	2.636	4.035	6.565	4.058	1.092
	0.5	2.555	3.808	6.329	3.687	0.931
	0.8	3.142	4.085	6.448	4.374	1.092
100	0.2	1.281	1.709	3.690	1.724	0.542
	0.5	1.239	1.608	3.102	1.483	0.453
	0.8	1.451	1.740	4.111	1.789	0.572
200	0.2	0.614	0.800	1.863	1.028	0.259
	0.5	0.579	0.725	1.659	0.888	0.207
	0.8	0.662	0.784	2.022	0.983	0.268

Table 6: $1000 \times IMSE$ for DGP B.3

N	τ	$\hat{m}_{uc}$	$\hat{m}_{np}$	$\hat{m}_{si}$	$\hat{m}_{nl}$
50	0.2	4.106	4.167	3.609	0.419
	0.5	3.835	3.904	3.046	0.300
	0.8	4.449	4.918	4.445	0.427
100	0.2	2.422	3.286	2.393	0.190
	0.5	2.037	3.116	2.174	0.146
	0.8	2.416	3.800	2.597	0.193
200	0.2	1.562	2.206	1.881	0.093
	0.5	1.286	1.975	1.157	0.087
	0.8	1.539	2.226	1.409	0.091

Table 7: $1000 \times IMSE$ for DGP C

	M1	M2	M3	M4	M5	M6
PRED $\times 10^5$	24.659	24.610	21.811	21.131	21.634	21.107

Table 8: PRED $\times 10^5$ for each prediction method

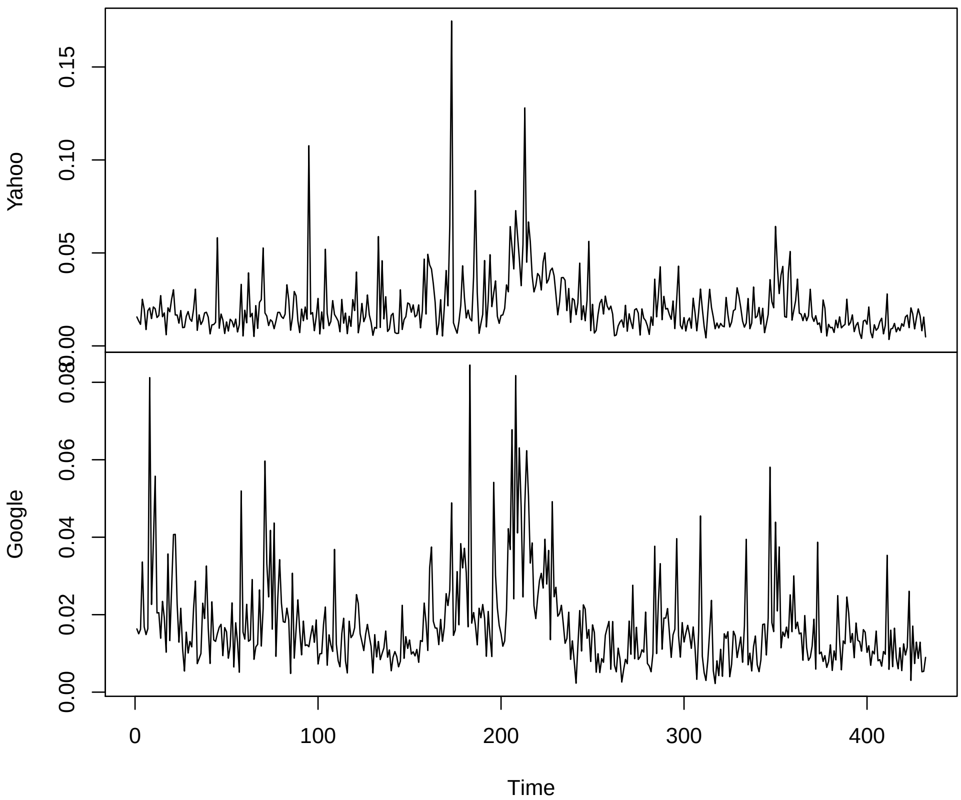

Figure 2: Plot of the historical volatilities for both companies.