Available online at www.sciencedirect.com
SCIENCE DIRECT
ELSEVIER
Journal of Econometrics 130 (2006) 307–335
JOURNAL OF Econometrics
www.elsevier.com/locate/jeconom

Estimation of copula-based semiparametric time series models

Xiaohong Chen[a,*,1], Yanqin Fan[b,2]

[a]*Department of Economics, New York University, 269 Mercer Street, New York, NY 10003, USA*
[b]*Department of Economics, Vanderbilt University, VU Station B 351819, 2301 Vanderbilt Place, Nashville, TN 37235, USA*

Accepted 21 March 2005
Available online 17 May 2005

Abstract

This paper studies the estimation of a class of copula-based semiparametric stationary Markov models. These models are characterized by nonparametric marginal distributions and parametric copula functions, while the copulas capture all the scale-free temporal dependence of the processes. Simple estimators of the marginal distribution and the copula parameter are provided, and their asymptotic properties are established under easily verifiable conditions. These results are used to obtain root-n consistent and asymptotically normal estimators of important features of the transition distribution such as the (nonlinear) conditional moments and conditional quantiles. The semiparametric conditional quantile estimators are automatically monotonic across quantiles, which is attractive for portfolio conditional value-at-risk calculations.

JEL classification: C14; C22

Keywords: Copula; Nonlinear Markov models; Semiparametric estimation; Conditional moment; Conditional quantile

*Corresponding author. Tel.: +212 998 8970; fax: +212 995 4186.
E-mail addresses: xiaohong.chen@nyu.edu (X. Chen), yanqin.fan@vanderbilt.edu (Y. Fan).
[1]Supported by the National Science Foundation and the C.V. Starr Center at New York University.
[2]Supported by the National Science Foundation.

0304-4076/$ - see front matter
doi:10.1016/j.jeconom.2005.03.004

1. Introduction

A copula function is a multivariate distribution function with standard uniform marginals. By Sklar's (1959) theorem, one can always model any multivariate distribution by modeling its marginal distributions and its copula function separately, where the copula captures all the scale-free dependence in the multivariate distribution. Because of this flexibility, copulas have gained popularity in the finance and insurance community[3] in the past few years, where modeling and estimating the dependence structure between several univariate time series are of great interest; see Frees and Valdez (1998) and Embrechts et al. (2002) for reviews.[4]

While modeling the contemporaneous dependence between several univariate time series is important, it is also important to model the temporal dependence of a univariate (nonlinear) time series. In the copula approach to univariate time series modeling, the finite dimensional distributions of the time series are generated by copulas. By coupling different marginal distributions with different copula functions, copula-based time series models are able to model a wide variety of marginal behaviors (such as skewness and fat tails) and dependence properties (such as clusters, positive or negative tail dependence). Darsow et al. (1992) provide a necessary and sufficient condition for a copula-based time series to be a Markov process. Joe (1997) proposes a class of parametric stationary Markov models based on parametric copulas and parametric marginal distributions, and provides an application to daily air quality measurements.

In this paper, we study a class of univariate copula-based semiparametric stationary Markov models, in which copulas are parameterized, but the invariant (or marginal) distributions are left unspecified. Models in this class are completely characterized by two unknown parameters: the copula dependence parameter α^* (i.e., the finite-dimensional parameter in the copula function); and the invariant (or marginal) distribution function $G^*(\cdot)$. The unknown marginal distribution can be estimated by any one of the existing nonparametric methods, including the rescaled empirical distribution function and the kernel smoothed estimator of the distribution function. To estimate the copula dependence parameter α^*, we extend the two-step estimator proposed for bivariate copula models with i.i.d. observations[5] to our class of univariate copula-based semiparametric time series models. We establish the consistency and $\sqrt{n}$-asymptotic normality of the proposed estimators of (G^*, α^*) under easily verifiable conditions.

[3]Copulas have also proven to be useful in microeconometrics, see e.g., Lee (1982, 1983) on sample selection models, and Heckman and Honoré (1989) on competing risk models.

[4]Specific applications include Rosenberg (1999) and Cherubini and Luciano (2002) for multivariate option pricing, Hull and White (1998) and Embrechts et al. (2003) for portfolio Value-at-Risk, Li (2000) and Frey and McNeil (2001) for default and credit risk, Costinot et al. (2000) and Hu (2002) for contagion, to name only a few.

[5]Genest et al. (1995) and Shih and Louis (1995) study this approach independently, while the latter paper allows the i.i.d. observations generated from a bivariate copula model with random censoring. Both papers and Hu (1998) present the asymptotic normality of their two-step estimators for i.i.d. observations.

In economic and financial applications, estimating the dependence parameter is often not the ultimate aim; one is often interested in estimating or forecasting certain features of the transition distribution of the time series such as the (nonlinear) conditional moment and conditional quantile functions. For example, estimating the conditional value-at-risk (VaR) of portfolios of assets, or equivalently the conditional quantile of portfolios of assets, has become routine in risk management, see e.g., Duffie and Pan (1997), Gourieroux and Jasiak (2002) and Engle and Manganelli (2002). This can be easily accomplished for copula-based semiparametric time series models, as the transition distribution of a time series in this class is completely characterized by the marginal distribution and the copula dependence parameter. Given the estimators of the marginal distribution and the copula dependence parameter, one can easily construct an estimator of the transition distribution of the time series and hence estimators of any (nonlinear) conditional moment and conditional quantile functions. Moreover, given the joint asymptotic distribution of the estimators of (G^*, α^*), one can easily establish the $\sqrt{n}$-consistency and asymptotic normality of the resulting estimators of the nonlinear conditional moment and conditional quantile functions. It is interesting to note that although the conditional distribution of a copula-based semiparametric stationary Markov model depends on the unknown marginal distribution, estimators of the nonlinear conditional moment and conditional quantiles are still $\sqrt{n}$-consistent and asymptotically normal. Moreover, the estimated conditional quantile functions are automatically monotonic across different quantiles, which is attractive for portfolio conditional VaR calculations.

In an unpublished working paper that is independently done from ours, Bouyé et al. (2002) also propose to use parametric copulas to model nonlinear autoregressive dependence of time series and provide applications to financial returns and transactions based forex data.[6] They briefly mention the two-step procedure of Genest et al. (1995)[7] for estimating the copula dependence parameter without establishing its large sample properties. Moreover, they did not study the estimation of any conditional moment and conditional quantile functions of a copula-based semiparametric time series model.

The rest of this paper is organized as follows. In Section 2, we present the class of copula-based semiparametric time series models considered in this paper, and study their β-mixing property. We also point out the close relation between these models and the generalized semiparametric regression transformation models. In Section 3, we introduce the semiparametric estimator of the copula dependence parameter and estimators of the conditional moment and conditional quantile functions. Section 4 establishes the asymptotic properties of the estimators proposed in Section 3. In Section 5, we verify the conditions for the consistency and asymptotic normality of the two-step estimator for three widely used copulas. Section 6 concludes with discussions of several extensions. All the proofs are relegated to the Appendix.

[6]Recently and independently, Gagliardini and Gourieroux (2002) have proposed a class of stationary Markov duration time series models with proportional hazard and discussed its link to copulas.

[7]It is referred to as the canonical maximum likelihood (CML) estimation method in Bouyé et al. (2002).

2. Copula-based Markov models of order 1

Let $\{Y_t\}$ be a stationary Markov process of order 1 with continuous state space. Then its probabilistic properties are completely determined by the joint distribution function of Y_{t-1} and Y_t, $H(y_1, y_2)$ (say). By Sklar's theorem, one can express $H(y_1, y_2)$ in terms of the marginal distribution function of Y_t and the copula function of Y_{t-1} and Y_t. This suggests the copula approach as an alternative to modeling a stationary Markov process: instead of specifying the joint distribution function of Y_{t-1} and Y_t directly, one specifies the marginal distribution function of Y_t and the copula function of Y_{t-1} and Y_t. The advantage of the copula approach is that one has the freedom to choose the marginal distribution and the copula function separately; the former characterizes the marginal behavior such as the fat-tails of the time series $\{Y_t\}$, while the latter characterizes the scale-free temporal dependence property such as nonlinear, asymmetric dependence, of the time series.

In this paper, we will work with the class of copula-based, semiparametric time series models in which the marginal distribution is left unspecified, but the copula function has a parametric form.

Assumption 1. $\{Y_t : t = 1, \ldots, n\}$ is a sample of a stationary first-order Markov process generated from $(G^*(\cdot), C(\cdot, \cdot; \alpha^*))$, where $G^*(\cdot)$ is the true invariant distribution which is absolutely continuous with respect to Lebesgue measure on the real line; $C(\cdot, \cdot; \alpha^*)$ is the true parametric copula for (Y_{t-1}, Y_t) up to unknown value α^*, is absolutely continuous with respect to Lebesgue measure on $[0, 1]^2$, and is neither the Fréchet–Hoeffding upper nor lower bound.

It is known that if the copula of Y_{t-1} and Y_t is either the Fréchet–Hoeffding upper bound ($C(u_1, u_2) = \min(u_1, u_2)$) or the lower bound ($C(u_1, u_2) = \max(u_1 + u_2 - 1, 0)$), then Y_t is almost surely a monotonic function of Y_{t-1}; the resulting time series is deterministic and under stationarity, $Y_t = Y_{t-1}$ for the upper bound and $Y_t = G^{*-1}(1 - G^*(Y_{t-1}))$ for the lower bound. Assumption 1 rules out these two cases.

Remark. One standard approach that has been used to construct semiparametric time series models is to specify a parametric conditional density of Y_t given Y_{t-1} with an unspecified marginal distribution of Y_{t-1}. Our approach specifies the conditional density of Y_t given Y_{t-1} semiparametrically via

$$h^*(y_t|y_{t-1}) = g^*(y_t)c(G^*(y_{t-1}), G^*(y_t); \alpha^*), \tag{2.1}$$

where $h^*(\cdot|y_{t-1})$ is the true conditional density function of Y_t given $Y_{t-1} = y_{t-1}$, $c(\cdot, \cdot; \alpha^*)$ is the copula density of $C(\cdot, \cdot; \alpha^*)$, and $g^*(\cdot)$ is the density of the marginal distribution $G^*(\cdot)$, which is unspecified. One obvious advantage of the copula approach over the standard approach is to separate out the temporal dependence structure[8] from the marginal behavior. The copula approach also allows us to take advantage of numerous existing parametric copulas to construct new conditional

[8]In this paper, we refer the temporal dependence structure of a time series as that characterized by the copula function which is invariant to any increasing transformation of the time series.

density functions via (2.1), see Joe (1997) and Nelsen (1999) for expressions of many commonly used parametric copulas.

We note that under Assumption 1, the transformed process, $\{U_t : U_t \equiv G^*(Y_t)\}$, is a stationary parametric Markov process of order 1 in which the joint distribution of U_t and U_{t-1} is given by the copula $C(u_0, u_1; \alpha^*)$, and the conditional density of U_t given $U_{t-1} = u_0$ is $f_{U_t|U_{t-1}=u_0}(u) = c(u_0, u; \alpha^*)$. This implies that Assumption 1 is consistent with the following *generalized semiparametric regression transformation* model:

$$\Lambda_{1,\theta_1}(G^*(Y_t)) = \Lambda_{2,\theta_2}(G^*(Y_{t-1})) + \varepsilon_t, \quad \mathrm{E}\{\varepsilon_t | Y_{t-1}\} = 0, \tag{2.2}$$

where $\Lambda_{1,\theta_1}(\cdot)$ is a parametric increasing function, $\Lambda_{2,\theta_2}(u_0) \equiv \mathrm{E}\{\Lambda_{1,\theta_1}(G^*(Y_t)) | G^*(Y_{t-1}) = u_0\}$, and the conditional density of ε_t given $G^*(Y_{t-1}) = u_0$ satisfies:

$$f_{\varepsilon_t|G^*(Y_{t-1})=u_0}(\varepsilon) = c(u_0, \Lambda_{1,\theta_1}^{-1}(\varepsilon + \Lambda_{2,\theta_2}(u_0)); \alpha^*) \div \frac{\mathrm{d}\Lambda_{1,\theta_1}(\varepsilon + \Lambda_{2,\theta_2}(u_0))}{\mathrm{d}\varepsilon}.$$

It is clear that the functional form of $\Lambda_{2,\theta_2}(\cdot)$ is completely pinned down by $\Lambda_{1,\theta_1}(\cdot)$ and the copula density $c(\cdot, \cdot; \alpha^*)$:

$$\Lambda_{2,\theta_2}(u_0) \equiv \mathrm{E}\{\Lambda_{1,\theta_1}(U_t) | U_{t-1} = u_0\} = \int_0^1 \Lambda_{1,\theta_1}(u) \times c(u_0, u; \alpha^*)\, \mathrm{d}u.$$

A special case of (2.2) is given by $\Lambda_{1,\theta_1}(u_1) = u_1$, the identity mapping. Then $\Lambda_{2,\theta_2}(u_0) = \mathrm{E}(U_t | U_{t-1} = u_0) = 1 - \int_0^1 \frac{\partial C(u_0, u; \alpha^*)}{\partial U_{t-1}} \mathrm{d}u$. For some commonly used copulas including the Plackett copula and the Farlie–Gumbel–Morgenstern (F–G–M) copula, $\mathrm{E}(U_t | U_{t-1} = u_0)$ has simple expressions, see e.g., Hutchinson and Lai (1990). However, for many copulas, transformations $\Lambda_{1,\theta_1}(u_1)$ that are different from the identity mapping will lead to simpler generalized semiparametric regression transformation models (2.2).

Example 1. Suppose the copula of Y_{t-1} and Y_t is the Gaussian copula:

$$C(v_1, v_2; \alpha) = \Phi_\alpha(\Phi^{-1}(v_1), \Phi^{-1}(v_2)), \tag{2.3}$$

where $0 \leqslant v_1, v_2 \leqslant 1$, $\Phi(\cdot)$ is the distribution function of a standard normal random variable, and $\Phi_\alpha(\cdot, \cdot)$ is the distribution function of the bivariate normal distribution with means zero, variances 1, and correlation coefficient α. Then the process $\{Y_t\}$ satisfies

$$\Phi^{-1}(G^*(Y_t)) = \alpha \Phi^{-1}(G^*(Y_{t-1})) + \varepsilon_t, \tag{2.4}$$

where $\varepsilon_t \sim N(0, 1 - \alpha^2)$ and is independent of Y_{t-1}. If in addition the marginal distribution $G^*(\cdot)$ is the standard normal, then $\{Y_t\}$ is a linear AR(1) process. By allowing $G^*(\cdot)$ to be non-normal such as Student's t, (2.4) is able to generate first order Markov processes characterized by the Gaussian copula, but non-normal marginal distributions.

It is known that the Gaussian copula is symmetric and has no tail dependence and hence cannot be used to model economic and financial time series exhibiting

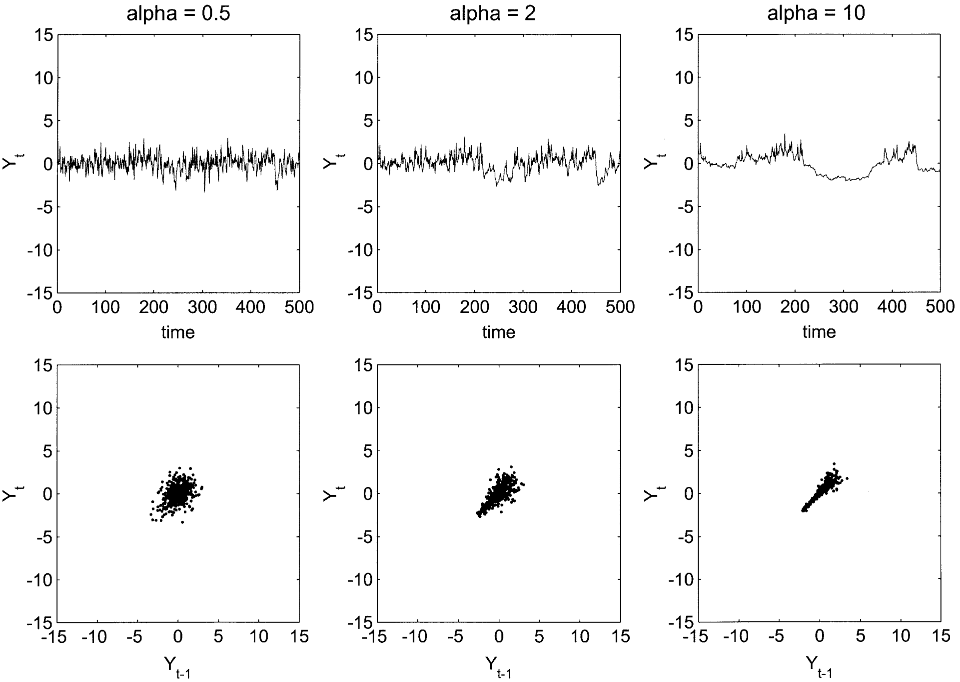

Fig. 1.

complicated nonlinear asymmetric dependence and clusters of large and/or small values. Joe (1997) and Nelsen (1999) provide many non-Gaussian copulas that might be used for this purpose.

Example 2. The Clayton copula given by

$$C(u_1, u_2; \alpha) = [u_1^{-\alpha} + u_2^{-\alpha} - 1]^{-1/\alpha}, \quad \text{where} \quad \alpha > 0, \tag{2.5}$$

has the lower tail dependence parameter $\tau^L = 2^{-1/\alpha}$ and the upper tail dependence parameter[9] $\tau^{\mathrm{U}} = 0$. The lower tail dependence of the Clayton copula increases as α increases. When coupled with fat-tailed marginal distributions such as the Student's t distribution, this class of models can generate time series with clusters of small values and hence provide alternative models for economic and financial time series that do exhibit such clusters. To illustrate, Figs. 1 and 2 present time series plots[10] and the corresponding scatter plots of realizations of time series models with the Clayton

[9]The coefficients of lower and upper tail dependence of a bivariate copula C are defined as: $\tau^L = \lim_{q\to 0}[C(q,q)/q]$ and $\tau^{\mathrm{U}} = \lim_{q\to 1}[\{1 - 2q + C(q,q)\}/(1-q)]$.

[10]The realizations are generated by a modification of the conditional approach described in Nelsen (1999) to time series models. Alternative algorithms are available for generating random variables from specific copulas; see Devroye (1986), Johnson (1987), and Nelsen (1999).

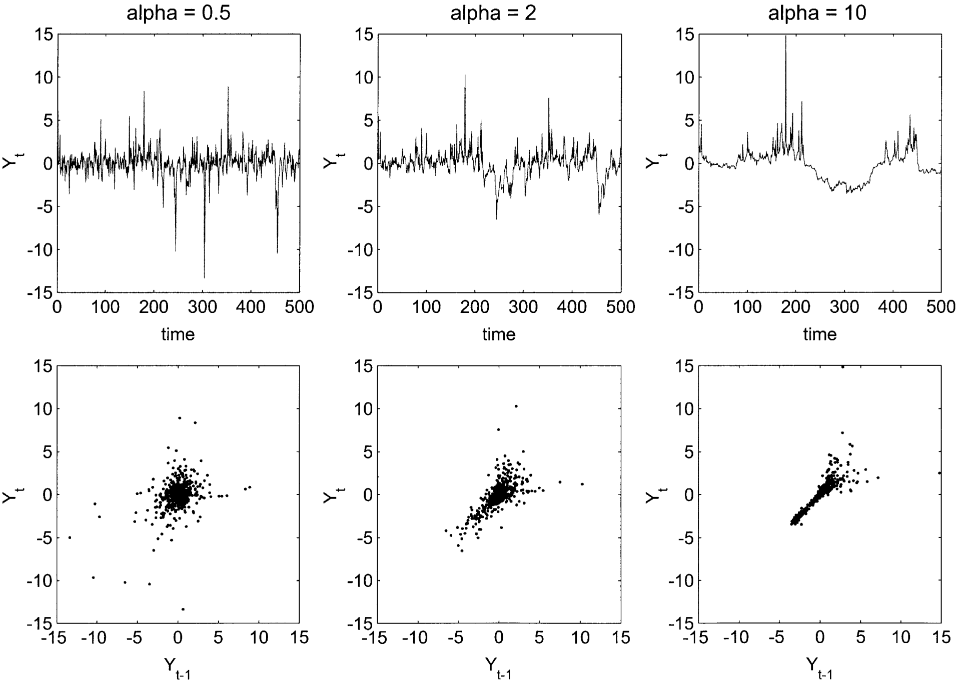

Fig. 2.

copula with $\alpha = (0.5, 2, 10)$ and the marginal distributions given by the standard normal distribution (Fig. 1) and the Student's t distribution with degrees of freedom 3 (Fig. 2), respectively. These figures demonstrate that: (1) the Clayton copula produces time series with asymmetric dependence structure and the degree of asymmetry becomes stronger as α increases; (2) as α increases, the lower tail dependence increases leading to smooth time series plots corresponding to small realizations; (3) coupled with fat-tailed marginal distributions such asthe Student's t distribution with 3 degrees of freedom, the Clayton copula with large α produces clusters of small values.

To close this section, we present the β-mixing property of a copula-based Markov process $\{Y_t\}$ below; see e.g., Bradley (1986) for the definition of β-mixing. A real-valued function Λ is called *norm-like* if the closure of the set $\{x : \Lambda(x) \leqslant B\}$ is compact for each $B > 0$.

Proposition 2.1. *Under Assumption* 1, *if* $c(u_1, u_2; \alpha^*)$ *is positive on* $(0,1)^2$, *then* (i) *and* (ii) *hold*:

(i) *If there are constants* $0 < \lambda < 1$, $0 < d < \infty$, *a norm-like function* $\Lambda(\cdot) \geqslant 1$, *and a small set* $\mathbf{K}$ *such that* $\int_0^1 \Lambda(u) \times c(U_{t-1}, u; \alpha^*) \mathrm{d}u \leqslant \lambda \Lambda(U_{t-1}) + d 1_{\mathbf{K}}(U_{t-1})$, *then* $\{Y_t\}$ *is* β*-mixing with the exponential decay rate*: $\beta_t \leqslant const \times \exp\{-at\}$ *for some* $a > 0$;

(ii) *If there are constants* $\lambda \in [0,1)$, $0<a,d<\infty$, *a norm-like function* $\Lambda(\cdot)\geqslant 1$, *and a small set* $\mathbf{K}$ *such that* $\int_0^1 \Lambda(u)\times c(U_{t-1},u;\alpha^*)\mathrm{d}u \leqslant \Lambda(U_{t-1}) - a[\Lambda(U_{t-1})]^{\lambda} + d1_{\mathbf{K}}(U_{t-1})$, *then* $\{Y_t\}$ *is* β*-mixing with the polynomial decay rate* β_t: $\beta_t(1+t)^{\lambda/(1-\lambda)} \to 0$ *as* $t\to\infty$.

The assumption that $c(u_1,u_2;\alpha^*)$ is positive on $(0,1)^2$ ensures that any process satisfying Assumption 1 with copula density given by $c(u_1,u_2;\alpha^*)$ is β-mixing, since any strictly stationary, recurrent, aperiodic Markov process is β-mixing, albeit the β-mixing decay rate could be very slow (see e.g., Bradley, 1986). The conditions in Proposition 2.1 on the copula are sufficient to ensure that time series with such a copula is β-mixing with at least a polynomial decay rate. Unlike many first-order nonlinear stationary Markov models for which conditions that ensure β-mixing with certain decay rates involve the invariant distributions (see e.g. Chen et al., 1998 for diffusion models), conditions for β-mixing in Proposition 2.1 do not depend on the invariant distribution G^*, but only depend on the copula specification. For instance, by applying Proposition 2.1(i) to Example 1, one can easily verify that the time series $\{Y_t\}$ generated by the Gaussian copula is β-mixing with the exponential decay rate as long as $|\alpha|<1$, regardless of its marginal distribution.

3. Estimation

In this section we first present estimators of model parameters (G^*,α^*) and then introduce estimators of the conditional moment and conditional quantile functions of Y_t given Y_{t-1}.

3.1. Estimation of model parameters

A semiparametric copula-based time series model is completely determined by (G^*,α^*). The unknown marginal distribution G^* can be estimated by $G_n(\cdot)$, the rescaled empirical distribution function defined as

$$G_n(y) = \frac{1}{n+1}\sum_{t=1}^{n} I\{Y_t \leqslant y\}. \tag{3.1}$$

Under Assumption 1, the true joint distribution function of Y_{t-1} and Y_t is of a semiparametric form: $H^*(y_1,y_2) = C(G^*(y_1),G^*(y_2);\alpha^*)$ and the conditional density of Y_t given Y_{t-1} is $h^*(Y_t|Y_{t-1}) = g^*(Y_t)c(G^*(Y_{t-1}),G^*(Y_t);\alpha^*)$. Hence, if the marginal distribution $G^*(\cdot)$ is completely known, then the log-likelihood function is given by

$$L(\alpha) = \frac{1}{n}\sum_{t=1}^{n}\log g^*(Y_t) + \frac{1}{n}\sum_{t=2}^{n}\log c(G^*(Y_{t-1}),G^*(Y_t);\alpha). \tag{3.2}$$

Ignoring the first term and replacing G^* with G_n in the second term on the right-hand side of (3.2) motivate the semiparametric estimator $\tilde{\alpha}$ of α^*:

$$\tilde{\alpha} = \operatorname{argmax}_{\alpha}\tilde{L}(\alpha), \quad \tilde{L}(\alpha) = \frac{1}{n}\sum_{t=2}^{n}\log c(G_n(Y_{t-1}),G_n(Y_t);\alpha). \tag{3.3}$$

The estimator $\tilde{\alpha}$ extends that in Genest et al. (1995) for an i.i.d. random sample $\{(X_i, Y_i)\}_{i=1}^n$ from a bivariate distribution $H(x,y) = C(F(x), G(y); \alpha^*)$ to a univariate time series satisfying Assumption 1. We note that the rescaled empirical distribution $G_n(\cdot)$ is used in criterion (3.3) instead of the standard empirical distribution $n^{-1}\sum_{t=1}^n I\{Y_t \leqslant \cdot\}$; this is a neat device to ensure that the criterion function is well defined for all finite n. As the partial derivatives of $\log c(u_1, u_2; \alpha)$ are infinity at $u_i = 0$ or 1 for $i = 1, 2$ for many copula densities, the use of the rescaled empirical distribution also ensures that the first order condition of criterion (3.3) is well defined for all finite n.

3.2. Estimation of conditional moment and conditional quantile functions

In economic and financial applications, one is often interested in estimating or forecasting certain characteristics of Y_t given Y_{t-1}. These can be easily obtained from the conditional density function $h^*(\cdot|Y_{t-1})$ of Y_t given Y_{t-1}. For example, the conditional kth moment of Y_t given Y_{t-1} can be calculated via

$$\mathrm{E}(Y_t^k | Y_{t-1} = y) = \int z^k h^*(z|y)\,\mathrm{d}z = \int z^k c(G^*(y), G^*(z); \alpha^*)\,\mathrm{d}G^*(z). \tag{3.4}$$

More generally, we may be interested in estimating a vector of conditional moment functions $\mathrm{E}[\psi(Y_t)|Y_{t-1}]$, where ψ is a vector of known measurable functions of Y_t. Since

$$\mathrm{E}[\psi(Y_t)|Y_{t-1} = y] = \int \psi(z) c(G^*(y), G^*(z); \alpha^*)\,\mathrm{d}G^*(z), \tag{3.5}$$

it can be estimated by the following simple plug-in estimator:

$$\widetilde{\mathrm{E}}[\psi(Y_t)|Y_{t-1} = y] = \int \psi(z) c(G_n(y), G_n(z); \widetilde{\alpha})\,\mathrm{d}G_n(z). \tag{3.6}$$

Another important characteristic of the conditional distribution of Y_t given Y_{t-1} is the conditional quantile of Y_t given Y_{t-1} or the conditional VaR of Y_t. Estimating the conditional VaR of portfolios of assets has become routine in risk management, see Gourieroux and Jasiak (2002). For $\{Y_t\}$ satisfying Assumption 1, the conditional quantile function of Y_t given Y_{t-1} can be easily estimated. To see this, note that $Y_t = G^{*-1}(U_t)$ is a monotonic transformation of U_t. Hence the qth conditional quantile of Y_t given Y_{t-1} is given by

$$Q_q^Y(Y_{t-1}) = G^{*-1}(Q_q(G^*(Y_{t-1}); \alpha^*)), \tag{3.7}$$

where $Q_q(u; \alpha^*)$ is the conditional quantile function of U_t given $U_{t-1} = u$:

$$Q_q(u; \alpha^*) = C_{2|1}^{-1}(q|u; \alpha^*), \tag{3.8}$$

in which $C_{2|1}(\cdot|u; \alpha^*) = \frac{\partial}{\partial u_1} C(u, \cdot; \alpha^*) \equiv C_1(u, \cdot; \alpha^*)$ is the conditional distribution of U_t given $U_{t-1} = u$. Bouyé and Salmon (2002) provide explicit expressions of the conditional quantile functions $Q_q(\cdot; \alpha)$ for several specific copulas including the Gaussian copula, the Frank copula, and the Clayton copula.

It follows from (3.7) and (3.8) that the plug-in estimator of the conditional quantile $Q_q(u;\alpha^*)$ of U_t given $U_{t-1}=u$ is

$$\widetilde{Q}_q(u) = Q_q(u;\tilde{\alpha}) = C_{2|1}^{-1}(q|u;\widetilde{\alpha}), \tag{3.9}$$

and the plug-in estimator of the conditional quantile $Q_q^Y(y)$ of Y_t given $Y_{t-1}=y$ is

$$\widetilde{Q}_q^Y(y) = G_n^-(\widetilde{Q}_q(G_n(y))) = G_n^-(C_{2|1}^{-1}(q|G_n(y);\widetilde{\alpha})), \tag{3.10}$$

where $G_n^-(v)=\inf\{y : G_n(y)\geqslant v\}$. For specific copulas, explicit expressions for the conditional quantile estimators are available. For example, for the Clayton copula,

$$\widetilde{Q}_q(u) = [(q^{-\tilde{\alpha}/(1+\tilde{\alpha})}-1)u^{-\tilde{\alpha}}+1]^{-1/\tilde{\alpha}},$$
$$\widetilde{Q}_q^Y(y) = G_n^{-1}([(q^{-\tilde{\alpha}/(1+\tilde{\alpha})}-1)G_n(y)^{-\tilde{\alpha}}+1]^{-1/\tilde{\alpha}}).$$

In general, the conditional quantile function $Q_q(\cdot;\alpha^*)$ is nonlinear. But as it is derived from the conditional distribution of U_t given U_{t-1}, it is automatically monotonic across different quantiles. As a result, the semiparametric conditional quantile function for $\{Y_t\}$ also satisfies the monotonicity property and so does its plug-in estimator $\widetilde{Q}_q^Y(y)$. This is a nice feature of the copula-based approach. Although Koenker and Bassett's (1978) linear quantile regression estimator satisfies this monotonicity property, the nonlinear quantile regression extension typically fails to be monotonic across quantiles.

Remark. Instead of using the rescaled empirical distribution function $G_n(\cdot)$ to estimate $G^*(\cdot)$, we could use the following kernel estimator of the distribution function defined as

$$\widehat{G}_n(y) = \frac{1}{n}\sum_{t=1}^{n} K\left(\frac{y-Y_t}{a_n}\right),$$

where $K(x)=\int_{-\infty}^{x} k(z)\,\mathrm{d}z$ for a kernel density function $k:\mathscr{R}\to[0,\infty)$, and a_n is the bandwidth going to zero at a certain rate as $n\to\infty$. Likewise, we could estimate α^*, $\mathrm{E}[\psi(Y_t)|Y_{t-1}]$ and $Q_q^Y(Y_{t-1})$ using $\widehat{G}_n(\cdot)$ instead of $G_n(\cdot)$. According to the general theory of Newey (1994) on semiparametric two-step estimation, the first order limiting distributions of the estimators based on $\widehat{G}_n(\cdot)$ will be the same as those based on $G_n(\cdot)$ under appropriate conditions.

4. Large sample properties

The main difficulty in establishing the asymptotic properties of the semiparametric estimator $\widetilde{\alpha}$ is that the score function and its derivatives could blow up to infinity near the boundaries. To overcome this difficulty, we first establish convergence of $G_n(\cdot)$ in a weighted metric and then use it to establish the consistency and asymptotic normality of $\tilde{\alpha}$. Finally we present the joint asymptotic distribution of $G_n(\cdot)$ and $\tilde{\alpha}$ which can be used together with the Delta method to establish the asymptotic properties of the conditional moment and conditional quantile estimators.

4.1. Estimators of model parameters

In the following we define $\widetilde{U}_n(v) \equiv G_n(G^{*-1}(v))$ for $v \in (0,1)$. Let $W^*(\cdot)$ be a zero-mean tight Gaussian process in $D[0,1]$ such that $W^*(0) = W^*(1) = 0$, and

$$\mathrm{E}\{W^*(v_1)W^*(v_2)\} = \min\{v_1,v_2\} - v_1v_2 + \sum_{k=2}^{\infty}\{\mathrm{Cov}[I(U_1 \leqslant v_1), I(U_k \leqslant v_2)] + \mathrm{Cov}[I(U_k \leqslant v_1), I(U_1 \leqslant v_2)]\}.$$

Lemma 4.1. *Suppose* $\{Y_t\}$ *satisfies Assumption* 1 *and is* β*-mixing. Let* $w(\cdot)$ *be a continuous function on* $[0,1]$ *which is strictly positive on* $(0,1)$, *symmetric at* $v = 1/2$, *and increasing on* $(0, 1/2]$.

(1) *If* $\beta_t = \mathrm{O}(t^{-b})$ *for some* $b>0$ *and* $\int_0^1 \frac{1}{w(v)} \log(1 + \frac{1}{w(v)})\,\mathrm{d}v < \infty$, *then*

$$\sup_{v\in[0,1]} \left| \frac{\widetilde{U}_n(v) - v}{w(v)} \right| = o_{a.s.}(1), \quad \sup_y \left| \frac{G_n(y) - G^*(y)}{w(G^*(y))} \right| = o_{a.s.}(1).$$

(2) *If either* (i) $\beta_t = \mathrm{O}(t^{-b})$ *for some* $b>\gamma/(\gamma-1)$ *with* $\gamma>1$ *and* $\int_0^1 (\frac{1}{w(v)})^{2\gamma}\,\mathrm{d}v < \infty$; *or* (ii) $\beta_t = \mathrm{O}(b^{-t})$ *for some* $b>1$ *and* $\int_0^1 (\frac{1}{w(v)})^2 \log(1 + \frac{1}{w(v)})\,\mathrm{d}v < \infty$, *then*

$$\sqrt{n}(\widetilde{U}_n(\cdot) - \cdot)/w(\cdot) \to_{dist} W^*(\cdot)/w(\cdot) \quad \text{in } D[0,1],$$

$$\sqrt{n}\sup_y \left| \frac{G_n(y) - G^*(y)}{w(G^*(y))} \right| = \mathrm{O_p}(1).$$

The results in Lemma 4.1 are more general than the standard results: $\sup_y|G_n(y) - G^*(y)| = o_{a.s.}(1)$ and $\sqrt{n}\sup_y|G_n(y) - G^*(y)| = \mathrm{O_p}(1)$. Obviously, choosing $w(v) \equiv 1$ in Lemma 4.1 leads to the latter results. More importantly, weight functions of the form: $w(v) = [v(1-v)]^{1-\xi}$ for all $v \in (0,1)$ and for some $\xi \in (0,1)$, also satisfy the conditions of Lemma 4.1 for appropriate choices of ξ. Such weight functions approach zero when v approaches 0 or 1. Hence, the results in Lemma 4.1 are stronger than the standard results, allowing us to handle unbounded score functions. Previously Shao and Yu (1996, Theorem 2.2) obtained results similar to our Lemma 4.1(2) for stationary strong mixing processes with decay rate $\mathrm{O}(t^{-b})$, $b>1+\sqrt{2}$. Our assumption on the β-mixing decay rate and the method of proof are different from theirs. According to our private communication with Shao and Yu, there is no existing result similar to Lemma 4.1(1).

Define $\mathscr{G}$ as the space of probability distributions over the support of Y_t [say $\mathscr{R}$]. For any $G \in \mathscr{G}$ we let $\|G - G^*\|_{\mathscr{G}} = \sup_y|\{G(y) - G^*(y)\}/w(G^*(y))|$ with $w(\cdot)$ satisfying the condition in Lemma 4.1(1). Let $\mathscr{G}_\delta = \{G \in \mathscr{G} : \|G - G^*\|_{\mathscr{G}} \leqslant \delta\}$ for a small $\delta > 0$.

Let $\mathscr{A} \subset R^d$ be the parameter space. For $\alpha \in \mathscr{A}$, we use $\|\alpha - \alpha^*\|$ to denote the usual Euclidean metric. In addition, let $l(v_1, v_2; \alpha) = \log c(v_1, v_2; \alpha)$. Denote $l_\alpha(v_1, v_2; \alpha) \equiv \frac{\partial l(v_1,v_2;\alpha)}{\partial \alpha}$, $l_{\alpha,\alpha}(v_1, v_2; \alpha) \equiv \frac{\partial^2 l(v_1,v_2;\alpha)}{\partial\alpha\partial\alpha'}$ and $l_{\alpha,j}(v_1, v_2; \alpha) \equiv \frac{\partial^2 l(v_1,v_2;\alpha)}{\partial v_j \partial\alpha}$ for $j = 1, 2$.

Proposition 4.2. *Suppose Assumption* 1 *and the following conditions hold*:

C1. (i) $\alpha^* \in \mathscr{A}$, $\mathscr{A}$ *is a compact subset of* $\mathscr{R}^d$; (ii) $\mathrm{E}[l_\alpha(U_{t-1}, U_t; \alpha)] = 0$ *if and only if* $\alpha = \alpha^*$;

C2. (i) $l_\alpha(v_1, v_2; \alpha)$ *is well-defined for* $(v_1, v_2; \alpha) \in (0,1)^2 \times \mathscr{A}$, *and for all* $\alpha \in \mathscr{A}$, $l_\alpha(U_{t-1}, U_t; \alpha)$ *is Lipschitz continuous at* α *with probability one*; (ii) $l_{\alpha,j}(v_1, v_2; \alpha)$, $(j = 1, 2)$ *are well-defined and continuous in* $(v_1, v_2; \alpha) \in (0,1)^2 \times \mathscr{A}$;

C3. $\{Y_t : t = 1, 2, \ldots\}$ *is* β*-mixing with mixing decay rate* $\beta_t = \mathrm{O}(t^{-b})$ *for some* $b > 0$;

C4. $\mathrm{E}\{\sup_{\overline{\alpha} \in \mathscr{A}} \|l_\alpha(U_{t-1}, U_t; \overline{\alpha})\| \log[1 + \|l_\alpha(U_{t-1}, U_t; \overline{\alpha})\|]\} < \infty$;

C5. *for* $j = 1, 2$, $\mathrm{E}\{\sup_{\overline{\alpha} \in \mathscr{A}, G \in \mathscr{G}_\delta} \|l_{\alpha,j}(G(Y_{t-1}), G(Y_t); \overline{\alpha})\| w(U_{t-2+j})\} < \infty$, *where* $w(\cdot)$ *satisfies the condition in Lemma* 4.1(1).

Then: $\|\widetilde{\alpha} - \alpha^*\| = \mathrm{o_p}(1)$.

We now discuss conditions C1–C5. The first two conditions are standard. The third condition,[11] C3, requires that the process $\{Y_t\}$ be β-mixing with polynomial decay rate, which may be verified via Proposition 2.1. Roughly speaking, C4 is a moment condition on the score function. C5 states that the partial derivatives of the score function with respect to the first two arguments must be dominated by a function which has a finite first moment when weighted by a weighting function $w(\cdot)$ satisfying the condition in Lemma 4.1(1). If the partial derivatives of the score function are bounded, then one can choose the identity weighting function and C5 is automatically satisfied. However, as the partial derivatives of the score function can be unbounded for some copula functions, C5 may not be satisfied with the identity weighting, but may be satisfied with other weighting functions such as $w(v) = [v(1-v)]^{1-\xi}$ for all $v \in (0,1)$ and for some $\xi \in (0,1)$.

Denote $\mathscr{F}_\delta = \{(\alpha, G) \in \mathscr{A} \times \mathscr{G}_\delta : \|\alpha - \alpha^*\| \leqslant \delta\}$ for a small $\delta > 0$. Let $\{G_\eta : \eta \in [0,1]\} \subset \mathscr{G}_\delta$ be a one-dimensional smooth path in $\mathscr{G}_\delta$ with $G_\eta|_{\eta=0} = G^*$. In particular we can take $G_\eta = G^* + \eta[G - G^*]$ for $G \in \mathscr{G}_\delta$. Let $\{(\alpha_\eta, G_\eta) : \eta \in [0,1]\} \subset \mathscr{F}_\delta$ be a one-dimensional smooth path in $\mathscr{F}_\delta$ with $(\alpha_\eta, G_\eta)|_{\eta=0} = (\alpha^*, G^*)$. We also define

$$A_n^* \equiv \frac{1}{n-1} \sum_{t=2}^{n} [l_\alpha(U_{t-1}, U_t; \alpha^*) + W_1(U_{t-1}) + W_2(U_t)], \tag{4.1}$$

$$W_1(U_{t-1}) \equiv \int_0^1 \int_0^1 [I\{U_{t-1} \leqslant v_1\} - v_1] l_{\alpha,1}(v_1, v_2; \alpha^*) c(v_1, v_2; \alpha^*) \, \mathrm{d}v_1 \mathrm{d}v_2, \tag{4.2}$$

$$W_2(U_t) \equiv \int_0^1 \int_0^1 [I\{U_t \leqslant v_2\} - v_2] l_{\alpha,2}(v_1, v_2; \alpha^*) c(v_1, v_2; \alpha^*) \, \mathrm{d}v_1 \mathrm{d}v_2. \tag{4.3}$$

The following set of conditions are sufficient to ensure the $\sqrt{n}$-asymptotic normality of $\widetilde{\alpha}$:

A1. (i) condition C1 is satisfied with $\alpha^* \in int(\mathscr{A})$; (ii) $B \equiv -\mathrm{E}[l_{\alpha,\alpha}(U_{t-1}, U_t; \alpha^*)]$ is positive definite; (iii) $\Sigma \equiv \lim_{n \to \infty} Var(\sqrt{n} A_n^*)$ is positive definite; (iv)

[11]We could replace this condition with a strong mixing condition by using the result in Shao and Yu (1996) mentioned earlier. However the conditions on the strong mixing decay rate and the existence of finite higher order moments of the score function and its partial derivatives will be stronger than those for β-mixing processes. As many copula models have score functions blowing up at a fast rate, it is essential to maintain minimal requirements for the existence of finite higher order moments. This motivates us to use the β-mixing condition instead of the strong mixing.

$\|\widetilde{\alpha} - \alpha^*\| = o_p(1)$, and $\sup_y |\{G_n(y) - G^*(y)\}/w_2(G^*(y))| = O_p(n^{-1/2})$, where $w_2(\cdot)$ satisfies the condition in Lemma 4.1(2);

A2. $l_{\alpha,\alpha}(v_1, v_2; \alpha)$ is well-defined and continuous in $(v_1, v_2; \alpha) \in (0,1)^2 \times int(\mathscr{A})$;

A3. the interchange of differentiation and integration of $l_\alpha(G_\eta(Y_{t-1}), G_\eta(Y_t); \alpha_\eta)$ with respect to $\eta \in (0,1)$ is valid;

A4. (i) $\{Y_t : t = 1, 2, \ldots\}$ is stationary β-mixing with mixing decay rate $\beta_t = O(t^{-b})$ for some $b > \gamma/(\gamma-1)$, in which $\gamma > 1$; (ii) $E\{\|W_1(U_{t-1}) + W_2(U_t)\|^{2\gamma}\} < \infty$;

(iii) $E\{\sup_{(\overline{\alpha},G)\in\mathscr{F}_\delta} \|l_\alpha(G(Y_{t-1}), G(Y_t); \overline{\alpha})\|\}^{2\gamma} < \infty$;

A4′. (i) $\{Y_t : t = 1, 2, \ldots\}$ is stationary β-mixing with mixing decay rate $\beta_t = O(b^{-t})$ for some $b > 1$; (ii) $E\{\|W_1(U_{t-1}) + W_2(U_t)\|^2 \log[1 + \|W_1(U_{t-1}) + W_2(U_t)\|]\} < \infty$;

(iii) $E\{\sup_{(\overline{\alpha},G)\in\mathscr{F}_\delta} \|l_\alpha(G(Y_{t-1}), G(Y_t); \overline{\alpha})\|^2 \log[1 + \|l_\alpha(G(Y_{t-1}), G(Y_t); \overline{\alpha})\|]\} < \infty$;

A5. $E\{\sup_{(\overline{\alpha},G)\in\mathscr{F}_\delta} \|l_{\alpha,\alpha}(G(Y_{t-1}), G(Y_t); \overline{\alpha})\|\}^2 < \infty$;

A6. $E\{\sup_{(\overline{\alpha},G)\in\mathscr{E}_\delta} \|l_{\alpha,j}(G(Y_{t-1}), G(Y_t); \overline{\alpha})\| w(U_{t-2+j})\}^2 < \infty$ for $j = 1, 2$, where $w(\cdot)$ satisfies the condition in Lemma 4.1(1) and $E\{[\frac{w_2(U_t)}{w(U_t)}]^2\} < \infty$.

We now comment on conditions A1 and A6; the other conditions are similar to those in Proposition 4.2. Condition A1(i) requires that α^* be in the interior of the parameter space. This is also assumed in Genest et al. (1995) and is a typical condition in classical parametric and semiparametric models, see the conclusion section for further discussion on this. A1(ii) and A1(iii) are also standard regularity conditions. A1(iv) requires that $G_n(\cdot)$ converge uniformly to $G^*(\cdot)$ at a rate $n^{-1/2}$ in the weighted metric with the weight $w_2(\cdot)$ satisfying the condition in Lemma 4.1(2). This condition implies that $w_2(\cdot)$ could go to zero at a slower rate than that in Lemma 4.1(1). Similar to C5, A6 requires that the partial derivatives of the score function be dominated by a function which has a finite second moment when weighted by the weight function $w(\cdot)$ satisfying the condition in Lemma 4.1(1). The assumption $\int_0^1 [\frac{w_2(v)}{w(v)}]^2 \mathrm{d}v < \infty$ in A6 restricts the relative decay rate of $w(\cdot)$ in A6 to $w_2(\cdot)$ in A1(iv); when the time series $\{Y_t\}$ is stationary β-mixing with the exponential decay rate, we can take $w_2(v) \approx \sqrt{w(v)}$, see e.g., the Gaussian copula example in Section 5. The fact that $w(\cdot)$ could go to zero at a fast rate is very important for copula models in which $\sup_\alpha \|l_{\alpha,j}(v_1, v_2; \alpha)\|$ $(j = 1, 2)$ can blow up to infinity at a fast rate.

Proposition 4.3. *Under Assumption* 1 *and conditions* A1–A3, A4 (*or* A4′), A5–A6, *we have*: (1) $\tilde{\alpha} - \alpha^* = B^{-1}A_n^* + o_p(n^{-1/2})$; (2) $\sqrt{n}(\tilde{\alpha} - \alpha^*) \to N(0, B^{-1}\Sigma B^{-1})$ *in distribution, where* B *and* Σ *are defined in* A1 *and* A_n^* *in* (4.1).

The additional terms $W_1(U_{t-1})$ and $W_2(U_t)$ in A_n^* are introduced by the need to estimate the marginal distribution function $G^*(\cdot)$. In the case where the distribution $G^*(\cdot)$ is completely known, both terms disappear from A_n^*.

4.2. Conditional moment and conditional quantile estimators

Asymptotic properties of the conditional moment and conditional quantile estimators can be established from the joint asymptotic distribution of $G_n(\cdot)$ and $\tilde{\alpha}$

via the Delta method. Lemma 4.1(2), Proposition 4.3(1) and the Cramér–Wold device lead to the following result.

Proposition 4.4. *Under the conditions of Proposition* 4.3,

$$\sqrt{n}\left(\frac{G_n(\cdot)-G^*(\cdot)}{w(G^*(\cdot))},[\tilde{\alpha}-\alpha^*]\right) \to \left(\frac{W^*(G^*(\cdot))}{w(G^*(\cdot))},Z^*\right) \quad \textit{in distribution},$$

where $(\frac{W^*(\cdot)}{w(\cdot)},Z^*)$ *is a bivariate Gaussian process on* $D[0,1]\times\mathscr{R}^d$ *and* $Z^*\sim \mathrm{N}(0,B^{-1}\Sigma B^{-1})$.

The covariance of $(\frac{W^*(\cdot)}{w(\cdot)},Z^*)$ can be derived by using the expression for $G_n(\cdot)$ and Proposition 4.3(1). The expression is tedious and thus omitted. Proposition 4.4 and the following expansions can be used to establish the asymptotic distributions of the conditional moment and conditional quantile estimators. In particular, they show that even though the transition distribution of the time series model is semiparametric, the conditional moment and conditional quantile functions can still be consistently estimated at the parametric $\sqrt{n}$-rate and the estimators are asymptotically normally distributed.

Under mild conditions, one can show that the conditional moment estimator (3.6) satisfies

$$\begin{aligned}
&\widetilde{\mathrm{E}}[\psi(Y_t)|Y_{t-1}=y]-\mathrm{E}[\psi(Y_t)|Y_{t-1}=y]\\
&\quad=\int\psi(z)c(G^*(y),G^*(z);\alpha^*)\mathrm{d}[G_n(z)-G^*(z)]\\
&\qquad+\int\psi(z)c_1(G^*(y),G^*(z);\alpha^*)[G_n(y)-G^*(y)]\,\mathrm{d}G^*(z)\\
&\qquad+\int\psi(z)c_2(G^*(y),G^*(z);\alpha^*)[G_n(z)-G^*(z)]\,\mathrm{d}G^*(z)\\
&\qquad+\int\psi(z)c_\alpha(G^*(y),G^*(z);\alpha^*)\,\mathrm{d}G^*(z)\times(\widetilde{\alpha}-\alpha^*)+\mathrm{o_p}(n^{-1/2}),
\end{aligned}$$

where $c_j(\cdot,\cdot;\alpha^*)$ denotes the partial derivative of c with respect to the j argument, $j=1,2,\alpha$.

Similarly, one can show that under mild conditions, the conditional quantile estimator (3.9) of U_t given $U_{t-1}=u$ satisfies

$$\widetilde{Q}_q(u)-Q_q(u;\alpha^*)=\frac{\partial C_{2|1}^{-1}(q|u;\alpha^*)}{\partial\alpha}(\widetilde{\alpha}-\alpha^*)+\mathrm{o_p}(n^{-1/2}).$$

Again the asymptotic distribution of the estimator of the conditional quantile of U_t given U_{t-1} does not depend on the marginal distribution G^*. Nevertheless, the fact that G^* is unknown and is estimated by G_n does affect the asymptotic variance of $\widetilde{Q}_q(u)$ via its impact on $(\widetilde{\alpha}-\alpha^*)$.

Finally after tedious calculations, we have for the conditional quantile estimator (3.10) of Y_t given $Y_{t-1} = y$:

$$\begin{aligned}\widetilde{Q}_q^Y(y) - Q_q^Y(y) = {} & \frac{1}{g^*(Q_q^Y(y))}\{G_n(Q_q^Y(y)) - G^*(Q_q^Y(y))\} \\ & + \frac{1}{g^*(Q_q^Y(y))}\left\{\frac{\partial C_{2|1}^{-1}(q|u;\alpha^*)}{\partial u_1}[G_n(y) - G^*(y)]\right\} \\ & + \frac{1}{g^*(Q_q^Y(y))}\left\{\frac{\partial C_{2|1}^{-1}(q|u;\alpha^*)}{\partial \alpha}(\widetilde{\alpha} - \alpha^*)\right\} \\ & + o_p(n^{-1/2}), \quad \text{with } u = G^*(y).\end{aligned}$$

Again the conditional quantile of Y_t given Y_{t-1} can be estimated consistently at the parametric $\sqrt{n}$-rate. Unfortunately the limiting distribution of its estimator depends on the marginal density $g^*(Q_q^Y(y))$.

4.3. Statistical inference

The asymptotic distributions of the estimators established in this section may be used to construct inference procedures for the underlying population quantities of interest. The unknown asymptotic variances of the estimators of α^* and of $E[\psi(Y_t)|Y_{t-1} = y]$ can be simply estimated by any existing heteroscedasticity autocorrelation consistent (HAC) covariance estimators, see e.g., Newey and West (1987) and Andrews (1991). The asymptotic variance of the estimator of the conditional quantile $Q_q^Y(y)$ can be obtained by combining a consistent estimator (say a kernel estimator) of the marginal density $g^*(Q_q^Y(y))$ with a HAC estimator, see e.g., Robinson (1983), Powell (1991), Newey (1994) and Engle and Manganelli (2002). Alternatively, some bootstrap methods may be used to approximate the asymptotic distributions of the estimators of interest directly.

For the class of copula-based semiparametric time series models, one convenient bootstrap procedure is the semiparametric bootstrap which takes advantage of the fact that $Y_t = G^{*-1}(U_t)$, where $\{U_t\}_{t=1}^n$ is a stationary first-order Markov process with the copula $C(u_1, u_2; \alpha^*)$ being the joint distribution of (U_1, U_2). The semiparametric bootstrap procedure involves:

Step 1: Generate n independent $U(0,1)$ random variables $\{X_t\}_{t=1}^n$.

Step 2: Generate $U_1^b = X_1$ and $U_t^b = C_{2|1}^{-1}(X_t|U_{t-1}^b; \tilde{\alpha})$ for $t = 2, \ldots, n$. This leads to one bootstrap sample $\{U_t^b\}_{t=1}^n$.

Step 3: Let $Y_t^b = \widehat{G}_n^{-1}(U_t^b)$, where $\widehat{G}_n(y)$ is the kernel estimator defined in Section 3. Compute the corresponding estimate using the bootstrap sample $\{Y_t^b\}_{t=1}^n$.

Step 4: Repeat Steps 1–3 a large number of times and use the empirical distribution of the centered bootstrap values of the estimator to approximate its distribution.

Observing that conditional on the time series $\{Y_t\}_{t=1}^n$, the bootstrap time series $\{Y_t^b\}$ satisfies Assumption 1 with the continuous marginal distribution $\widehat{G}_n(\cdot)$ and the copula function $C(\cdot,\cdot;\tilde{\alpha})$ and hence under the conditions of Proposition 4.3, bootstrap works for all the estimators we proposed in the sense that the conditional distribution of the bootstrap estimator converges in probability to the asymptotic distribution of the corresponding estimator based on the original data. Consequently, inference procedures can be constructed from the bootstrap distribution.

5. Verification of conditions for some copula families

In this section we verify the conditions of Propositions 4.2 and 4.3 for three copulas: the Gaussian copula, the Frank copula, and the Clayton copula. The Gaussian copula is widely used and turns out to be the most difficult to check, as its score function blows up faster than most other copulas. By choosing the weighting functions in A1(iv) and A6 carefully, we are able to verify them for the Gaussian copula. Unlike the Gaussian copula, the Frank copula has bounded score functions. As a result, the identity weighting is enough to verify the conditions of Propositions 4.2 and 4.3 for the Frank copula. The Clayton copula also has unbounded score functions. Similar arguments used to verify conditions for the Gaussian copula can be used to show that the Clayton copula also satisfies the conditions of Propositions 4.2 and 4.3 for appropriate choices of the weighting functions.

5.1. The Gaussian copula

From (2.1), it follows that the copula density of the Gaussian copula is given by

$$c(v_1,v_2;\alpha)=\frac{\phi_\alpha(\Phi^{-1}(v_1),\Phi^{-1}(v_2))}{\phi(\Phi^{-1}(v_1))\phi(\Phi^{-1}(v_2))},$$

where $\phi_\alpha(\cdot,\cdot)$ is the density function of $\Phi_\alpha(\cdot,\cdot)$ and $\phi(\cdot)$ is the density function of $\Phi(\cdot)$. Apart from a constant term, we get

$$\begin{aligned} l(v_1,v_2,\alpha)= & -\frac{1}{2}\ln(1-\alpha^2)-\frac{1}{2(1-\alpha^2)}\{[\Phi^{-1}(v_1)]^2 \\ & +[\Phi^{-1}(v_2)]^2-2\alpha\Phi^{-1}(v_1)\Phi^{-1}(v_2)\}. \end{aligned}$$

As a result, the first and second order partial derivatives of $l(v_1,v_2,\alpha)$ are given by

$$l_\alpha(v_1,v_2,\alpha)=\frac{\alpha(1-\alpha^2)-\alpha\{[\Phi^{-1}(v_1)]^2+[\Phi^{-1}(v_2)]^2\}+(1+\alpha^2)\Phi^{-1}(v_1)\Phi^{-1}(v_2)}{(1-\alpha^2)^2},$$

$$\begin{aligned} & l_{\alpha,\alpha}(v_1,v_2,\alpha) \\ & \quad =\frac{1+\alpha^2}{(1-\alpha^2)^2}+\frac{(6\alpha+2\alpha^3)\Phi^{-1}(v_1)\Phi^{-1}(v_2)-(1+3\alpha^2)\{[\Phi^{-1}(v_1)]^2+[\Phi^{-1}(v_2)]^2\}}{(1-\alpha^2)^3}, \end{aligned}$$

$$l_{\alpha,1}(v_1,v_2,\alpha)=\frac{(1+\alpha^2)\Phi^{-1}(v_2)-2\alpha\Phi^{-1}(v_1)}{(1-\alpha^2)^2\phi(\Phi^{-1}(v_1))},$$
$$l_{\alpha,2}(v_1,v_2,\alpha)=\frac{(1+\alpha^2)\Phi^{-1}(v_1)-2\alpha\Phi^{-1}(v_2)}{(1-\alpha^2)^2\phi(\Phi^{-1}(v_2))}.$$

5.1.1. Consistency

We first establish the consistency of $\widetilde{\alpha}$ for α^* by verifying conditions C1–C5 of Proposition 4.2. Suppose $|\alpha^*|<1$, especially, $\alpha^*\in int(\mathscr{A})$ with $\mathscr{A}=[-1+\eta,1-\eta]$ for an arbitrarily small $\eta>0$. Then condition C1(i) is satisfied. Conditions C1(ii), C2, and C3 are trivially satisfied. It remains to verify conditions C4 and C5. We first notice that there are constants M_1, $M_2>0$ and small $\varepsilon>0$ such that for all $v\in(0,1)$, the following inequalities hold:

$$\left|\frac{\Phi^{-1}(v)}{\phi(\Phi^{-1}(v))}\right|\leqslant[v(1-v)]^{-1},\quad |\Phi^{-1}(v)|\leqslant M_1[v(1-v)]^{-\varepsilon},$$
$$\frac{1}{\phi(\Phi^{-1}(v))}\leqslant M_2[v(1-v)]^{-1},$$

see e.g., Hu (1998, p. 132). Let $r(v)\equiv v(1-v)$, then there are constants $k_1,k_2>0$ such that

$$\sup_{\alpha\in\mathscr{A}}\|l_\alpha(v_1,v_2,\alpha)\|\leqslant k_1\{[r(v_1)r(v_2)]^{-\varepsilon}+[r(v_1)]^{-2\varepsilon}+[r(v_2)]^{-2\varepsilon}\}\leqslant k_2[r(v_1)r(v_2)]^{-2\varepsilon}.$$

Since $U_t\sim U(0,1)$, one can easily verify that condition C4 is satisfied as long as $\varepsilon\in(0,1/2)$ such that $\int_0^1[r(v)]^{-2\varepsilon}\{1+\log([r(v)]^{-2\varepsilon})\}\mathrm{d}v<\infty$. For condition C5, since

$$\sup_{\alpha\in\mathscr{A}}\|l_{\alpha,1}(v_1,v_2,\alpha)\|\leqslant k_1\,\frac{[r(v_2)]^{-\varepsilon}+1}{r(v_1)},\quad \sup_{\alpha\in\mathscr{A}}\|l_{\alpha,2}(v_1,v_2,\alpha)\|\leqslant k_2\,\frac{[r(v_1)]^{-\varepsilon}+1}{r(v_2)},$$

for some constants $k_1,k_2>0$, it suffices to show that for an arbitrarily small $\delta>0$,

$$\mathrm{E}\left[\sup_{G\in\mathscr{G}_\delta}\{[r(G(Y_{t-1}))]^{-1}[r(G(Y_t))]^{-\varepsilon}\}w(U_{t-1})\right]<\infty,$$

for a weighting function $w(\cdot)$ satisfying the condition for Lemma 4.1(1). By the definition of $\mathscr{G}_\delta$, one can show that the following inequalities hold almost surely:

$$\frac{1}{G^*(Y_t)-\delta w(G^*(Y_t))}\geqslant\frac{1}{G(Y_t)}\geqslant\frac{1}{G^*(Y_t)+\delta w(G^*(Y_t))},$$

$$\frac{1}{1-G^*(Y_t)-\delta w(G^*(Y_t))}\geqslant\frac{1}{1-G(Y_t)}\geqslant\frac{1}{1-G^*(Y_t)+\delta w(G^*(Y_t))}.$$

Hence, we get

$$\frac{1}{r(U_{t-1})-\delta w(U_{t-1})}\geqslant\frac{1}{[1-U_{t-1}-\delta w(U_{t-1})][U_{t-1}-\delta w(U_{t-1})]}\geqslant\frac{1}{r(G(Y_{t-1}))},$$

$$\frac{1}{\{r(U_t)-\delta w(U_t)\}^{\varepsilon}} \geqslant \frac{1}{\{[1-U_t-\delta w(U_t)][U_t-\delta w(U_t)]\}^{\varepsilon}} \geqslant \frac{1}{\{r(G(Y_t))\}^{\varepsilon}}.$$

Let $w(v)=[r(v)]^{1-\xi}$ for some $\xi\in(0,1)$. By Holder's inequality, we have

$$\begin{aligned}&\mathrm{E}\left[\frac{w(U_{t-1})}{[r(U_{t-1})-\delta w(U_{t-1})]\{r(U_t)-\delta w(U_t)\}^{\varepsilon}}\right]\\&\quad\leqslant\{\mathrm{E}[\{[r(U_t)]^{\xi}-\delta\}^{-p}]\}^{1/p}\{\mathrm{E}[\{r(U_t)-\delta[r(U_t)]^{1-\xi}\}^{-\varepsilon q}]\}^{1/q},\end{aligned}$$

where $p,q>1$ and $\frac{1}{p}+\frac{1}{q}=1$. Hence condition C5 is satisfied as long as $\xi\in(0,1/p)$ and $\varepsilon\in(0,1/q)$. Proposition 4.2 now implies that $\widetilde{\alpha}-\alpha^*=o_p(1)$.

5.1.2. $\sqrt{n}$-normality

We now establish $\sqrt{n}$-asymptotic normality of $\widetilde{\alpha}$ by verifying conditions A1–A6 of Proposition 4.3. Obviously A1(i) is satisfied. One can easily verify that

$$\begin{aligned}&B=\frac{1+\alpha^{*2}}{(1-\alpha^{*2})^2},\quad W_1(U_{t-1})=\frac{\alpha^*\{[\Phi^{-1}(U_{t-1})]^2-1\}}{2(1-\alpha^{*2})},\\&W_2(U_t)=\frac{\alpha^*\{[\Phi^{-1}(U_t)]^2-1\}}{2(1-\alpha^{*2})},\end{aligned}$$

$$A_n^*=\frac{-1}{n-1}\sum_{t=2}^{n}\frac{\alpha^*(1+\alpha^{*2})([\Phi^{-1}(U_{t-1})]^2+[\Phi^{-1}(U_t)]^2)-2(1+\alpha^{*2})\Phi^{-1}(U_{t-1})\Phi^{-1}(U_t)}{2(1-\alpha^{*2})^2}.$$

Hence conditions A1(ii)(iii) are satisfied. Since the time series generated from Assumption 1 with the Gaussian copula is stationary β-mixing with the exponential decay rate, condition A1(iv) is satisfied with the weighting function $w_2(v)=[r(v)]^{(1-\xi)/2}$ for some $\xi\in(0,1)$. Conditions A2, A3 and A4$'$(i)(ii) are satisfied. It remains to check conditions A4$'$(iii), A5 and A6. Since $\sup_{\alpha\in\mathscr{A}}||l_{\alpha,\alpha}(v_1,v_2,\alpha)||\leqslant k[r(v_1)r(v_2)]^{-2\varepsilon}$, similar to condition C5, one can conclude that conditions A4$'$(iii) and A5 are satisfied if

$$\mathrm{E}\{[\{r(U_{t-1})-\delta w(U_{t-1})\}\{r(U_t)-\delta w(U_t)\}]^{-4\varepsilon}(1+\log[r(U_t)-\delta w(U_t)]^{-2\varepsilon})\}<\infty,$$

which is satisfied for some $\varepsilon\in(0,1/8)$. Finally let $w(v)=[r(v)]^{1-\xi}$ for some $\xi\in(0,1)$ satisfying the condition in Lemma 4.1(1). Then $\mathrm{E}\{[\frac{w_2(U_t)}{w(U_t)}]^2\}=\int_0^1\frac{1}{[r(v)]^{1-\xi}}\,\mathrm{d}v<\infty$. Also for any $p,q>1$ with $\frac{1}{p}+\frac{1}{q}=1$ we have

$$\begin{aligned}&\mathrm{E}\left[\frac{w(U_{t-1})}{\{r(U_{t-1})-\delta w(U_{t-1})\}\{r(U_t)-\delta w(U_t)\}^{\varepsilon}}\right]^2\\&\quad\leqslant\{\mathrm{E}[\{[r(U_t)]^{\xi}-\delta\}^{-2p}]\}^{1/p}\{\mathrm{E}[\{r(U_t)-\delta[r(U_t)]^{1-\xi}\}^{-2\varepsilon q}]\}^{1/q}<\infty,\end{aligned}$$

where the last inequality holds as long as $\xi \in (0, \frac{1}{2p})$ and $\varepsilon \in (0, \frac{1}{2q})$. Hence condition A6 is satisfied. Consequently, the following result holds:

$$\sqrt{n}(\tilde{\alpha} - \alpha^*) = B^{-1} A_n^* + o_p(1) \to N(0, 1 - \alpha^{*2}) \text{ in distribution.}$$

5.2. The Frank copula

The Frank copula density function is

$$\begin{aligned} c(v_1, v_2; \alpha) = {} & \log(\alpha^{-1}) \frac{\alpha^{v_1}\alpha^{v_2}}{1-\alpha}\left[1 - \frac{(1-\alpha^{v_1})(1-\alpha^{v_2})}{1-\alpha}\right]^{-2} \quad \text{if } \alpha > 0, \ \ \alpha \neq 1, \\ = {} & 1 \quad \text{if } \alpha = 1. \end{aligned}$$

This copula generates positive dependence between Y_{t-1} and Y_t when $\alpha \in (0, 1)$, negative dependence when $\alpha > 1$, and independence when $\alpha = 1$, see Nelsen (1999) for additional properties. We assume $\alpha^* \in int(\mathscr{A})$ with $\mathscr{A} = [A^{-1}, A]$ for a large $A > 1$.

If $\alpha > 0, \alpha \neq 1$, then

$$\begin{aligned} l(v_1, v_2, \alpha) = {} & \log\log(\alpha^{-1}) - \log(1-\alpha) + (v_1 + v_2)\log\alpha \\ & - 2\log\left(1 - \frac{(1-\alpha^{v_1})(1-\alpha^{v_2})}{1-\alpha}\right). \end{aligned}$$

Hence,

$$\begin{aligned} l_\alpha(v_1, v_2, \alpha) = {} & \frac{1}{\alpha\log\alpha} + \frac{1}{1-\alpha} + \frac{v_1 + v_2}{\alpha} \\ & - \frac{2[(1-\alpha)\{(1-\alpha^{v_2})\alpha^{v_1} v_1 + (1-\alpha^{v_1})\alpha^{v_2} v_2\} - \alpha(1-\alpha^{v_1})(1-\alpha^{v_2})]}{\alpha(1-\alpha)[(1-\alpha) - (1-\alpha^{v_1})(1-\alpha^{v_2})]}, \end{aligned}$$

$$\begin{aligned} & l_{\alpha,1}(v_1, v_2, \alpha) \\ & = \frac{1}{\alpha} + \frac{2(1-\alpha^{v_2})\alpha^{v_1}\log\alpha[(1-\alpha)\{(1-\alpha^{v_2})\alpha^{v_1} v_1 + (1-\alpha^{v_1})\alpha^{v_2} v_2\} - \alpha(1-\alpha^{v_1})(1-\alpha^{v_2})]}{\alpha(1-\alpha)[(1-\alpha) - (1-\alpha^{v_1})(1-\alpha^{v_2})]} \\ & \quad - \frac{2[(1-\alpha)\{(1-\alpha^{v_2})\alpha^{v_1}(1+\log\alpha) - \alpha^{v_2} v_2 \alpha^{v_1}\log\alpha\} + \alpha^{1+v_1}(1-\alpha^{v_2})\log\alpha]}{\alpha(1-\alpha)[(1-\alpha) - (1-\alpha^{v_1})(1-\alpha^{v_2})]}, \end{aligned}$$

$$\begin{aligned} & l_{\alpha,\alpha}(v_1, v_2, \alpha) \\ & = \frac{[(1-\alpha)\{(1-\alpha^{v_2})\alpha^{v_1} v_1 + (1-\alpha^{v_1})\alpha^{v_2} v_2\} - \alpha(1-\alpha^{v_1})(1-\alpha^{v_2})]^2}{\alpha^2(1-\alpha)^2[(1-\alpha) - (1-\alpha^{v_1})(1-\alpha^{v_2})]^2} \\ & \quad - \frac{2(1-\alpha^{v_1})\{[2\alpha v_2(1-\alpha) + \alpha^2 - (1-\alpha)^2 v_2(1-v_2)]\alpha^{v_2} - \alpha^2\}}{\alpha^2(1-\alpha)^2[(1-\alpha) - (1-\alpha^{v_1})(1-\alpha^{v_2})]} \end{aligned}$$

$$-\frac{2(1-\alpha^{v_2})\{[2\alpha v_1(1-\alpha)+\alpha^2-(1-\alpha)^2 v_1(1-v_1)]\alpha^{v_1}-\alpha^2\}-4(1-\alpha)^2 v_1 v_2 \alpha^{v_1}\alpha^{v_2}}{\alpha^2(1-\alpha)^2[(1-\alpha)-(1-\alpha^{v_1})(1-\alpha^{v_2})]}$$
$$-\frac{1+\log\alpha}{(\alpha\log\alpha)^2}+\frac{1}{(1-\alpha)^2}-\frac{v_1+v_2}{\alpha^2}.$$

If $\alpha=1$, then $l(v_1,v_2,\alpha)=0$; $l_\alpha(v_1,v_2,\alpha)=v_1+v_2-2v_1v_2-1/2$; $l_{\alpha,1}(v_1,v_2,\alpha)=-2v_2+1$; and $l_{\alpha,\alpha}(v_1,v_2,\alpha)=2(v_1v_2)^2-2(v_1v_2)(v_1+v_2-2)-(v_1+v_2)+5/12$.

It is easy to see that Conditions C1, C2, A2 and A3 are automatically satisfied. Although the score function and its derivatives are in complicated forms, one can show that $|l_\alpha(v_1,v_2,\alpha)|$, $|l_{\alpha,\alpha}(v_1,v_2,\alpha)|$, $|l_{\alpha,j}(v_1,v_2,\alpha)|$ for $j=1,2$, are all bounded uniformly in $v_1,v_2\in[0,1]$ and $\alpha\in int(\mathscr{A})$. Hence Conditions C4, C5, A4(iii) or A4′(iii), A5 and A6 are trivially satisfied with the identity weighting function $w(\cdot)=1$. Assuming condition A4(i) or A4′(i), then conditions A1(ii)(iii)(iv) with $w_2(\cdot)=1$, and A4(ii) or A4′(ii) are trivially satisfied. We can now apply Proposition 4.2 to establishing the consistency of $\widetilde{\alpha}$, and apply Proposition 4.3 to obtain its $\sqrt{n}$-asymptotic normality.

5.3. The Clayton copula

The copula density of the Clayton copula is given by

$$c(v_1,v_2;\alpha)=(1+\alpha)v_1^{-(\alpha+1)}v_2^{-(\alpha+1)}[v_1^{-\alpha}+v_2^{-\alpha}-1]^{-(\alpha^{-1}+2)},\quad \text{where } \alpha>0.$$

Hence, the log-copula density and its derivatives are:

$$\begin{aligned} l(v_1,v_2;\alpha)=\ & \log(1+\alpha)-(\alpha+1)\log v_1-(\alpha+1)\log v_2\\ & -(\alpha^{-1}+2)\log(v_1^{-\alpha}+v_2^{-\alpha}-1). \end{aligned}$$

$$\begin{aligned} l_\alpha(v_1,v_2;\alpha)=\ & \frac{1}{1+\alpha}-\log(v_1v_2)+\frac{\log(v_1^{-\alpha}+v_2^{-\alpha}-1)}{\alpha^2}\\ & +\left(\frac{1}{\alpha}+2\right)\frac{v_1^{-\alpha}\log v_1+v_2^{-\alpha}\log v_2}{v_1^{-\alpha}+v_2^{-\alpha}-1},\\ l_{\alpha,1}(v_1,v_2;\alpha)=\ & \frac{-1}{v_1}+\frac{(1+2\alpha)[v_2^{-\alpha}(\log v_2-\log v_1)+\log v_1]+2(v_1^{-\alpha}+v_2^{-\alpha}-1)}{v_1^{\alpha+1}(v_1^{-\alpha}+v_2^{-\alpha}-1)^2},\\ l_{\alpha,\alpha}(v_1,v_2;\alpha)=\ & -\frac{1}{(1+\alpha)^2}-\frac{2}{\alpha^3}\log(v_1^{-\alpha}+v_2^{-\alpha}-1)-\frac{2(v_1^{-\alpha}\log v_1+v_2^{-\alpha}\log v_2)}{\alpha^2(v_1^{-\alpha}+v_2^{-\alpha}-1)}\\ & +\left(\frac{1}{\alpha}+2\right)\left\{\frac{(v_1^{-\alpha}\log v_1+v_2^{-\alpha}\log v_2)^2}{(v_1^{-\alpha}+v_2^{-\alpha}-1)^2}-\frac{v_1^{-\alpha}(\log v_1)^2+v_2^{-\alpha}(\log v_2)^2}{(v_1^{-\alpha}+v_2^{-\alpha}-1)}\right\}. \end{aligned}$$

We note that there are constants $k_1,k_2>0$ and small $\gamma>0$ such that the following inequalities hold for all $v_i\in(0,1)$, $i=1,2$ and all $\alpha>0$:

$$|\log v_i|\leqslant k_1v_i^{-\gamma},\ 0\leqslant\log(v_1^{-\alpha}+v_2^{-\alpha}-1)\leqslant k_2(v_1^{-\gamma}+v_2^{-\gamma}),0\leqslant\frac{v_i^{-\alpha}}{v_1^{-\alpha}+v_2^{-\alpha}-1}\leqslant 1.$$

The remaining verifications of the conditions in Propositions 4.2 and 4.3 for the Clayton copula are very similar to those for the Gaussian copula and are omitted due to space limitations.

6. Conclusions and extensions

In this paper, we have studied the temporal dependence properties and the estimation of a class of semiparametric stationary Markov time series models; a member of this class is completely characterized by a parametric copula and a nonparametric marginal distribution. We have proposed simple estimators of the unknown marginal distribution and the copula dependence parameter, and have established their large sample properties under easily verifiable conditions. In addition, we have demonstrated that features of the transition distribution of models in this class such as the (nonlinear) conditional moment and conditional quantile functions can be easily estimated and their asymptotic properties can be established from those of the estimators $(G_n(\cdot), \tilde{\alpha})$.

As this class of semiparametric Markov models is relatively new, much work remains to be done. We now list a few of them, some of which will be addressed in other papers.

α^* on the boundary: The results established in this paper can be used to construct tests for the correct density forecasts and for the serial independence of a time series that are robust to misspecification of the marginal distribution, see Chen and Fan (2004a). Regarding tests for the serial independence of a time series, one limitation of the asymptotic results obtained in this paper is due to Condition A1(i): the true parameter value α^* is in the interior of the parameter space. If a parametric copula is such that it equals to the independence copula when the parameter takes its value on the boundary of the parameter space, then our Proposition 4.3 is not applicable. In this case, one may establish the limiting distribution result by following Andrews' (2001) approach.

Choice of copula: An important issue faced by an applied researcher interested in using the class of semiparametric copula-based time series models is the choice of an appropriate parametric copula. In different contexts, (1) Chen et al. (2003) propose two simple tests for the correct specification of a parametric copula in the context of modeling the contemporaneous dependence between several univariate time series[12] and of the innovations of univariate GARCH models used to filter each univariate time series; (2) Chen and Fan (2004b) establish pseudo-likelihood ratio tests for selection of parametric copula models for multivariate i.i.d. observations under copula misspecification. Extensions of these tests to time series models considered in this paper will be addressed in a separate paper.

Markov processes of higher order: In principle, the results in this paper can be extended to copula-based semiparametric Markov processes of any finite order. For modeling higher order Markov processes, the parametric copula approach has an

[12] Fermanian (2003) has proposed another copula specification test in this context.

additional appealing feature. That is, the finite dimensional distribution of such processes depends on nonparametric functions of only one dimension and hence achieves dimension reduction. This is particularly useful when the dimension is high due to the curse of dimensionality associated with fully nonparametric modeling. Student's t copula and its extensions should prove useful in high dimensions, see Demarta and McNeil (2004).

Time-varying copulas: Patton (2002a, b; 2004) extended Sklar's theorem to multivariate conditional distributions and applied parametric conditional copulas to model the time-varying dependence between different exchange rates; see Rockinger and Jondeau (2002) and Granger et al. (2003) for similar applications. We can borrow their idea to let the copula function be time-varying in a parametric manner. Alternatively we could let the copula dependence parameter to be time-varying in a Markov-switching manner. Of course both of these extensions will make the resulting time series no longer stationary Markovian. We shall investigate these models in future work.

Acknowledgements

We thank a co-editor, an associate editor and three anonymous referees for detailed suggestions which greatly improved the paper. We are also grateful to R. Engle, S. Goncalves, J. Hausman, J. Hidalgo, R. Koenker, N. Meddahi, A. Patton, Q. Shao, J. Stock, Z. Xiao, seminar participants at LSE, Harvard/MIT, Stanford, Caltech, UC Berkeley, UC Riverside, Stern/NYU, Wisconsin, Rochester and University of Central Florida, conference participants at 2002 NBER/NSF Time Series Conference, 2003 North American Econometric Society Winter Meetings and 2003 Montreal Financial Econometrics Conference for helpful comments. We thank J. Zhang for research assistance on simulation.

Appendix A. Technical proofs

Proof of Proposition 2.1. First, Assumption 1 with a positive copula density function c and conditions in (i) imply that the Markov process $\{U_t\}$ satisfies all the conditions for Theorem 5.2 in Down et al. (1995), hence $\{U_t\}$ is geometric ergodic. This and the definition of beta-mixing imply that $\{U_t\}$ is beta-mixing with the exponential decay rate.

Second, Assumption 1 with a positive copula density function c and conditions in (ii) imply that the Markov process $\{U_t\}$ satisfies all the conditions for Theorem 3.6 in Jarner and Robert (2001), hence $\{U_t\}$ is ergodic with the polynomial decay rate. This and the definition of beta-mixing imply that $\{U_t\}$ is beta-mixing with the polynomial decay rate.

Since $G^*(\cdot)$ is a continuous probability distribution, and by the definition of beta-mixing, $\{Y_t\}$ is beta-mixing with certain decay rate as long as $\{U_t\}$ is beta-mixing with the same decay rate. Hence we obtain the results (i) and (ii). $\square$

Proof of Lemma 4.1. For result (1), we first consider the class of functions $\{\frac{1}{w(v)}I(U_t \leqslant v) : v \in (0, 1/2]\}$. Denote $F(U_t) \equiv \sup_{v\in(0,1/2]}|\frac{1}{w(v)}I(U_t \leqslant v)|$ as the envelop function. Since $\frac{1}{w(v)}$ is decreasing in $v \in (0, 1/2]$, we have $F(U_t) \leqslant \frac{1}{w(U_t)}$. Hence $\mathrm{E}[\{F(U_t)\log[1 + F(U_t)]\}] < \infty$ by the assumption on $w(\cdot)$ and that $\{U_t\}$ is uniformly distributed over $(0, 1)$. Now we can apply Rio's (1995, p. 924) Theorem 1 and Application 5, and obtain $|\{\widetilde{U}_n(v) - v\}/w(v)| = o_{a.s.}(1)$ for any fixed $v \in (0, 1/2]$. Now for any small $\varepsilon > 0$, we form a grid of points $v_0 = 0 < v_1 < \cdots < v_m = 1/2$ such that $\Pr\{\frac{1}{w(v)}I(U_t \leqslant v) : v \in (v_i, v_{i+1})\} < \varepsilon$ for each $i \in \{0, 1, \ldots, m\}$. Then $\sup_{v\in(0,1/2]}|\{\widetilde{U}_n(v) - v\}/ w(v)| \leqslant \max_i|\{\widetilde{U}_n(v_i) - v_i\}/w(v_i)| + \varepsilon$. Hence $\limsup_n\{\sup_{v\in(0,1/2]}|\{\widetilde{U}_n(v) - v\}/w(v)|\} \leqslant \varepsilon$ almost surely. By taking a sequence of small $\varepsilon_m \to 0$, we see that $\limsup_n\{\sup_{v\in(0,1/2]}|\{\widetilde{U}_n(v) - v\}/w(v)|\} = 0$ almost surely. Hence $\{\frac{1}{w(v)}I(U_t \leqslant v) : v \in (0, 1/2]\}$ is a Glivenko–Cantelli class. To show that $\{\frac{1}{w(v)}I(U_t \leqslant v) : v \in (1/2, 1)\}$ is also a Glivenko–Cantelli class, we note that $\frac{1}{w(v)}$ is symmetric about $\frac{1}{2}$, decreasing in $v \in [0, 1/2]$, and $\int_0^1 \frac{1}{w(v)}\mathrm{d}v < \infty$. As a result, it suffices to show that $\{\frac{1}{w(v)}[1 - I(U_t \leqslant v)] : v \in (1/2, 1)\}$ is a Glivenko–Cantelli class, which can be established in the same way as that for $v \in (0, 1/2]$.

For result (2), by the same reasoning as above, it suffices to show that $\{\frac{1}{w(v)}I(U_t \leqslant v) : v \in (0, 1/2]\}$ is a Donsker class. Again by the assumption on $w(\cdot)$, we have that the envelop function $F(U_t) \leqslant \frac{1}{w(U_t)}$. Also by the assumption on $w(\cdot)$ and that $\{U_t\}$ is stationary β-mixing and U_t is a uniform $(0, 1)$ random variable, we have either $\mathrm{E}[F(U_t)]^{2\gamma} < \infty$ with $\gamma > 1$ for β-mixing with the polynomial decay, or $\mathrm{E}\{[F(U_t)]^2 \log[1 + F(U_t)]\} < \infty$ for β-mixing with the exponential decay. Now we can apply Theorem 1 in Doukhan et al. (1995) to conclude that $\{\frac{1}{w(v)}I(U_t \leqslant v) : v \in (0, 1/2]\}$ is a Donsker class. $\square$

In the following let $\mu_n(f) \equiv \frac{1}{n-1}\sum_{t=2}^n [f(Y_{t-1}, Y_t) - Ef(Y_{t-1}, Y_t)]$ be the empirical process indexed by f. Also let $\mathbf{U}_t \equiv (G^*(Y_{t-1}), G^*(Y_t))$ and $\widetilde{\mathbf{U}}_t \equiv (G_n(Y_{t-1}), G_n(Y_t))$.

Proof of Proposition 4.2. Notice that by Assumption 1, condition (C3) and Lemma 4.1, we have $\|G_n - G^*\|_{\mathscr{G}} = \mathrm{o_p}(1)$ for the weight function $w(\cdot)$ stated in condition (C5). Under condition (C1), $\widetilde{\alpha}$ solves $\inf_{\alpha\in\mathscr{A}}\widetilde{Q}(\alpha)$ with $\widetilde{Q}(\alpha) = \{\frac{1}{n}\sum_{t=1}^n l_\alpha(\widetilde{\mathbf{U}}_t, \alpha)\}'\{\frac{1}{n}\sum_{t=1}^n l_\alpha(\widetilde{\mathbf{U}}_t, \alpha)\}$, and α^* solves $\inf_{\alpha\in\mathscr{A}} Q(\alpha)$ with $Q(\alpha) = \{E[l_\alpha(\mathbf{U}_t, \alpha)]\}'\ \{\mathrm{E}[l_\alpha(\mathbf{U}_t, \alpha)]\}$.Again under conditions (C1) and (C2.i), it suffices to show that

$$\sup_{\alpha\in\mathscr{A}}\left\|\frac{1}{n}\sum_{t=1}^n l_\alpha(\widetilde{\mathbf{U}}_t, \alpha) - E[l_\alpha(\mathbf{U}_t, \alpha)]\right\| = \mathrm{o_p}(1).$$

First by conditions (C2), (C3), (C5) and Assumption 1,

$$\begin{aligned}
&\sup_{\alpha\in\mathscr{A}}\left\|\frac{1}{n}\sum_{t=1}^{n}\{l_\alpha(\widetilde{\mathbf{U}}_t,\alpha)-l_\alpha(\mathbf{U}_t,\alpha)\}\right\| \\
&\quad\leqslant \frac{1}{n}\sum_{t=1}^{n}\sup_{\alpha\in\mathscr{A}}\|l_\alpha(\widetilde{\mathbf{U}}_t,\alpha)-l_\alpha(\mathbf{U}_t,\alpha)\| \\
&\quad= \frac{1}{n}\sum_{t=1}^{n}\sup_{\alpha\in\mathscr{A}}\left\|\sum_{j=1}^{2} l_{\alpha,j}(G_\eta(Y_{t-1}),G_\eta(Y_t),\alpha)[G_n(Y_{t-2+j})-G^*(Y_{t-2+j})]\right\| \\
&\quad\leqslant \sum_{j=1}^{2}\left(\frac{1}{n}\sum_{t=1}^{n}\sup_{\overline{\alpha}\in\mathscr{A},G\in\mathscr{G}_\delta}\{|l_{\alpha,j}(G(Y_{t-1}),G(Y_t),\overline{\alpha})|w(G^*(Y_{t-2+j}))\}\right) \\
&\qquad\times\|G_n-G^*\|_{\mathscr{G}}=\mathrm{o_p}(1).
\end{aligned}$$

It remains to show that

$$(\ast)\quad \sup_{\alpha\in\mathscr{A}}\left\|\frac{1}{n}\sum_{t=1}^{n} l_\alpha(\mathbf{U}_t,\alpha)-\mathrm{E}[l_\alpha(\mathbf{U}_t,\alpha)]\right\|=\mathrm{o_p}(1).$$

Under conditions (C1.i) and (C2.i), we know that for any $\varepsilon>0$, there exists $\delta>0$ and m finite integers such that $\{\alpha_1,\ldots,\alpha_m\}$ forms a δ-covering of $\mathscr{A}$, and

$$\sup_{\alpha\in\mathscr{A},\|\alpha-\alpha_i\|\leqslant\delta}\|l_\alpha(\mathbf{U}_t,\alpha)-l_\alpha(\mathbf{U}_t,\alpha_i)\|\leqslant\varepsilon,\qquad \sup_{\alpha\in\mathscr{A},\|\alpha-\alpha_i\|\leqslant\delta}\|\mathrm{E}\{l_\alpha(\mathbf{U}_t,\alpha)-l_\alpha(\mathbf{U}_t,\alpha_i)\}\|\leqslant\varepsilon.$$

Hence

$$\sup_{\alpha\in\mathscr{A},\|\alpha-\alpha_i\|\leqslant\delta}\left\|\frac{1}{n}\sum_{t=1}^{n}\{l_\alpha(\mathbf{U}_t,\alpha)-l_\alpha(\mathbf{U}_t,\alpha_i)\}\right\|\leqslant\varepsilon,$$

$$\sup_{\alpha\in\mathscr{A},\|\alpha-\alpha_i\|\leqslant\delta}\|\mu_n(l_\alpha(\mathbf{U}_t,\alpha))-\mu_n(l_\alpha(\mathbf{U}_t,\alpha_i))\|\leqslant 2\varepsilon.$$

Under conditions (C3) and (C4), we have by Theorem 1 and Application 5 in Rio (1995),

$$\max_{1\leqslant i\leqslant m}\|\mu_n(l_\alpha(\mathbf{U}_t,\alpha_i))\|=\mathrm{o_p}(1).$$

Hence $(\ast)$ is valid. $\square$

Recall that $\|G-G^*\|_{\mathscr{G}}\equiv\sup_y|\{G(y)-G^*(y)\}/w(G^*(y))|$ where $w(\cdot)$ satisfies the condition in Lemma 4.1(1). In the following we also denote $\|G-G^*\|_{\mathscr{G},w_2}\equiv\sup_y|\{G(y)-G^*(y)\}/w_2(G^*(y))|$ where $w_2()$ satisfies the condition in Lemma 4.1(2).

Lemma A.1. Suppose Assumption 1, conditions A1–A3, A4 or A4′, and the following hold:

(a) uniformly over $(\overline{\alpha},G)\in\mathscr{F}_\delta$,

$$\mu_n(l_\alpha(G(Y_{t-1}),G(Y_t),\overline{\alpha})-l_\alpha(\mathbf{U}_t,\alpha^*))=\mathrm{o_p}(n^{-1/2}),$$

(b) uniformly over $(\overline{\alpha}, G) \in \mathscr{F}_\delta$ with $\|G - G^*\|_{\mathscr{G},w_2} = O_p(n^{-1/2})$,

$$\left| \begin{array}{l} E\{l_\alpha(G(Y_{t-1}), G(Y_t), \overline{\alpha})\} - E\{l_{\alpha,\alpha}(\mathbf{U}_t, \alpha^*)[\overline{\alpha} - \alpha^*]\} \\ -\sum_{j=1}^{2} E\{l_{\alpha,j}(\mathbf{U}_t, \alpha^*)[G(Y_{t-2+j}) - G^*(Y_{t-2+j})]\} \end{array} \right|$$
$$= o(\|\overline{\alpha} - \alpha^*\|) + o(\|G - G^*\|_{\mathscr{G},w_2}).$$

Then: $\tilde{\alpha} - \alpha^* = B^{-1}A_n^* + o_p(n^{-1/2})$.

Proof. By condition A1(i) and the first order condition, we have

$$\frac{1}{n-1}\sum_{t=2}^{n} l_\alpha(G_n(Y_{t-1}), G_n(Y_t); \tilde{\alpha}) = 0.$$

In the following we denote $Z = (Y_{t-1}, Y_t)$. By condition (a) we have

$$E_Z[l_\alpha(G_n(Y_{t-1}), G_n(Y_t), \widetilde{\alpha})] + \mu_n(l_\alpha(\mathbf{U}_t, \alpha^*)) = o_p(n^{-1/2}).$$

By condition (b) we have uniformly over $(\overline{\alpha}, G) \in \mathscr{F}_\delta$ with $\|G - G^*\|_{\mathscr{G},w_2} = O_p(n^{-1/2})$,

$$E_Z\{l_{\alpha,\alpha}(\mathbf{U}_t, \alpha^*)[\widetilde{\alpha} - \alpha^*]\} + \sum_{j=1}^{2} E_Z\{l_{\alpha,j}(\mathbf{U}_t, \alpha^*)[G_n(Y_{t-2+j}) - G^*(Y_{t-2+j})]\}$$
$$+ o(\|\widetilde{\alpha} - \alpha^*\|) + o(\|G_n - G^*\|_{\mathscr{G},w_2}) + \mu_n(l_\alpha(\mathbf{U}_t, \alpha^*))$$
$$= o_p(n^{-1/2}).$$

Since $\|G_n - G^*\|_{\mathscr{G},w_2} = O_p(n^{-1/2})$ and $\|\widetilde{\alpha} - \alpha^*\| = o_p(1)$ by condition A1(iv), we have

$$-E_Z\{l_{\alpha,\alpha}(\mathbf{U}_t, \alpha^*)\}[\widetilde{\alpha} - \alpha^*] + o_p(\|\widetilde{\alpha} - \alpha^*\|) = A_n^* + o_p(n^{-1/2}).$$

By conditions A1(i)(iii), A4 or A4′, and the definition of A_n^*, applying Theorem 1 of Doukhan et al. (1995), we have $\sqrt{n}A_n^* \to N(0, \Sigma)$. Now condition A1(ii) implies for any fixed $\lambda \neq 0$, all $\widetilde{\alpha}$ with $\|\widetilde{\alpha} - \alpha^*\| = o_p(1)$,

$$\sqrt{n}\lambda'[\widetilde{\alpha} - \alpha^*] + \sqrt{n} \times o_p(|\lambda'[\widetilde{\alpha} - \alpha^*]|) = \sqrt{n}\lambda' B^{-1}A_n^* + o_p(1),$$

which could hold only if $\sqrt{n}|\lambda'[\widetilde{\alpha} - \alpha^*]|$ is bounded in probability since $\sqrt{n}\lambda' B^{-1}A_n^* \to N(0, B^{-1}\Sigma B^{-1})$. Thus we obtain $\sqrt{n}(\widetilde{\alpha} - \alpha^*) = \sqrt{n}B^{-1}A_n^* + o_p(1)$. □

Lemma A.2. Condition (a) is implied by Assumption 1, conditions A1–A3, A4 or A4′, A5–A6.

Proof. We first show that $\{l_\alpha(G(Y_{t-1}), G(Y_t), \overline{\alpha}) : (\overline{\alpha}, G) \in \mathscr{F}_\delta\}$ is a Donsker class by applying Theorem 1 of Doukhan et al. (1995). Define the envelop function $F(Y_{t-1}, Y_t) = \sup_{(\overline{\alpha},G)\in\mathscr{F}_\delta} |l_\alpha(G(Y_{t-1}), G(Y_t), \overline{\alpha})|$. Then $E_Z\{[F(Y_{t-1}, Y_t)]^{2\gamma}\} < \infty$, $\gamma > 1$ by condition A4(i)(iii) for beta mixing with polynomial decay rate, or

$E_Z\{[F(Y_{t-1}, Y_t)]^2 \log[1+F(Y_{t-1}, Y_t)]\}<\infty$ by condition A4′(i)(iii) for beta mixing with exponential decay rate. By condition A3,

$$\begin{aligned}
&|l_\alpha(G(Y_{t-1}), G(Y_t), \overline{\alpha}) - l_\alpha(\mathbf{U}_t, \alpha^*)| \\
&\quad \leqslant |l_{\alpha,\alpha}(G_\eta(Y_{t-1}), G_\eta(Y_t), \alpha_\eta)| \times \|\overline{\alpha} - \alpha^*\| \\
&\qquad + |l_{\alpha,1}(G_\eta(Y_{t-1}), G_\eta(Y_t), \alpha_\eta) w(G^*(Y_{t-1}))| \times \|G - G^*\|_{\mathscr{G}} \\
&\qquad + |l_{\alpha,2}(G_\eta(Y_{t-1}), G_\eta(Y_t), \alpha_\eta) w(G^*(Y_t))| \times \|G - G^*\|_{\mathscr{G}} \\
&\quad \leqslant \left\{ \sup_{(\alpha_\eta, G_\eta)\in \mathscr{F}_\delta} |l_{\alpha,\alpha}(G_\eta(Y_{t-1}), G_\eta(Y_t), \alpha_\eta)| \right\} \times \|\overline{\alpha} - \alpha^*\| \\
&\qquad + \left\{ \sup_{(\alpha_\eta, G_\eta)\in \mathscr{F}_\delta} |l_{\alpha,1}(G_\eta(Y_{t-1}), G_\eta(Y_t), \alpha_\eta) w(G^*(Y_{t-1}))| \right\} \times \|G - G^*\|_{\mathscr{G}} \\
&\qquad + \left\{ \sup_{(\alpha_\eta, G_\eta)\in \mathscr{F}_\delta} |l_{\alpha,2}(G_\eta(Y_{t-1}), G_\eta(Y_t), \alpha_\eta) w(G^*(Y_t))| \right\} \times \|G - G^*\|_{\mathscr{G}}.
\end{aligned}$$

Hence by conditions A5 and A6,

$$\begin{aligned}
&\log N_{[]}(\varepsilon, \{l_\alpha(G(Y_{t-1}), G(Y_t), \overline{\alpha}) : (\overline{\alpha}, G) \in \mathscr{F}_\delta\}, L_2(P)) \\
&\quad \leqslant K_1 \log N(\varepsilon, \{\overline{\alpha} \in \mathscr{A} : \|\overline{\alpha} - \alpha^*\| \leqslant \delta\}, \|\cdot\|) \\
&\qquad + K_2 \log N(\varepsilon, \mathscr{G}_\delta, \|\cdot\|_{\mathscr{G}}) \leqslant const. \times \left\{ \ln\left(\frac{1}{\varepsilon}\right) + \frac{1}{\varepsilon} \right\},
\end{aligned}$$

this and condition A4(i)(iii) or A4′(i)(iii) imply that all the conditions for Theorem 1 of Doukhan et al. (1995) are satisfied, hence $\{l_\alpha(G(Y_{t-1}), G(Y_t), \overline{\alpha}) : (\overline{\alpha}, G) \in \mathscr{F}_\delta\}$ is a Donsker class, moreover for any $\delta_n \to 0$,

$$\sup_{E_Z[l_\alpha(G,\overline{\alpha}) - l_\alpha(\mathbf{U}_t,\alpha^*)]^2 < \delta_n} \mu_n(l_\alpha(G(Y_{t-1}), G(Y_t), \overline{\alpha}) - l_\alpha(\mathbf{U}_t, \alpha^*)) = o_p(n^{-1/2}).$$

Under conditions A5 and A6, $E_Z\{|l_\alpha(G(Y_{t-1}), G(Y_t), \overline{\alpha}) - l_\alpha(\mathbf{U}_t, \alpha^*)|^2\} \to 0$ as $\|\overline{\alpha} - \alpha^*\| \to 0$ and $\|G - G^*\|_{\mathscr{G}} \to 0$. This implies condition (a). □

Lemma A.3. Condition (b) is implied by conditions A1(i)(iv), A2, A3, A5 and A6.

Proof. By conditions A1(i) and A2, $l_\alpha(G(Y_{t-1}), G(Y_t), \overline{\alpha})$ is continuously Gateaux differentiable in a neighborhood of (α^*, G^*). By Proposition A5.1.E of Bickel et al. (1993, p. 455), condition (b) is implied by: ($\diamond$) for some small $\varepsilon > 0$,

$$\begin{aligned}
&\sup\left\{ \left| \frac{\mathrm{d}E_Z\{l_\alpha(U_{t-1} + \eta\Delta G(Y_{t-1}), U_t + \eta \Delta G(Y_t), \alpha^* + \eta\Delta\alpha)\}}{\mathrm{d}\eta} \right| \right. \\
&\quad \left. : \|\Delta\alpha\| + \|\Delta G\|_{\mathscr{G},w_2} \leqslant 1, |\eta| \leqslant \varepsilon \right\} < \infty.
\end{aligned}$$

By condition A3,

$$\begin{aligned}
&\left|\frac{\mathrm{d}E_Z\{l_\alpha(U_{t-1}+\eta\Delta G(Y_{t-1}), U_t+\eta\Delta G(Y_t), \alpha^*+\eta\Delta\alpha)\}}{\mathrm{d}\eta}\right| \\
&\quad = \left|E_Z\left(\frac{\mathrm{d}l_\alpha(U_{t-1}+\eta\Delta G(Y_{t-1}), U_t+\eta\Delta G(Y_t), \alpha^*+\eta\Delta\alpha)}{\mathrm{d}\eta}\right)\right| \\
&\quad \leqslant E_Z\left(\left|\frac{\mathrm{d}l_\alpha(U_{t-1}+\eta\Delta G(Y_{t-1}), U_t+\eta\Delta G(Y_t), \alpha^*+\eta\Delta\alpha)}{\mathrm{d}\eta}\right|\right) \\
&\quad \leqslant E_Z(|l_{\alpha,\alpha}(U_{t-1}+\eta\Delta G(Y_{t-1}), U_t+\eta\Delta G(Y_t), \alpha^*+\eta\Delta\alpha)|)\times\|\Delta\alpha\| \\
&\qquad + E_Z(|l_{\alpha,1}(U_{t-1}+\eta\Delta G(Y_{t-1}), U_t+\eta\Delta G(Y_t), \\
&\qquad \alpha^*+\eta\Delta\alpha)w_2(G^*(Y_{t-1}))|)\times\|\Delta G\|_{\mathscr{G},w_2} \\
&\qquad + E_Z(|l_{\alpha,2}(U_{t-1}+\eta\Delta G(Y_{t-1}), U_t+\eta\Delta G(Y_t), \\
&\qquad \alpha^*+\eta\Delta\alpha)w_2(G^*(Y_t))|)\times\|\Delta G\|_{\mathscr{G},w_2}.
\end{aligned}$$

By Holder inequality,

$$\begin{aligned}
&E_Z(|l_{\alpha,1}(U_{t-1}+\eta\Delta G(Y_{t-1}), U_t+\eta\Delta G(Y_t), \alpha^*+\eta\Delta\alpha)w_2(G^*(Y_{t-1}))|) \\
&\quad \leqslant \sqrt{E_Z\{|l_{\alpha,1}(U_{t-1}+\eta\Delta G(Y_{t-1}), U_t+\eta\Delta G(Y_t), \alpha^*+\eta\Delta\alpha)w(U_{t-1})|\}^2} \\
&\qquad \times\sqrt{E\left[\frac{w_2(U_{t-1})}{w(U_{t-1})}\right]^2}.
\end{aligned}$$

Hence ($\diamond$) is satisfied given conditions A5 and A6. □

Proof of Proposition 4.3. Result (1) follows directly from Lemmas A.1, A.2 and A.3, Lemma 4.1 and Proposition 4.2. Result (2) follows from result (1) and conditions A1 and A4 (or A4′) and a standard central limit theorem for stationary beta-mixing processes. □

References

Andrews, D., 1991. Heteroskedasticity and autocorrelation consistent covariance matrix estimation. Econometrica 59, 817–858.

Andrews, D., 2001. Testing when a parameter is on the boundary of the maintained hypothesis. Econometrica 69, 683–734.

Bickel, P., Klaassen, C., Ritov, Y., Wellner, J., 1993. Efficient and Adaptive Estimation for Semiparametric Models. The Johns Hopkins University Press.

Bouyé, E., Salmon, M., 2002. Dynamic copula quantile regression and tail area dynamic dependence in forex markets. Manuscript, Financial Econometrics Research Center.

Bouyé, E., Gaussel, N., Salmon, M., 2002. Investigating dynamic dependence using copulae. Manuscript, Financial Econometrics Research Center.

Bradley, R., 1986. Basic properties of strong mixing conditions. In: Eberlein, E., Taqqu, M.S. (Eds.), Dependence in Probability and Statistics. Birkhauser, Boston, pp. 165–192.

Chen, X., Fan, Y., 2004a. Evaluating density forecasts via the copula approach. Finance Research Letters 1, 74–84.
Chen, X., Fan, Y., 2004b. Pseudo-likelihood ratio tests for model selection in semiparametric multivariate copula models. Canadian Journal of Statistics, forthcoming.
Chen, X., Hansen, L.P., Carrasco, M., 1998. Nonlinearity and temporal dependence. Working Paper, University of Chicago.
Chen, X., Fan, Y., Patton, A., 2003. Simple tests for models of dependence between multiple financial time series, with applications to U.S. equity returns and exchange rates. Manuscript, London School of Economics.
Cherubini, U., Luciano, E., 2002. Multivariate option pricing with copulas. Applied Mathematical Finance, forthcoming.
Costinot, A., Roncalli, T., Teiletche, J., 2000. Revisiting the dependence between financial markets with copulas. Working Paper, Crédit Lyonnais.
Darsow, W., Nguyen, B., Olsen, E., 1992. Copulas and Markov processes. Illinois Journal of Mathematics 36, 600–642.
Demarta, S., McNeil, A.J., 2004. The t copula and related copulas. Manuscript.
Devroye, L., 1986. Non-Uniform Random Variate Generation. Springer, New York.
Doukhan, P., Massart, P., Rio, E., 1995. Invariance principles for absolutely regular empirical processes. Ann. Inst. Henri Poincare 31, 393–427.
Down, D., Meyn, S.P., Tweedie, R.L., 1995. Exponential and uniform ergodicity of Markov processes. The Annals of Probability 23, 1671–1691.
Duffie, D., Pan, J., 1997. An overview of value at risk. Journal of Derivatives 4, 7–49.
Embrechts, P., McNeil, A., Straumann, D., 2002. Correlation and dependence properties in risk management: properties and pitfalls. In: Dempster, M. (Ed.), Risk Management: Value at Risk and Beyond. Cambridge University Press, Cambridge, pp. 176–223.
Embrechts, P., Hoing, A., Juri, A., 2003. Using copulae to bound the value-at-risk for functions of dependent risks. Finance and Stochastics 7 (2), 145–167.
Engle, R., Manganelli, S., 2002. CAViaR: conditional autoregressive value at risk by regression quantiles. Journal of Business and Economic Statistics, forthcoming.
Fermanian, J.-D., 2003. Goodness of fit tests for copulas. Manuscript, CREST.
Frees, E.W., Valdez, E.A., 1998. Understanding relationships using copulas. North American Actuarial Journal 2, 1–25.
Frey, R., McNeil, A., 2001. Modeling dependent defaults. Working Paper, Department of Mathematics, ETHZ.
Gagliardini, P., Gourieroux, C., 2002. Duration time series models with proportional hazard. Manuscript, CREST and University of Toronto.
Genest, C., Ghoudi, K., Rivest, L., 1995. A semiparametric estimation procedure of dependence parameters in multivariate families of distributions. Biometrika 82 (3), 543–552.
Gourieroux, C., Jasiak, J., 2002. Value at risk. Manuscript, University of Toronto.
Granger, C.W.J., Terasvirta, T., Patton, A., 2003, Common factors in conditional distributions. Journal of Econometrics, forthcoming.
Heckman, J.J., Honoré, B.E., 1989. The identifiability of the competing risks model. Biometrika 76 (2), 325–330.
Hu, H., 1998. Large sample theory for pseudo-maximum likelihood estimates in semiparametric models. Ph.D. Thesis, University of Washington.
Hu, L., 2002. Dependence patterns across financial markets: methods and evidence. Manuscript, Yale University.
Hull, J., White, A., 1998. Value at risk when daily changes in market variables are not normally distributed. Journal of Derivatives 5, 9–19.
Hutchinson, T., Lai, C., 1990. Continuous Bivariate Distributions, Emphasizing Applications. Rumsby Scientific Publishing, Adelaide.
Jarner, S., Robert, G., 2001. Polynomial convergence rates of Markov chains. Working Paper, Lancaster University.

Joe, H., 1997. Multivariate Models and Dependence Concepts. Chapman & Hall/CRC.
Johnson, M., 1987. Multivariate Statistical Simulation. Wiley, New York.
Koenker, R., Bassett, G., 1978. Regression quantiles. Econometrica 46, 33–50.
Lee, L., 1982. Some approaches to the correction of selectivity bias. Review of Economic Studies 49, 355–372.
Lee, L., 1983. Generalized econometric models with selectivity. Econometrica 51, 507–512.
Li, D., 2000. On default correlation: a copula function approach. Journal of Fixed Income 9, 43–54.
Nelsen, R., 1999. An Introduction to Copulas. Springer.
Newey, W.K., 1994. The asymptotic variance of semiparametric estimators. Econometrica 62, 1349–1382.
Newey, W., West, K., 1987. Hypothesis testing with efficient method of moments estimation. International Economic Review 28, 777–787.
Patton, A.J., 2002a. Modeling time-varying exchange rate dependence using the conditional copula. Working Paper 01-09, Department of Economics, University of California, San Diego.
Patton, A.J., 2002b. Estimation of copula models for time series of possibly different lengths. Working Paper 01-17, Department of Economics, University of California, San Diego.
Patton, A.J., 2004. On the importance of skewness and asymmetric dependence in stock returns for asset allocation. Journal of Financial Econometrics, forthcoming.
Powell, J., 1991. Estimation of monotonic regression models under quantile restriction. In: Barnett, W., Powell, J., Tauchen, G. (Eds.), Non-parametric and Semi-parametric Methods in Econometrics and Statistics. Cambridge University Press, New York.
Rio, E., 1995. A maximal inequality and dependent Marcinkiewicz–Zygmund strong laws. Annals of Probability 23, 918–937.
Robinson, P., 1983. Nonparametric estimators for time series. Journal of Time Series Analysis 4, 185–207.
Rockinger, M., Jondeau, E., 2002. Conditional dependency of financial series: an application of copulas. Manuscript, HEC School of Management, France.
Rosenberg, J., 1999. Semiparametric pricing of multivariate contingent claims. Working Paper, Department of Finance, Stern School of Business, New York University.
Shao, Q., Yu, H., 1996. Weak convergence for weighted empirical processes of dependent sequences. The Annals of Probability 24, 2098–2127.
Shih, J., Louis, T., 1995. Inferences on the association parameter in copula models for bivariate survival data. Biometrics 51, 1384–1399.
Sklar, A., 1959. Fonctions de r'epartition 'a *n* dimensionset leurs marges. Publ. Inst. Statis. Univ. Paris 8, 229–231.