Estimation of Copula-Based Semiparametric Time Series Models*

Xiaohong Chen† — New York University

Yanqin Fan‡ — Vanderbilt University

First version: October 2002; This version: February 2004

Abstract

This paper studies the estimation of a class of copula-based semiparametric stationary Markov models. These models are characterized by nonparametric invariant (or marginal) distributions and parametric copula functions that capture the temporal dependence of the processes; the implied transition distributions are all semiparametric. Models in this class are easy to simulate, and can be expressed as semiparametric regression transformation models. One advantage of this copula approach is to separate out the temporal dependence (such as tail dependence) from the marginal behavior (such as fat tailedness) of a time series. We present conditions under which processes generated by models in this class are β-mixing; naturally, these conditions depend only on the copula specification. Simple estimators of the marginal distribution and the copula parameter are provided, and their asymptotic properties are established under easily verifiable conditions. Estimators of important features of the transition distribution such as the (nonlinear) conditional moments and conditional quantiles are easily obtained from estimators of the marginal distribution and the copula parameter; their $\sqrt{n}-$ consistency and asymptotic normality can be obtained using the Delta method. In addition, the semiparametric conditional quantile estimators are automatically monotonic across quantiles.

JEL Classification: C14; C22

KEY WORDS: Copula; Nonlinear Markov models; β-Mixing; Weighted empirical process; Semiparametric estimation; Conditional moment; Conditional quantile

---

*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. Hidalgo, R. Koenker, N. Meddahi, A. Patton, Q. Shao, Z. Xiao, seminar participants at London School of Economics, University of Central Florida and Rochester, 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. Both authors acknowledge financial support from the National Science Foundation, and Chen acknowledges support from the C.V. Starr Center at NYU.

†Xiaohong Chen, Department of Economics, New York University, 269 Mercer Street, New York, NY 10003, USA; Tel: (212) 998-8970; Fax: (212) 995-4186; E-mail: xiaohong.chen@nyu.edu

‡Yanqin Fan, Department of Economics, Vanderbilt University, VU Station B 351819, 2301 Vanderbilt Place, Nashville, TN 37235, USA; Tel: (615) 322-3796; Fax: (615) 343-8495; E-mail: yanqin.fan@vanderbilt.edu

1 Introduction

Copulas have gained popularity in finance and insurance community[1] in the past few years because of the flexibility they offer in modeling the distribution of multivariate random variables; see e.g., Frees and Valdez (1998) and Embrechts et al. (2002) for reviews. A copula connects a multivariate distribution to its marginals in such a way that it captures the entire dependence structure in the multivariate distribution. The importance of copulas in modeling the distribution of a multivariate random variable is justified by the Sklar's (1959) theorem: any multivariate distribution can be expressed as its copula function evaluated at its marginal distribution functions; and any copula function when evaluated at any marginal distributions is a multivariate distribution. Hence the information in the joint distribution is decomposed into those in the marginal distributions and that in the copula function. Consequently copulas allow one to model the marginal distributions and the dependence structure of a multivariate random variable separately. Moreover, the copula measure of dependence is invariant to any increasing transformation of individual series.

Papers that apply copulas in the finance and insurance literatures include Rosenberg (1999) and Cherubini and Luciano (2002) which analyze multivariate option pricing; Hull and White (1998) and Embrechts, et al. (2003) which study the portfolio Value-at-Risk; Li (2000) and Frey and McNeil (2001) which develop models of default and credit risk; and Costinot, et al. (2000) and Hu (2002) which investigate contagion, to mention just a few. Patton (2002a, b, 2004) extend Sklar's theorem to conditional distributions and apply conditional copulas to modeling the time-varying dependence between different exchange rates, among other things; see Rockinger and Jondeau (2002) and Granger et al. (2003) for similar applications.

While the afore-mentioned papers use copulas to model the contemporaneous dependence between multiple time series, there are a few published papers proposing to use copulas to model temporal dependence within a time series. 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; Darsow, et al. (1992) provide a necessary and sufficient condition for a copula-based time series to be a Markov process. In the copula approach to 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 the dependence structure and the marginal behavior of a time series separately, allowing for a wide variety of marginal behaviors (such as skewness and fat tailedness) and dependence properties (such as asymmetric dependence and positive tail dependence).

[1]Copulas have also proven to be useful in microeconometrics, see e.g. Lee (1982a,b, 1983) and Ray, et al. (1980) on bivariate logit and sample selection models, Heckman and Honore (1989) on competing risk models.

This separate modelling of the temporal dependence and the marginal behavior is particularly important when it is known that the dependence structure and the marginal properties of a time series are affected by different exogenous variables, which can be easily modeled via the parametric copula approach by letting the copula parameter depend on X_t (say) and the marginal distribution depend on Z_t (say, which may differ from X_t).

In this paper, we study a class of univariate copula-based semiparametric stationary Markov models, in which copulas are parameterized and are used to model the dependence between the adjacent observations in a univariate time series, but the invariant (or marginal) distributions are left unspecified. Our specification is more general than Joe's (1997) in that we do not parameterize the marginal distribution, and hence our estimation and inference is robust to misspecification of marginals. Nevertheless, both ours and Joe's (1997) specifications impose strict stationarity, while the most general copula-based Markov models proposed in Darsow, et al. (1992) can allow for marginal distributions to vary over time. However, Darsow, et al. (1992) only studied some probabilistic properties of their copula-based Markov models. Given that we only observe a finite sample of the time series once, it is impossible to estimate marginal distributions fully nonparametrically if we also allow for arbitrarily time-varying marginal distributions.

Although we restrict our attention to a class of strictly stationary Markov models, we shall demonstrate that many flexible semiparametric regression transformation models belong to this class of copula-based semiparametric stationary Markov models. Unlike the standard approach of specifying either the finite dimensional joint distribution or the transition distribution of a stationary Markov process parametrically, our class of models implies a semiparametric specification of the transition distribution. Moreover, the abundance of parametric copula specifications will generate many new forms of transition distributions, and hence many more nonlinear Markov models which are easy to simulate.[2] In Section 2 we also provide conditions under which processes generated by models in this class are $\beta-mixing$.[3] Given that the main advantage of a copula-based approach is to separate out the temporal dependence from the marginal behavior of a time series, it is natural that our sufficient conditions for processes in this class to be $\beta-mixing$ with polynomial decay rates depend only on the copula specification.

A member of the class of copula-based semiparametric stationary Markov models is completely characterized by two unknown parameters: the copula dependence parameter α^* (i.e., the finite-dimensional parameter in the copula function specification); and the invariant (or marginal) distribution function $G^*(\cdot)$. The unknown marginal distribution can be estimated by any one of the

[2] One important application of copulas in probability and statistics literature is in simulating new multivariate models.

[3] $\beta-mixing$ is one popular measure of temporal persistency for nonlinear Markov time series models.

existing nonparametric methods, including the rescaled empirical distribution function and the kernel smoothed estimator of the distribution function. The copula dependence parameter can then be estimated by the pseudo maximum likelihood method. Since the sample pseudo likelihood criterion depends on the first-step estimator of the marginal distribution function, the resulting estimator of the dependence parameter α^* is semiparametric and is often called a two-step estimator. In particular, we focus on the two-step estimator of α^* where the rescaled empirical distribution function is used as the first step estimator of $G^*(\cdot)$ in the pseudo likelihood criterion. This method extends the two-step approach commonly used in bivariate copula models for i.i.d. observations[4] to our class of univariate copula-based semiparametric time series models. We establish the consistency and $\sqrt{n}-$asymptotic normality of the semiparametric estimators of (G^*, α^*) under easily-verifiable conditions. Interestingly, the asymptotic variance of the two-step estimator of the copula dependence parameter α^* does not depend on the functional form of the marginal distribution G^*, and hence any marginal density behavior (such as fat tailedness) has no impact on the large sample inference using the two-step estimator of α^*. As in the i.i.d. case, these results are not easy to establish under primitive conditions, as the score functions and their derivatives blow up to infinity for many widely used copula functions including the Gaussian copula, the Students t-copula, and the Clayton copula. The conditions presented in this paper are weak enough to allow for such copula functions.[5] We overcome the technical difficulty by making use of the asymptotic properties of the rescaled empirical distribution function in a weighted metric. This technique should also be useful in establishing asymptotic properties of estimators in other models in which the score function blows up to infinity.

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 semiparametric estimators of the

[4]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 to be censored. Both papers and Hu (1998) present the asymptotic normality of their semiparametric estimators for i.i.d. observations.

[5]Although the conditions and propositions are stated for copula-based univariate time series models in this paper, they are also applicable to bivariate time series models where the parametric copula functions are used to model the contemporaneous dependence between the two univariate stationary time series.

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 semiparametric estimators of (G^*, α^*) and by applying the Delta method, 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, the estimators of the nonlinear conditional moment and conditional quantile are still $\sqrt{n}$-consistent and asymptotically normal. Moreover, the estimated conditional quantile functions are automatically monotonic across different quantiles. These are nice features of the copula-based semiparametric time series models.

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. They briefly mention the two-step procedure of Genest et al. (1995)[6] for estimating the copula dependence parameter without establishing its large sample properties. Moreover, they didn't study the estimation of any nonlinear 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 provide a brief introduction to copulas, present the class of copula-based semiparametric time series models considered in this paper, and study their $\beta-$mixing property. We also discuss the close relation between the copula-based semiparametric time series models and the 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 semiparametric estimator for three widely used copulas. Section 6 concludes with discussions of several extensions. All the proofs are relegated to the Appendix.

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

2 Copula-Based Time Series Models

2.1 A Brief Review of Copulas

A copula is a multivariate distribution whose marginal distributions are uniform distributions on the interval $(0, 1)$. The importance of copulas in modeling the distribution of multivariate random variables is justified by the Sklar's theorem. For simplicity, let's consider the bivariate case. Let $H(x, y)$ denote the joint distribution function of random variables X and Y whose marginal distribution functions are continuous, denoted as F and G respectively. Sklar's theorem states that there exists a unique copula function $C(v_1, v_2) = H(F^{-1}(v_1), G^{-1}(v_2))$ that connects $H(x, y)$ to $F(x)$ and $G(y)$ via $H(x, y) = C(F(x), G(y))$. Hence the information in the joint distribution $H(x, y)$ is decomposed into those in the marginal distributions and that in the copula function, where the latter captures the dependence structure between X and Y. On the other hand, for any copula function C and any univariate distribution functions F and G, the function $C(F(x), G(y))$ is a bivariate distribution function. Consequently copulas allow one to model the marginal distributions and the dependence structure of a multivariate random variable separately. For more discussions on the theory of copulas and specific examples of copulas, see Joe (1997) and Nelsen (1999).

One copula that we will refer to frequently in this paper is the Gaussian copula. Let $\Phi_\alpha(\cdot, \cdot)$ be the distribution function of the bivariate normal distribution with means zero, variances 1, and correlation coefficient α. Then the Gaussian copula is given by

$$C(v_1, v_2; \alpha) = \Phi_\alpha(\Phi^{-1}(v_1), \Phi^{-1}(v_2)), \tag{2.1}$$

where $0 \leq v_1, v_2 \leq 1$ and $\Phi(\cdot)$ is the distribution function of a standard normal random variable. By Sklar's theorem, for any two marginal distribution functions $F(\cdot)$ and $G(\cdot)$, the distribution defined as

$$H(x, y) = C(F(x), G(y); \alpha) = \Phi_\alpha(\Phi^{-1}(F(x)), \Phi^{-1}(G(y))) \tag{2.2}$$

is a bivariate distribution function whose marginals are $F(\cdot)$ and $G(\cdot)$ respectively, and the copula that connects $H(\cdot, \cdot)$ to $F(\cdot)$ and $G(\cdot)$ is the Gaussian copula. Hence Sklar's theorem allows one to construct bivariate distributions with non-Normal marginal distributions and the Gaussian copula.

Different copulas typically exhibit different dependence properties. Joe (1997) and Nelsen (1999) contain excellent discussions of various dependence measures and of dependence properties of numerous parametric copulas. One useful dependence measure in modeling financial time series is that of tail dependence; this is a measure of dependence between random variables in the extreme lower and upper joint tails. For example, the coefficients of lower and upper tail dependence of a bivariate copula C of (X, Y) are defined as:

$$\tau^L = \lim_{q\to 0} \Pr[G(Y) \le q | F(X) \le q] = \lim_{q\to 0} \frac{C(q,q)}{q} \tag{2.3}$$

$$\tau^U = \lim_{q\to 1} \Pr[G(Y) > q | F(X) > q] = \lim_{q\to 1} \frac{1-2q+C(q,q)}{1-q}. \tag{2.4}$$

Heuristically, if (X, Y) denotes returns on two assets, then the coefficients of upper (lower) tail dependence of the copula C measure the probability of an extremely large positive (negative) return on one asset (Y) given that the other asset has yielded an extremely large positive (negative) return (X). If the two assets have a bivariate Gaussian copula, then both upper and lower tail dependence coefficients are zero, i.e., the bivariate Gaussian copula generates zero tail dependence.[7]

2.2 Copula-Based Semiparametric Time Series Models and Their Mixing Properties

Let $\{Y_t\}$ be a stationary Markov process of order one. Then its statistical properties are completely determined by the joint distribution 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 of Y_t and the copula function of Y_{t-1} and Y_t. As a result, the statistical properties of a stationary Markov process $\{Y_t\}$ are completely determined by its marginal distribution and the copula of Y_{t-1} and Y_t. This suggests the copula approach as an alternative approach to modeling a stationary Markov process: instead of specifying the joint distribution of Y_{t-1} and Y_t, one specifies the marginal distribution 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-tailedness of the time series $\{Y_t\}$, while the latter characterizes the 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 G^* is left unspecified, but the copula function has a parametric form. 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. To avoid these trivial cases, we shall rule out the perfect dependent copulas in this paper.

Assumption 1: $\{Y_t : t = 1, ..., n\}$ *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*

[7]However, this does not mean that the bivariate Gaussian copula (with correlation coefficient α) goes to the independence copula (i.e., $C(u, v) = uv$) unless $\alpha = 0$. More generally, *tail independence* (i.e., $\tau^L = 0, \tau^U = 0$) is not equivalent to *independence in the tail* (i.e., $\lim_{x,y}\{H(x, y)/[F(x)G(y)]\} = 1$).

with respect to Lebesgue measure on 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 does not satisfy the Fréchet-Hoeffding upper or lower bound.*

If the marginal distribution $G^*(\cdot)$ belongs to a parametric class of distributions, then Assumption 1 specifies a class of stationary, parametric Markov processes, which was studied in Joe (1997). Otherwise, it specifies a class of stationary, semiparametric Markov processes which is robust to misspecification of marginals.

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

$$h^*(y_t|y_{t-1}) = g^*(y_t)c(G^*(y_{t-1}), G^*(y_t); \alpha^*), \tag{2.5}$$

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 our copula approach over the standard approach is to separate out the temporal dependence structure from the marginal behavior. This is particularly important when it is known that the dependence structure and the marginal properties of the time series are affected by different exogenous variables, which can be easily modeled via the copula approach by letting the copula parameter α^* depend on X_t (say) and the marginal distribution G^* depend on Z_t (which may differ from X_t). A related advantage is that the copula measure of temporal dependence is invariant to any increasing transformation of the time series.

We observe that the transformed process, $\{U_t : U_t \equiv G^*(Y_t)\}$, is a stationary parametric Markov process. Under Assumption 1, 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 $f_{U_t|U_{t-1}=u_0}(u) = c(u_0, u; \alpha^*)$.

We now study the temporal persistency properties of a time series satisfying Assumption 1.

Definition 1 *(Davydov (1973)) For a stationary Markov process* $\{X_t\}$*, its* β*-mixing coefficients are given by:* $\beta_t = E\{\ \sup_{0\le\psi\le1} |\ E[\psi(X_t)|X_0] - E[\psi(X_t)]\ |\ \}$*. The process* $\{X_t\}$ *is* β*-mixing if* $\lim_{t\to\infty}\beta_t = 0$*.*

The following result shows that the β-mixing property of a copula-based Markov process $\{Y_t\}$ is completely determined by its copula density function $c(\cdot,\cdot;\alpha^*)$. In the following a real-valued function Λ is called *norm-like* if the closure of the set $\{x : \Lambda(x) \le B\}$ is compact for each $B > 0$.

Proposition 2.1 *Under Assumption 1, if* $c(u_1, u_2; \alpha^*)$ *is aperiodic, then the following (i) and (ii) hold:*

(i) If there are constants $0 < \lambda < 1$, $0 < d < \infty$, a norm-like function $\Lambda(\cdot) \geq 1$, and a small set $\mathbf{K}$ such that

$$\int_0^1 \Lambda(u) \times c(U_{t-1}, u; \alpha^*) du \leq \lambda \Lambda(U_{t-1}) + d 1_{\mathbf{K}}(U_{t-1}),$$

then $\{Y_t\}$ is β-mixing with the exponential decay rate: $\beta_t \leq const \times \exp\{-at\}$ for some $a > 0$;

(ii) If there are constants $\lambda \in [0, 1)$, $0 < a, d < \infty$, a norm-like function $\Lambda(\cdot) \geq 1$, and a small set $\mathbf{K}$ such that

$$\int_0^1 \Lambda(u) \times c(U_{t-1}, u; \alpha^*) du \leq \Lambda(U_{t-1}) - a[\Lambda(U_{t-1})]^\lambda + d 1_{\mathbf{K}}(U_{t-1}),$$

then $\{Y_t\}$ is β-mixing with the polynomial decay rate: $\beta_t (1+t)^{\lambda/(1-\lambda)} \to 0$ as $t \to \infty$.

The assumption that $c(u_1, u_2; \alpha^*)$ is aperiodic 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 the time series with such a copula is β-mixing with at least a polynomial decay rate.

For many first-order nonlinear stationary Markov models, the conditions that ensure β-mixing with certain decay rates will involve the invariant distributions, see e.g. Chen, et al. (1998) for diffusion models. It is interesting to note that the conditions for β-mixing in Proposition 2.1 do not depend on the invariant distribution G^*, but only depend on the copula specification.

2.3 Semiparametric Regression Transformation Models

As discrete-time Markov models in econometrics are typically expressed as regression models, we now provide such representations for the copula-based stationary Markov time series models.

Example 1: Let the copula $C(\cdot, \cdot; \alpha)$ be the Gaussian copula defined in (2.1). Then the process $\{\Phi^{-1}(G^*(Y_t))\}$ is a Gaussian process that can be represented by

$$\Phi^{-1}(G^*(Y_t)) = \alpha \Phi^{-1}(G^*(Y_{t-1})) + \varepsilon_t, \tag{2.6}$$

where $\varepsilon_t \sim N(0, 1-\alpha^2)$, and is independent of Y_{t-1}. If the marginal distribution $G^*(\cdot)$ is left unspecified, then we have the class of semiparametric time series models generated by the Gaussian copula. If 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, we obtain first order Markov processes characterized by the Gaussian copula, but non-normal marginal distributions. By applying Proposition 2.1(i) to this example with $\Lambda(u) = [1 + \Phi^{-1}(u)]I\{1/2 \leq u \leq 1\} + [1 - \Phi^{-1}(u)]I\{0 \leq u < 1/2\}$, 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.

Other examples satisfying Assumption 1 can be constructed from the following class of *regression transformation* models:

$$\Lambda_1(Y_t) = \Lambda_2(Y_{t-1}) + \sigma(Y_{t-1})e_t, \tag{2.7}$$

where $\Lambda_1(\cdot)$ is an increasing function, $\inf_y \sigma(y) > 0$, and $\{e_t\}$ is an i.i.d. sequence with mean zero and variance one, and e_t is independent of Y_{t-1}.

Example 2: Clearly (2.7) includes the following *semiparametric regression transformation* model:

$$\Lambda_{1,\theta_1}(G^*(Y_t)) = \Lambda_{2,\theta_2}(G^*(Y_{t-1})) + \sigma_{\theta_3}(G^*(Y_{t-1}))e_t, \tag{2.8}$$

where $G^*(\cdot)$ is the unknown probability distribution function of Y_t, $\Lambda_{1,\theta_1}(\cdot)$ is a parametric increasing function, $\Lambda_{2,\theta_2}(\cdot)$ and $\sigma_{\theta_3}(\cdot) > 0$ are also parametric functions, e_t is independent of Y_{t-1}, and $\{e_t\}$ is i.i.d. with a parametric probability density $f_e(\cdot;\theta_4)$ satisfying mean zero and variance 1. It is easy to see that $\{Y_t\}$ generated from (2.8) satisfies our Assumption 1 with the parametric copula density function given by:

$$c(u_0, u_1; \alpha^*) = f_e\left(\frac{\Lambda_{1,\theta_1}(u_1) - \Lambda_{2,\theta_2}(u_0)}{\sigma_{\theta_3}(u_0)}; \theta_4\right) \times \frac{d\Lambda_{1,\theta_1}(u_1)}{du_1},$$

where α^* consists of the distinct elements of $\theta_1, \theta_2, \theta_3, \theta_4$. For instance, the stationary Markov process with the Gaussian copula in Example 1 with a nonparametric marginal distribution $G^*(\cdot)$ is an example of model (2.8) in which $\Lambda_{1,\theta_1}(u_1) = \Phi^{-1}(u_1)$, $\Lambda_{2,\theta_2}(u_0) = \alpha\Phi^{-1}(u_0)$, $\sigma_{\theta_3}(u_0) = \sqrt{1-\alpha^2}$, $f_e(\cdot;\theta_4)$ is the standard normal density, and $\alpha^* = \alpha = \theta_2 = \theta_3$.

On the other hand, Assumption 1 is consistent with the following *generalized semiparametric regression transformation* model without the independence restriction between the error term and Y_{t-1}:

$$\Lambda_{1,\theta_1}(G^*(Y_t)) = \Lambda_{2,\theta_2}(G^*(Y_{t-1})) + \varepsilon_t, \quad E\{\varepsilon_t|Y_{t-1}\} = 0, \tag{2.9}$$

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

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

We note that the functional form of $\Lambda_{2,\theta_2}(\cdot)$ is completely pinned down by $\Lambda_{1,\theta_1}(\cdot)$ and the copula density $c(\cdot,\cdot;\alpha^*)$. To see this, we recall that $U_t \equiv G^*(Y_t)$. Hence

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

A special case of (2.9) is given by $\Lambda_{1,\theta_1}(u_1) = u_1$, the identity mapping. In this case,

$$G^*(Y_t) = m(G^*(Y_{t-1}); \alpha^*) + \varepsilon_t, \quad E\{\varepsilon_t|Y_{t-1}\} = 0,$$

where the conditional density of ε_t given Y_{t-1} is $f_{\varepsilon_t|Y_{t-1}}(\varepsilon) = c(G^*(Y_{t-1}), \varepsilon + m(G^*(Y_{t-1}); \alpha^*); \alpha^*)$ and

$$m(U_{t-1}; \alpha^*) = E(U_t|U_{t-1}) = \int_0^1 uc(U_{t-1}, u; \alpha^*)du = 1 - \int_0^1 \frac{\partial C(U_{t-1}, u; \alpha^*)}{\partial U_{t-1}}du.$$

Hutchinson and Lai (1990) point out that some commonly used copulas have simple expressions for $m(U_{t-1}; \alpha^*)$. For example, the Farlie-Gumbel-Morgenstern (F-G-M) copula

$$C(u_1, u_2; \alpha^*) = u_1u_2[1 + \alpha^*(1 - u_1)(1 - u_2)], \quad |\alpha^*| \leq 1,$$

and the Plackett copula[8]

$$C(u_1, u_2; \alpha^*) = \frac{[1 + (\alpha^* - 1)(u_1 + u_2)] - \{[1 + (\alpha^* - 1)(u_1 + u_2)]^2 - 4\alpha^*(\alpha^* - 1)u_1u_2\}^{1/2}}{2(\alpha^* - 1)},$$

have $m(U_{t-1}; \alpha^*)$ being linear in U_{t-1}. Noting that $E[U_t|U_{t-1}] = \frac{3-\alpha}{6} + \frac{\alpha}{3}U_{t-1}$ for the F-G-M copula, one can apply Proposition 2.1 (i) with $\Lambda(u) = 1 + u$ to conclude that $\{Y_t\}$ generated by the F-G-M copula is always β-mixing with the exponential decay rate.

In closing this subsection, we point out that the copula-based time series specifications lead naturally to semiparametric quantile regression models. For example, the following quantile regression model holds for $\{Y_t\}$ satisfying Assumption 1:

$$G^*(Y_t) = Q_q(G^*(Y_{t-1}); \alpha^*) + \eta_t, \quad \Pr(\eta_t \leq 0|Y_{t-1}) = q \in (0, 1),$$

where the q -th conditional quantile function $Q_q(U_{t-1}; \alpha^*)$ of U_t given U_{t-1} can be solved from:

$$\int_0^{Q_q(U_{t-1};\alpha^*)} c(U_{t-1}, u; \alpha^*)du = q \quad \text{for all } U_{t-1},$$

or alternatively from

$$Q_q(U_{t-1}; \alpha^*) = C_{2|1}^{-1}(q|U_{t-1}; \alpha^*), \tag{2.10}$$

[8]Lee (1982b) has applied the Plackett copula to construct bivariate logit models. Ray, et al. (1980) have studied the sample selection models using the F-G-M copula and/or the Pareto copula with logistic marginals.

where $C_{2|1}(\cdot|U_{t-1};\alpha^*) = \frac{\partial}{\partial u_1}C(U_{t-1},\cdot;\alpha^*) \equiv C_1(U_{t-1},\cdot;\alpha^*)$ is the conditional distribution of U_t given U_{t-1}. 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.

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 above semiparametric quantile regression model for the time series $\{Y_t\}$ also satisfies the monotonicity property.

2.4 Simulating Copula-Based Time Series Models

Figure 1 presents time series plots and the corresponding scatter plots of realizations of three time series models with the Gaussian copula with $\alpha = 0.5$ and the marginal distributions given by the standard Normal distribution and the Student's t distribution with degrees of freedom 3 and 10 respectively. It is apparent from both the time series plots and the scatter plots that the dependence structure of time series characterized by the Gaussian copula is symmetric regardless of its marginal distribution. Coupled with fat-tailed marginal distributions such as the Student's t with 3 degrees of freedom, the time series model with the Gaussian copula produces large and small values. However no clustering of such large or small values occurs as the Gaussian copula does not have tail dependence. As the degrees of freedom of the Student's t distribution increases, the time series plot resembles more and more like the one with the Normal marginal distribution.

Figure 1 reveals the limitation of time series models with the Gaussian copula in the context of modeling economic and financial time series exhibiting complicated nonlinear asymmetric dependence and clusters of large and/or small values. Fortunately, a wide variety of non-Gaussian copulas is available to serve this purpose, see Joe (1997) and Nelsen (1999). For example, the Clayton copula is given by:

$$C(u_1, u_2; \alpha) = [u_1^{-\alpha} + u_2^{-\alpha} - 1]^{-1/\alpha}, \quad \text{where } \alpha > 0. \tag{2.11}$$

The lower tail dependence parameter for this family is $\tau^L = 2^{-1/\alpha}$ and the upper tail dependence parameter is $\tau^U = 0$. 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 family of copulas 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.

It is very easy to simulate a time series from a copula-based Markov model. Let $C_{2|1}(\cdot|u_1;\alpha^*) = \frac{\partial}{\partial u_1}C(u_1,\cdot;\alpha^*) \equiv C_1(u_1,\cdot;\alpha^*)$ be the conditional distribution of U_t given $U_{t-1} = u_1$. To generate a series $\{Y_t\}_{t=1}^n$ from a non-Gaussian copula-based time series model $(G^*(\cdot), C(\cdot,\cdot;\alpha^*))$, one may proceed as follows:

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

Step 2. Generate $U_1 = X_1$ and $U_t = C_{2|1}^{-1}(X_t|U_{t-1};\alpha^*)$ for $t = 2, ..., n$.

Step 3. Generate $Y_t = G^{*-1}(U_t)$ for $t = 1, \ldots, n$.

One can easily verify that the necessary and sufficient condition for $\{Y_t\}_{t=1}^n$ to be a realization from a Markov process is satisfied, see Darsow, et al. (1992).

For the Clayton copula, $C_{2|1}(u_2|u_1;\alpha) = u_1^{-(\alpha+1)}[u_1^{-\alpha} + u_2^{-\alpha} - 1]^{-(\alpha^{-1}+1)}$ and $C_{2|1}^{-1}(u_2|u_1;\alpha) = [(u_2^{-\alpha/(1+\alpha)} - 1)u_1^{-\alpha} + 1]^{-1/\alpha}$. Figures 2a and 2b present time series plots and the corresponding scatter plots of realizations of time series models with the Clayton copula with $\alpha = (0.5, 2, 10)$ and the marginal distributions given by the standard normal distribution (Figure 2a) and the Student's t distribution with degrees of freedom 3 (Figure 2b) respectively. These figures demonstrate that: (1) unlike the Gaussian copula, 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 as the Student's t distribution with 3 degrees of freedom, the Clayton copula with large α produces clusters of small values.

Alternative algorithms are available for generating random variables from specific copulas, see Devroye (1986), Johnson (1987), and Nelsen (1999). These can be adapted to generate time series observations from copula-based time series models.

3 Estimation of Copula-Based Semiparametric Time Series Models

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

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

It remains to estimate the copula parameter α^*. 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 on the right hand side of (3.2) 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^*:

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

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 $H(x, y) = C(F(x), G(y); \alpha^*)$ of (X, Y) to a univariate time series satisfying Assumption 1. We note that the rescaled empirical distribution $G_n(\cdot)$ is used in the criterion (3.3) instead of the standard empirical distribution $\frac{1}{n}\sum_{t=1}^{n} I\{Y_t \leq \cdot\}$; 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 popular copula densities, the use of the rescaled empirical distribution also ensures that the first order condition of the 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 may be 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 k-th mean of Y_t given Y_{t-1} can be calculated via

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

Equation (3.4) reveals that in general the conditional mean and the conditional variance of Y_t given Y_{t-1} are nonlinear functions of Y_{t-1}.

More generally, we may be interested in estimating a vector of conditional moment functions $E[\psi(Y_t)|Y_{t-1}]$, where ψ is a vector of known measurable functions of Y_t. For example, $\psi(Y_t) = (|Y_t|, |Y_t|^2)'$. Since

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

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

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

Another important measure 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 conditional VaR of portfolios of assets has become routine in risk management, see Gourieroux and Jasiak (2002).

Noting that $Y_t = G^{*-1}(U_t)$ is a monotonic transformation of U_t, the q-th 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 $Q_q(G^*(Y_{t-1}); \alpha^*)$ is defined in (2.10).

The plug-in estimator of the conditional quantile $Q_q(u; \alpha^*)$ of U_t given $U_{t-1} = u$ is defined as:

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

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

$$\widetilde{Q}_q^Y(y) = G_n^-(\widetilde{Q}_q(G_n(y))) = G_n^- \left(C_{2|1}^{-1}(q|G_n(y); \tilde{\alpha})\right), \tag{3.9}$$

where $G_n^-(v) = \inf\{y : G_n(y) \geq v\}$ is the generalized quantile function. For specific copulas, explicit expressions for the conditional quantile estimators are available. For example, for the Clayton copula, subsection 2.4 implies that

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

We note that this semiparametric conditional quantile estimator $\widetilde{Q}_q^Y(y)$ is always non-decreasing in q. 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:

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

where $K(x) = \int_{-\infty}^{x} k(z)dz$ for a kernel density function $k : \mathcal{R} \to [0, \infty)$, and a_n is the bandwidth going to zero at a certain rate as $n \to \infty$. Likewise, we could estimate α^*, $E[\psi(Y_t)|Y_{t-1}]$ and $Q_q^Y(Y_{t-1})$ using $\widehat{G}_n(\cdot)$ instead of $G_n(\cdot)$:

$$\widehat{\alpha} = \arg\max_{\alpha \in \mathcal{A}} \frac{1}{n}\sum_{t=2}^{n} \log c(\widehat{G}_n(Y_{t-1}), \widehat{G}_n(Y_t); \alpha),$$

$$\widehat{E}[\psi(Y_t)|Y_{t-1}] = \int \psi(z)c(\widehat{G}_n(Y_{t-1}), \widehat{G}_n(z); \widehat{\alpha})d\widehat{G}_n(z), \quad \widehat{Q}^Y_q(Y_{t-1}) = \widehat{G}_n^{-}\left(C^{-1}_{2|1}(q|\widehat{G}_n(Y_{t-1}); \widehat{\alpha})\right).$$

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 $G_n(\cdot)$.

4 Large Sample Properties of the Proposed Estimators

In principle, we could state the large sample properties of the proposed estimators by simply referring to the existing general theories on semiparametric two-step estimation such as Andrews (1994), Newey (1994), Newey and McFadden (1994), and Chen, et al. (2003). However, we would like to establish the asymptotic properties under primitive sufficient conditions. 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 Asymptotic Properties of $G_n(\cdot)$

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

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

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 $\beta_t = O(t^{-b})$ for some $b > 0$ and $\int_0^1 \frac{1}{w(v)} \log(1 + \frac{1}{w(v)})dv < \infty$, 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) $\beta_t = O(t^{-b})$ for some $b > \gamma/(\gamma-1)$ with $\gamma > 1$ and $\int_0^1 (\frac{1}{w(v)})^{2\gamma} dv < \infty$; or (ii) $\beta_t = O(b^{-t})$ for some $b > 1$ and $\int_0^1 (\frac{1}{w(v)})^2 \log(1 + \frac{1}{w(v)})dv < \infty$, then

$$\begin{aligned} \sqrt{n}\left(\widetilde{U}_n(\cdot) - \cdot\right)/w(\cdot) &\to_{dist} W^*(\cdot)/w(\cdot) \quad \text{in } D[0,1], \\ \sqrt{n}\sup_{y}\left|\frac{G_n(y) - G^*(y)}{w(G^*(y))}\right| &= O_p(1). \end{aligned}$$

The results in Lemma 4.1 are much more general than the standard results: $\sup_y |G_n(y) - G^*(y)| = o_{a.s.}(1)$ and $\sqrt{n}\sup_y |G_n(y) - G^*(y)| = O_p(1)$. 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 choice 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 $O(t^{-b})$, $b > 1 + \sqrt{2}$. 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).

4.2 Asymptotic Properties of $\tilde{\alpha}$

In the following, we shall define $\mathcal{G}$ as the space of continuous probability distributions over the support of Y_t [say $\mathcal{R}$]. For any $G \in \mathcal{G}$ we let $||G - G^*||_{\mathcal{G}} = \sup_y |\{G(y) - G^*(y)\}/w(G^*(y))|$ with $w(\cdot)$ satisfying the condition in Lemma 4.1(1). Let $\mathcal{G}_\delta = \{G \in \mathcal{G} : ||G - G^*||_{\mathcal{G}} \leq \delta\}$ for a small $\delta > 0$. Obviously $G^* \in \mathcal{G}_\delta$, and G_n will belong to $\mathcal{G}_\delta$ with probability approaching one. Let $\{G_\eta : \eta \in [0,1]\} \subset \mathcal{G}_\delta$ be a one-dimensional smooth path in $\mathcal{G}_\delta$ with $G_\eta|_{\eta=0} = G^*$. In particular we can take $G_\eta = G^* + \eta[G - G^*]$ for $G \in \mathcal{G}_\delta$.

Let $\mathcal{A} \subset R^d$ be the parameter space. For $\alpha \in \mathcal{A}$, 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) $\alpha^* \in \mathcal{A}$, $\mathcal{A}$ is a compact subset of $\mathcal{R}^d$; (ii) $E[l_\alpha(U_{t-1}, U_t; \alpha))] = 0$ if and only if $\alpha = \alpha^*$;*
C2. *(i) $l_\alpha(v_1, v_2; \alpha)$ is well-defined for $(\alpha, v_1, v_2) \in \mathcal{A} \times (0,1) \times (0,1)$, and for all $\alpha \in \mathcal{A}$, $l_\alpha(U_{t-1}, U_t; \alpha)$ is Lipschitz continuous at α with probability one; (ii) $l_{\alpha,j}(v_1, v_2; \alpha)$, $j = 1, 2$ are well-defined, and are continuous in $(\alpha, v_1, v_2) \in \mathcal{A} \times (0,1) \times (0,1)$;*
C3. *$\{Y_t : t = 1, 2, ...\}$ is β-mixing with the mixing decay rate $\beta_t = O(t^{-b})$ for some $b > 0$;*
C4. *$E\{\sup_{\overline{\alpha}\in\mathcal{A}} ||l_\alpha(U_{t-1}, U_t; \overline{\alpha})|| \log[1 + ||l_\alpha(U_{t-1}, U_t; \overline{\alpha})||]\} < \infty$;*
C5. *for $j = 1, 2$, $E\{\sup_{\overline{\alpha}\in\mathcal{A}, G\in\mathcal{G}_\delta} ||l_{\alpha,j}(G(Y_{t-1}), G(Y_t); \overline{\alpha})||w(U_{t-2+j})\} < \infty$, where $w(\cdot)$ satisfies the condition in Lemma 4.1(1).*
Then: $||\widetilde{\alpha} - \alpha^*|| = o_p(1)$.

We now discuss conditions C1-C5. The first two conditions are standard. The third condition,[9]

[9] We could replace this condition with a strong mixing condition by using the result in Shao and Yu (1996)

C3, requires that the process $\{Y_t\}$ is β-mixing with the 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 $\xi \in (0,1)$.

In the following we denote $\mathcal{F}_\delta = \{(\alpha, G) \in \mathcal{A} \times \mathcal{G}_\delta : ||\alpha - \alpha^*|| \leq \delta\}$ for a small $\delta > 0$. Let $\{(\alpha_\eta, G_\eta) : \eta \in [0,1]\} \subset \mathcal{F}_\delta$ be a one-dimensional smooth path in $\mathcal{F}_\delta$ with $(\alpha_\eta, G_\eta)|_{\eta=0} = (\alpha^*, G^*)$. We also define

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

$$W_1(U_{t-1}) \equiv \int_0^1\int_0^1 [I\{U_{t-1} \leq v_1\} - v_1] l_{\alpha,1}(v_1, v_2; \alpha^*) c(v_1, v_2; \alpha^*) dv_1 dv_2, \tag{4.2}$$

$$W_2(U_t) \equiv \int_0^1\int_0^1 [I\{U_t \leq v_2\} - v_2] l_{\alpha,2}(v_1, v_2; \alpha^*) c(v_1, v_2; \alpha^*) dv_1 dv_2. \tag{4.3}$$

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

A1. (i) condition C1 is satisfied with $\alpha^* \in int(\mathcal{A})$; (ii) $B \equiv -E\left[l_{\alpha,\alpha}(U_{t-1}, U_t; \alpha^*)\right]$ is positive definite; (iii) $\Sigma \equiv \lim_{n\to\infty} Var(\sqrt{n}A_n^*)$ is positive definite; (iv) $||\widetilde{\alpha} - \alpha^*|| = o_p(1)$, 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. $l_{\alpha,\alpha}(v_1, v_2; \alpha)$, $l_{\alpha,j}(v_1, v_2; \alpha)$, $j = 1, 2$, are well-defined, and are continuous in $(\alpha, v_1, v_2) \in int(\mathcal{A}) \times (0,1) \times (0,1)$;

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) $\{Y_t : t = 1, 2, ...\}$ is stationary β-mixing with the mixing decay rate $\beta_t = O(t^{-b})$ for some $b > \gamma/(\gamma - 1)$, in which $\gamma > 1$; (ii) $E\{||W_1(U_{t-1}) + W_2(U_t)||^{2\gamma}\} < \infty$ for some $\gamma > 1$;
(iii) $E\{\sup_{(\overline{\alpha},G)\in\mathcal{F}_\delta} ||l_\alpha(G(Y_{t-1}), G(Y_t); \overline{\alpha})||\}^{2\gamma} < \infty$ for some $\gamma > 1$;

A4'. (i) $\{Y_t : t = 1, 2, ...\}$ is stationary β-mixing with the mixing decay rate $\beta_t = O(b^{-t})$ for some $b > 1$; (ii) $E\{||W_1(U_{t-1}) + W_2(U_t)||^2 \log[1 + ||W_1(U_{t-1}) + W_2(U_t)||]\} < \infty$;

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.

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

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

A6. $E\{\sup_{(\overline{\alpha},G)\in\mathcal{F}_\delta} ||l_{\alpha,j}(G(Y_{t-1}),G(Y_t);\overline{\alpha})||w(U_{t-2+j})\}^2<\infty$ 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^* is 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 about this. A1(ii) and A1(iii) are also standard regularity conditions. A1(iv) requires that $G_n(\cdot)$ converges 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 are 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 $\int_0^1[\frac{w_2(v)}{w(v)}]^2dv<\infty$ 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^*. It is interesting to note that the asymptotic variance of $\tilde{\alpha}$ does not depend on the functional form of the marginal distribution G^*.

4.3 Asymptotic Properties of the Conditional Moment and Conditional Quantile Estimators

The 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* $D[0,1]\times\mathcal{R}^d$ *and* $Z^*\sim N(0, B^{-1}\Sigma B^{-1})$.

The covariance of $(\frac{W^*(\cdot)}{w(\cdot)}, Z^*)$ can be derived by using the expression of $G_n(\cdot)$ 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

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

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 (3.8) of U_t given $U_{t-1}=u$ satisfies

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

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 (3.9) of Y_t given $Y_{t-1}=y$:

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

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 $g^*(Q_q^Y(y))$.

4.4 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 $t = 2, ..., n$. 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 Examples

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

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

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

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

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

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

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, $\alpha^* \in int(\mathcal{A})$ with $\mathcal{A} = [-1+\eta, 1-\eta]$ 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 $\epsilon > 0$ such that for all $v \in (0,1)$, the following inequalities hold:

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

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

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

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

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

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

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

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

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

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

Hence, we get

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

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

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

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

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 $\epsilon \in (0, 1/q)$. 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

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

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

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 $\sup_{\alpha\in\mathcal{A}} ||l_{\alpha\alpha}(v_1,v_2,\alpha)|| \leq k[r(v_1)r(v_2)]^{-2\epsilon}$, similar to condition C5, one can conclude that conditions A4'(iii) and A5 are satisfied if

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

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

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

where the last inequality holds as long as $\xi \in (0, \frac{1}{2p})$ and $\epsilon \in (0, \frac{1}{2q})$. 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 $\alpha^* \in int(\mathcal{A})$ with $\mathcal{A} = [A^{-1}, A]$ for a large $A > 1$.

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

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

Hence,

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

$$
\begin{aligned}
l_{\alpha,1}(v_1,v_2,\alpha) &= \frac{1}{\alpha}+\frac{2\left[\frac{(1-\alpha^{v_2})\alpha^{v_1}v_1+(1-\alpha^{v_1})\alpha^{v_2}v_2}{\alpha(1-\alpha)}-\frac{(1-\alpha^{v_1})(1-\alpha^{v_2})}{(1-\alpha)^2}\right]\left[\frac{(1-\alpha^{v_2})\alpha^{v_1}\log\alpha}{1-\alpha}\right]}{[1-\frac{(1-\alpha^{v_1})(1-\alpha^{v_2})}{1-\alpha}]^2} \\
&\quad -\frac{2\left[\frac{(1-\alpha^{v_2})\alpha^{v_1}(1+\log\alpha)-\alpha^{v_2}v_2\alpha^{v_1}\log\alpha}{\alpha(1-\alpha)}+\frac{(1-\alpha^{v_2})\alpha^{v_1}\log\alpha}{(1-\alpha)^2}\right]}{1-\frac{(1-\alpha^{v_1})(1-\alpha^{v_2})}{1-\alpha}}, \\
l_{\alpha\alpha}(v_1,v_2,\alpha) &= \frac{2\left[\frac{(1-\alpha^{v_2})\alpha^{v_1}v_1+(1-\alpha^{v_1})\alpha^{v_2}v_2}{\alpha(1-\alpha)}-\frac{(1-\alpha^{v_1})(1-\alpha^{v_2})}{(1-\alpha)^2}\right]^2}{[1-\frac{(1-\alpha^{v_1})(1-\alpha^{v_2})}{1-\alpha}]^2} \\
&\quad -\frac{2\frac{(1-\alpha^{v_1})}{(1-\alpha)^3}\left\{\left(\frac{2v_2}{\alpha}-2v_2+1-\frac{(1-\alpha)^2v_2(1-v_2)}{\alpha^2}\right)\alpha^{v_2}-1\right\}}{1-\frac{(1-\alpha^{v_1})(1-\alpha^{v_2})}{1-\alpha}} \\
&\quad -\frac{2\frac{(1-\alpha^{v_2})}{(1-\alpha)^3}\left\{\left(\frac{2v_1}{\alpha}-2v_1+1-\frac{(1-\alpha)^2v_1(1-v_1)}{\alpha^2}\right)\alpha^{v_1}-1\right\}-\frac{4v_1v_2\alpha^{v_1}\alpha^{v_2}}{\alpha^2(1-\alpha)}}{1-\frac{(1-\alpha^{v_1})(1-\alpha^{v_2})}{1-\alpha}} \\
&\quad -\frac{1+\log\alpha}{(\alpha\log\alpha)^2}+\frac{1}{(1-\alpha)^2}-\frac{v_1+v_2}{\alpha^2}.
\end{aligned}
$$

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 $\alpha\in int(\mathcal{A})$. 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:

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

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

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

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 $\alpha > 0$:

$$
|\log v_i| \leq k_1 v_i^{-\gamma},\ 0 \leq \log(v_1^{-\alpha} + v_2^{-\alpha} - 1) \leq k_2(v_1^{-\gamma} + v_2^{-\gamma}),\ 0 \leq \frac{v_i^{-\alpha}}{v_1^{-\alpha} + v_2^{-\alpha} - 1} \leq 1.
$$

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 temporal dependence properties of a class of stationary semiparametric 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 easily 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 (2003). 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.

Efficient estimation: For a bivariate copula model with i.i.d. observations, Genest and Werker (2001), Klaassen and Wellner (2001) have shown that the two-step estimator is generally inefficient unless the copula is the Gaussian copula or the independence copula. Intuitively, the inefficiency results from the two step nature of the estimator and the use of the inefficient empirical distribution function in the first step. Recently, Chen and Fan (2002) and Gagliardini and Gourieroux (2002b) have independently considered the semiparametric efficient estimation of the copula parameter. For i.i.d. multivariate observations, Chen and Fan (2002) show that the sieve MLE joint estimation of the copula parameter and the unknown marginals are semiparametrically efficient. We expect that their result remains valid for copula-based semiparametric Markov time series models considered in this paper.

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 of multivariate time series[10] and of the innovations of univariate GARCH models used to filter each univariate time series; (2) Chen and Fan (2004) 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.

Semiparametric copula: We note that the parametric specification of the copula function does rule out some choices of $\Lambda_1(\cdot)$ and $\Lambda_2(\cdot)$ in the semiparametric regression transformation models described in Section 2. For example, Gagliardini and Gourieroux (2002a) have proposed a class of stationary Markov duration time series models with proportional hazard. Their class of models belongs to the regression transformation model (2.7) with $\Lambda_1(Y_t) = \log(\Lambda_0(Y_t))$ where $\Lambda_0(Y_t)$ is a baseline cumulated hazard, $\Lambda_2(Y_{t-1}) = \log(\frac{1}{a(Y_{t-1})})$, $\sigma(Y_{t-1}) = 1$ and ε_t has a Gompertz distribution (i.e., log of standard exponential). Their paper allows for $\Lambda_0(Y_t), a(Y_{t-1})$ to be fully nonparametric, which leads to a semi-nonparametric specification of the copula density function via the following relation: $c(G^*(y_0), G^*(y_1)) = f_\varepsilon(\Lambda_1(y_1) - \Lambda_2(y_0)) \times \Lambda_1'(y_1)/g^*(y_1)$. See also Gagliardini and Gourieroux (2002b) and Gagliardini and Gourieroux (2003). We are currently extending our analysis to allow for semiparametric specification of the copula function such as the Archimedian copulas.

Markov processes of higher order: 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 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.

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

Appendix: Technical Proofs

Proof. (**Proposition 2.1**) First, Assumption 1 with aperiodic 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 aperiodic 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 Roberts (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 $G^*()$ 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). ■

Proof. (**Lemma 4.1**) For result **(1)**, we first consider the class of functions $\{\frac{1}{w(v)}I(U_t \leq v) : v \in (0, 1/2]\}$. Denote $F(U_t) \equiv \sup_{v\in(0,1/2]} \left|\frac{1}{w(v)}I(U_t \leq v)\right|$ as the envelop function. Since $\frac{1}{w(v)}$ is decreasing in $v \in (0, 1/2]$, we have $F(U_t) \leq \frac{1}{w(U_t)}$. Hence $E[\{F(U_t)\log[1 + F(U_t)]\}] < \infty$ by the assumption on $w(\cdot)$ and that $\{U_t\}$ is uniformly distributed over $(0, 1)$. Now we can apply Rio's (1995, page 924) theorem 1 and application 5, and obtain $\left|\{\widetilde{U}_n(v) - v\}/w(v)\right| = o_{a.s.}(1)$ for any fixed $v \in (0, 1/2]$. Now for any small $\varepsilon > 0$, we form a grid of points $v_0 = 0 < v_1 < ... < v_m = 1/2$ such that $\Pr\{\frac{1}{w(v)}I(U_t \leq v) : v \in (v_i, v_{i+1})\} < \varepsilon$ for each $i \in \{0, 1, ..., m\}$. Then $\sup_{v\in(0,1/2]} \left|\{\widetilde{U}_n(v) - v\}/w(v)\right| \leq \max_i \left|\{\widetilde{U}_n(v_i) - v_i\}/w(v_i)\right| + \varepsilon$. Hence $\limsup_n\{\sup_{v\in(0,1/2]} \left|\{\widetilde{U}_n(v) - v\}/w(v)\right|\} \leq \varepsilon$ almost surely. By taking a sequence of small $\varepsilon_m \to 0$, we see that $\limsup_n\{\sup_{v\in(0,1/2]} \left|\{\widetilde{U}_n(v) - v\}/w(v)\right|\} = 0$ almost surely. Hence $\{\frac{1}{w(v)}I(U_t \leq v) : v \in (0, 1/2]\}$ is a Glivenko-Cantelli class. To show that $\{\frac{1}{w(v)}I(U_t \leq v) : v \in (1/2, 1)\}$ is also a Glivenko-Cantelli class, we note that $\frac{1}{w(v)}$ is symmetric about 1/2, decreasing in $v \in [0, 1/2]$, and $\int_0^1 \frac{1}{w(v)}dv < \infty$. As a result, it suffices to show that $\{\frac{1}{w(v)}[1 - I(U_t \leq v)] : v \in (1/2, 1)\}$ 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 $\{\frac{1}{w(v)}I(U_t \leq v) : v \in (0, 1/2]\}$ is a Donsker class. Again by the assumption on $w(\cdot)$, we have that the envelop function $F(U_t) \leq \frac{1}{w(U_t)}$. 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 $E[F(U_t)]^{2\gamma} < \infty$ with $\gamma > 1$ for β-mixing with the polynomial decay, or $E\{[F(U_t)]^2 \log[1+F(U_t)]\} < \infty$ for β-mixing with the exponential decay. Now we can apply theorem 1 in Doukhan, et al. (1995) to conclude that $\{\frac{1}{w(v)}I(U_t \leq v) : v \in (0, 1/2]\}$ is a Donsker class. ■

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.

Proof. (**Proposition 4.2**) Notice that by Assumption 1 and condition (C3) and Lemma 4.1, we have $||G_n - G^*||_{\mathcal{G}} = o_p(1)$ for the weight function $w(\cdot)$ stated in condition (C5). Under condition (C1), $\widetilde{\alpha}$ solves $\inf_{\alpha\in\mathcal{A}} \widetilde{Q}(\alpha)$, where

$$\widetilde{Q}(\alpha) = \{\frac{1}{n}\sum_{t=1}^{n} l_\alpha(G_n(Y_{t-1}), G_n(Y_t), \alpha))\}'\{\frac{1}{n}\sum_{t=1}^{n} l_\alpha(G_n(Y_{t-1}), G_n(Y_t), \alpha))\},$$

and α^* solves $\inf_{\alpha\in\mathcal{A}} Q(\alpha)$, where

$$Q(\alpha) = \{E[l_\alpha(G^*(Y_{t-1}), G^*(Y_t), \alpha))]\}'\{E[l_\alpha(G^*(Y_{t-1}), G^*(Y_t), \alpha))]\}.$$

Again under conditions (C1) and (C2.i), it suffices to show that

$$\sup_{\alpha\in\mathcal{A}} ||\frac{1}{n}\sum_{t=1}^{n} l_\alpha(G_n(Y_{t-1}), G_n(Y_t), \alpha) - E[l_\alpha(G^*(Y_{t-1}), G^*(Y_t), \alpha)]|| = o_p(1)$$

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

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

It remains to show that

$$(*) \quad \sup_{\alpha\in\mathcal{A}} ||\frac{1}{n}\sum_{t=1}^{n} l_\alpha(G^*(Y_{t-1}), G^*(Y_t), \alpha) - E[l_\alpha(G^*(Y_{t-1}), G^*(Y_t), \alpha)]|| = o_p(1)$$

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

$$\begin{aligned}
\sup_{\alpha\in\mathcal{A}, ||\alpha-\alpha_i||\leq\delta} ||l_\alpha(G^*(Y_{t-1}), G^*(Y_t), \alpha) - l_\alpha(G^*(Y_{t-1}), G^*(Y_t), \alpha_i)|| &\leq \varepsilon \\
\sup_{\alpha\in\mathcal{A}, ||\alpha-\alpha_i||\leq\delta} ||E\{l_\alpha(G^*(Y_{t-1}), G^*(Y_t), \alpha) - l_\alpha(G^*(Y_{t-1}), G^*(Y_t), \alpha_i)\}|| &\leq \varepsilon.
\end{aligned}$$

Hence,

$$\sup_{\alpha\in\mathcal{A}, ||\alpha-\alpha_i||\leq\delta} ||\frac{1}{n}\sum_{t=1}^{n}\{l_\alpha(G^*(Y_{t-1}), G^*(Y_t), \alpha) - l_\alpha(G^*(Y_{t-1}), G^*(Y_t), \alpha_i)\}|| \leq \varepsilon$$

$$\sup_{\alpha\in\mathcal{A}, ||\alpha-\alpha_i||\leq\delta} ||\mu_n\left(l_\alpha(G^*(Y_{t-1}), G^*(Y_t), \alpha)\right) - \mu_n\left(l_\alpha(G^*(Y_{t-1}), G^*(Y_t), \alpha_i)\right)|| \leq 2\varepsilon.$$

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

$$\max_{1\le i\le m} ||\mu_n\left(l_\alpha(G^*(Y_{t-1}),G^*(Y_t),\alpha_i)\right)|| = o_p(1).$$

Hence (*) is valid. ■

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

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

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

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

$$\begin{aligned} & \left|\begin{array}{c} E\{l_\alpha(G(Y_{t-1}),G(Y_t),\overline{\alpha})\}-E\{l_{\alpha,\alpha}(U_{t-1},U_t,\alpha^*)[\overline{\alpha}-\alpha^*]\} \\ -\sum_{j=1}^2 E\{l_{\alpha,j}(U_{t-1},U_t,\alpha^*)[G(Y_{t-2+j})-G^*(Y_{t-2+j})]\} \end{array}\right| \\ = \ & o(||\overline{\alpha}-\alpha^*||)+o(||G-G^*||_{\mathcal{G},w_2}). \end{aligned}$$

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:

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

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

$$\begin{aligned} & E_Z\{l_{\alpha,\alpha}(U_{t-1},U_t,\alpha^*)[\tilde{\alpha}-\alpha^*]\}+\sum_{j=1}^2 E_Z\{l_{\alpha,j}(U_{t-1},U_t,\alpha^*)[G_n(Y_{t-2+j})-G^*(Y_{t-2+j})]\} \\ & +o(||\tilde{\alpha}-\alpha^*||)+o(||G_n-G^*||_{\mathcal{G},w_2})+\mu_n(l_\alpha(G^*(Y_{t-1}),G^*(Y_t),\alpha^*)) \\ = \ & o_p(n^{-1/2}). \end{aligned}$$

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

$$-E_Z\{l_{\alpha,\alpha}(U_{t-1},U_t,\alpha^*)\}[\tilde{\alpha}-\alpha^*]+o_p\left(||\tilde{\alpha}-\alpha^*||\right)=A_n^*+o_p(n^{-1/2})$$

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 $\tilde{\alpha}$ with $||\tilde{\alpha}-\alpha^*||=o_p(1)$,

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

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

Lemma A.2: Condition **(a)** is implied by Assumption 1, conditions A1-A3, A4 or A4', A5-A6.

Proof. We first show that $\{l_\alpha(G(Y_{t-1}), G(Y_t), \overline{\alpha}) : (\overline{\alpha}, G) \in \mathcal{F}_\delta\}$ is a Donsker class by applying theorem 1 of Doukhan, et al. (1995). Define the envelop function $F(Y_{t-1}, Y_t) = \sup_{(\overline{\alpha},G)\in\mathcal{F}_\delta} |l_\alpha(G(Y_{t-1}), G(Y_t), \overline{\alpha})|$. 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,

$$
\begin{aligned}
& |l_\alpha(G(Y_{t-1}), G(Y_t), \overline{\alpha}) - l_\alpha(G^*(Y_{t-1}), G^*(Y_t), \alpha^*)| \\
\le\ & |l_{\alpha\alpha}(G_\eta(Y_{t-1}), G_\eta(Y_t), \alpha_\eta)| \times ||\overline{\alpha} - \alpha^*|| \\
& + |l_{\alpha,1}(G_\eta(Y_{t-1}), G_\eta(Y_t), \alpha_\eta) w(G^*(Y_{t-1}))| \times ||G - G^*||_{\mathcal{G}} \\
& + |l_{\alpha,2}(G_\eta(Y_{t-1}), G_\eta(Y_t), \alpha_\eta) w(G^*(Y_t))| \times ||G - G^*||_{\mathcal{G}} \\
\le\ & \{ \sup_{(\alpha_\eta, G_\eta)\in\mathcal{F}_\delta} |l_{\alpha\alpha}(G_\eta(Y_{t-1}), G_\eta(Y_t), \alpha_\eta)|\} \times ||\overline{\alpha} - \alpha^*|| \\
& + \{ \sup_{(\alpha_\eta, G_\eta)\in\mathcal{F}_\delta} |l_{\alpha,1}(G_\eta(Y_{t-1}), G_\eta(Y_t), \alpha_\eta) w(G^*(Y_{t-1}))|\} \times ||G - G^*||_{\mathcal{G}} \\
& + \{ \sup_{(\alpha_\eta, G_\eta)\in\mathcal{F}_\delta} |l_{\alpha,2}(G_\eta(Y_{t-1}), G_\eta(Y_t), \alpha_\eta) w(G^*(Y_t))|\} \times ||G - G^*||_{\mathcal{G}}.
\end{aligned}
$$

Hence by conditions A5 and A6,

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

this and condition A4(i)(iii) or A4'(i)(iii) imply that all the conditions for Theorem 1 of Doukhan, et al. (1995) is satisfied, hence $\{l_\alpha(G(Y_{t-1}), G(Y_t), \overline{\alpha}) : (\overline{\alpha}, G) \in \mathcal{F}_\delta\}$ is a Donsker class, moreover for any $\delta_n \to 0$,

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

Under conditions A5 and A6, $E_Z\{|l_\alpha(G(Y_{t-1}), G(Y_t), \overline{\alpha}) - l_\alpha(G^*(Y_{t-1}), G^*(Y_t), \alpha^*)|^2\} \to 0$ as $||\overline{\alpha} - \alpha^*|| \to 0$ and $||G - G^*||_{\mathcal{G}} \to 0$. 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, page 455), condition (b) is implied by: (*) for some small $\epsilon > 0$,

$$
\sup\left\{ \left|\frac{dE_Z\{l_\alpha(U_{t-1}+\eta\triangle G(Y_{t-1}), U_t+\eta\triangle G(Y_t), \alpha^*+\eta\triangle\alpha)\}}{d\eta}\right| : ||\triangle\alpha|| + ||\triangle G||_{\mathcal{G},w_2} \le 1, |\eta| \le \epsilon \right\} < \infty.
$$

By condition A3,

$$
\begin{aligned}
& \left|\frac{dE_Z\{l_\alpha(U_{t-1}+\eta\triangle G(Y_{t-1}),U_t+\eta\triangle G(Y_t),\alpha^*+\eta\triangle\alpha)\}}{d\eta}\right| \\
= & \left|E_Z\left(\frac{dl_\alpha(U_{t-1}+\eta\triangle G(Y_{t-1}),U_t+\eta\triangle G(Y_t),\alpha^*+\eta\triangle\alpha)}{d\eta}\right)\right| \\
\leq & \ E_Z\left(\left|\frac{dl_\alpha(U_{t-1}+\eta\triangle G(Y_{t-1}),U_t+\eta\triangle G(Y_t),\alpha^*+\eta\triangle\alpha)}{d\eta}\right|\right) \\
\leq & \ E_Z\left(|l_{\alpha\alpha}(U_{t-1}+\eta\triangle G(Y_{t-1}),U_t+\eta\triangle G(Y_t),\alpha^*+\eta\triangle\alpha)|\right)\times||\triangle\alpha|| \\
& +E_Z\left(|l_{\alpha,1}(U_{t-1}+\eta\triangle G(Y_{t-1}),U_t+\eta\triangle G(Y_t),\alpha^*+\eta\triangle\alpha)w_2(G^*(Y_{t-1}))|\right)\times||\triangle G||_{\mathcal{G},w_2} \\
& +E_Z\left(|l_{\alpha,2}(U_{t-1}+\eta\triangle G(Y_{t-1}),U_t+\eta\triangle G(Y_t),\alpha^*+\eta\triangle\alpha)w_2(G^*(Y_t))|\right)\times||\triangle G||_{\mathcal{G},w_2}
\end{aligned}
$$

By Holder inequality,

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

Hence (*) is satisfied given conditions A5 and A6. ■

Proof. (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

[1] Andrews, D. (1991), "Heteroskedasticity and Autocorrelation Consistent Covariance Matrix Estimation," *Econometrica* 59, 817-858.

[2] Andrews, D. (1994), "Asymptotics for Semi-parametric Econometric Models via Stochastic Equicontinuity," *Econometrica,* 62, 43-72.

[3] Andrews, D. (2001), "Testing When a Parameter is on the Boundary of the Maintained Hypothesis," *Econometrica* 69, 683-734.

[4] Bickel, P., C. Klaassen, Y. Ritov and J. Wellner (1993), *Efficient and Adaptive Estimation for Semiparametric Models*, The Johns Hopkins University Press.

[5] Bouyé, E., N. Gaussel and M. Salmon (2002), "Investigating Dynamic Dependence Using Copulae," Manuscript, Financial Econometrics Research Center.

[6] Bouyé, E. and M. Salmon (2002), "Dynamic Copula Quantile Regression and Tail Area Dynamic Dependence in Forex Markets," Manuscript, Financial Econometrics Research Center.

[7] Bradley, R. (1986), "Basic Properties of Strong Mixing Conditions," in *Dependence in Probability and Statistics*, Eberlein, E. and M.S. Taqqu (eds.), page 165-192. Birkhauser: Boston.

[8] Chen, X. and Y. Fan (2002), "Efficient Semiparametric Estimation of Copulas," Manuscript, New York University and Vanderbilt University.

[9] Chen, X. and Y. Fan (2003), "Evaluating Density Forecasts via the Copula Approach," Forthcoming in *Finance Research Letters*.

[10] Chen, X. and Y. Fan (2004), "Pseudo-Likelihood Ratio Tests for Model Selection in Semiparametric Multivariate Copula Models," Manuscript, NYU and Vanderbilt University.

[11] Chen, X., Y. Fan and A. Patton (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.

[12] Chen, X., L.P. Hansen and M. Carrasco (1998), "Nonlinearity and Temporal Dependence", working paper, University of Chicago.

[13] Chen, X., O. Linton and I. van Keilegom (2003), "Estimation of Semiparametric Models when the Criterion Function is not Smooth," *Econometrica*, 71, 1591-1608.

[14] Cherubini, U. and E. Luciano (2002), "Multivariate Option Pricing With Copulas," forthcoming in *Applied Mathematical Finance*.

[15] Costinot, A., T. Roncalli, and J. Teiletche (2000): "Revisiting the Dependence Between Financial Markets with Copulas," Working Paper, Crédit Lyonnais.

[16] Darsow, W., B. Nguyen and E. Olsen (1992), "Copulas and Markov Processes," *Illinois Journal of Mathematics* 36, 600-642.

[17] Davydov, Y. (1973), "Mixing Conditions for Markov Chains," *Theory of Probability and Its Applications* XVIII, 312-328.

[18] Devroye, L. (1986), *Non-Uniform Random Variate Generation*, Springer-Verlag, New York.

[19] Doukhan, P., P. Massart, and E. Rio (1995), "Invariance Principles for Absolutely Regular Empirical Processes," *Ann. Inst. Henri Poincare* 31, 393-427.

[20] Down, D., S. P. Meyn, and R. L. Tweedie (1995), "Exponential and Uniform Ergodicity of Markov Processes," *The Annals of Probability* 23, 1671-1691.

[21] Duffie, D. and J. Pan (1997), "An Overview of Value at Risk," *Journal of Derivatives*, 4, 7-49.

[22] Embrechts, P. , A. McNeil, and D. Straumann (2002), "Correlation and Dependence Properties in Risk Management: Properties and Pitfalls," in M. Dempster, ed., *Risk Management: Value at Risk and Beyond*, Cambridge University Press, 176-223.

[23] Embrechts, P. , A. Hoing, and A. Juri (2003), "Using Copulae to Bound the Value-at-Risk for Functions of Dependent Risks," *Finance and Stochastics* 7(2), 145-167.

[24] Engle, R. and S. Manganelli (2002), "CAViaR: Conditional Autoregressive Value ar Risk by Regression Quantiles," forthcoming in *Journal of Business and Economic Statistics*.

[25] Fermanian, J.-D. (2003), "Goodness of Fit Tests for Copulas", mimeo, CREST.

[26] Frees, Edward W. and E. A. Valdez (1998), "Understanding Relationships Using Copulas," *North American Actuarial Journal* 2, 1-25.

[27] Frey, R. and A. McNeil (2001), "Modeling Dependent Defaults," Working Paper, Department of Mathematics, ETHZ.

[28] Gagliardini, P. and C. Gourieroux (2002a), "Duration Time Series Models with Proportional Hazard," Working paper, CREST and University of Toronto.

[29] Gagliardini, P. and C. Gourieroux (2002b), "Efficient Nonparametric Estimation of Models With Nonlinear Dependence," Working paper, CREST and University of Toronto.

[30] Gagliardini, P. and C. Gourieroux (2003), "Constrained Nonparametric Dependence With Application to Finance and Insurance," Working paper, CREST and University of Toronto.

[31] Genest, C., K. Ghoudi, and L. Rivest (1995), "A Semiparametric Estimation Procedure of Dependence Parameters in Multivariate Families of Distributions," *Biometrika* 82 (3), 543-552.

[32] Genest, C. and B. Werker (2001), "Conditions for the Asymptotic Semiparametric Efficiency of an Omnibus Estimator of Dependence Parameters in Copula Models," Working paper.

[33] Gourieroux, C. and J. Jasiak (2002), "Value at Risk," Manuscript, University of Toronto.

[34] Granger, C.W.J., T. Terasvirta and A. Patton (2003): "Common Factors in Conditional Distributions", forthcoming in *Journal of Econometrics*.

[35] Heckman, J. J. and B. E. Honoré (1989), "The Identifiability of the Competing Risks Model," *Biometrika* 76(2), 325-330.

[36] Hu, H. (1998), "Large Sample Theory for Pseudo-Maximum Likelihood Estimates in Semiparametric Models," Ph.D. Thesis, University of Washington.

[37] Hu, L. (2002), "Dependence Patterns Across Financial Markets: Methods and Evidence," Manuscript, Yale University.

[38] Hull, J. and A. White (1998): "Value at Risk when Daily Changes in Market Variables are not Normally Distributed", *Journal of Derivatives*, 5, 9-19.

[39] Hutchinson, T. and C. Lai (1990), *Continuous Bivariate Distributions, Emphasizing Applications*, Rumsby Scientific Publishing, Adelaide.

[40] Jarner, S. and G. Robert (2001), "Polynomial Convergence Rates of Markov Chains," Working paper, Lancaster University.

[41] Joe, H. (1997), *Multivariate Models and Dependence Concepts*, Chapman & Hall/CRC.

[42] Johnson, M. (1987), *Multivariate Statistical Simulation*, John Wiley & Sons, New York.

[43] Klaassen, C. and J. Wellner (2001), "Efficient Estimation in the Bivariate Normal Copula Model: Normal Margins are Least-Favorable," *Bernoulli* 3, 55-77.

[44] Koenker, R. and G. Bassett (1978), "Regression Quantiles," *Econometrica* 46, 33-50.

[45] Lee, L. (1982a), "Some Approaches to the Correction of Selectivity Bias," *Review of Economic Studies* 49, 355-372.

[46] Lee, L. (1982b), "A Bivariate Logit Model," Mimeographed Paper, Center for Econometrics and Decision Sciences, University of Florida.

[47] Lee, L. (1983), "Generalized Econometric Models with Selectivity," *Econometrica* 51, 507-512.

[48] Li, D. (2000), "On Default Correlation: A Copula Function Approach," *Journal of Fixed Income* 9, 43-54.

[49] Nelsen, R. (1999), *An Introduction to Copulas*, Springer.

[50] Newey, W.K. (1994), "The Asymptotic Variance of Semiparametric Estimators," *Econometrica,* 62, 1349-1382.

[51] Newey, W.K. and D. F. McFadden (1994), "Large sample estimation and hypothesis testing", in R.F. Engle III and D.F. McFadden (eds.), *The Handbook of Econometrics*, vol. 4. North-Holland, Amsterdam.

[52] Newey, W. and K. West (1987), "Hypothesis Testing with Efficient Method of Moments Estimation," *International Economic Review*, 28, 777-787.

[53] 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.

[54] 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.

[55] Patton, A.J. (2004), "On the Importance of Skewness and Asymmetric Dependence in Stock Returns for Asset Allocation," forthcoming, *Journal of Financial Econometrics.*

[56] Powell, J. (1991) "Estimation of Monotonic Regression Models Under Quantile Restriction", in Barnett, W., J. Powell and G. Tauchen (eds.), *Non-parametric and Semi-parametric Methods in Econometrics and Statistics.* New York: Cambridge University Press.

[57] Ray, S., R. Berk, and W. Bielby (1980), "Correcting Sample Selection Bias for Bivariate Logistic Distribution of Disturbances," *American Statistical Association, 1980 Proceedings of the Business and Economic Statistics Section*, 456-459.

[58] Rio, E. (1995), "A Maximal Inequality and Dependent Marcinkiewicz-Zygmund Strong Laws," *Annals of Probability* 23, 918-937.

[59] Robinson, P. (1983), "Nonparametric Estimators for Time Series," *Journal of Time Series Analysis*, 4, 185-207.

[60] Rockinger, M. and E. Jondeau (2002), "Conditional Dependency of Financial Series: An Application of Copulas," Manuscript, HEC School of Management, France.

[61] Rosenberg, J. (1999), "Semiparametric Pricing of Multivariate Contingent Claims," Working Paper, Department of Finance, Stern School of Business, New York University.

[62] Shao, Q. and H. Yu (1996), "Weak Convergence for Weighted Empirical Processes of Dependent Sequences," *The Annals of Probability* 24, 2098-2127.

[63] Shih, J. and T. Louis (1995), "Inferences on the Association Parameter in Copula Models for Bivariate Survival Data," *Biometrics* 51, 1384-1399.

[64] Sklar, A. (1959), "Fonctions de r'epartition 'a n dimensionset leurs marges," *Publ. Inst. Statis. Univ. Paris* 8, 229-231.

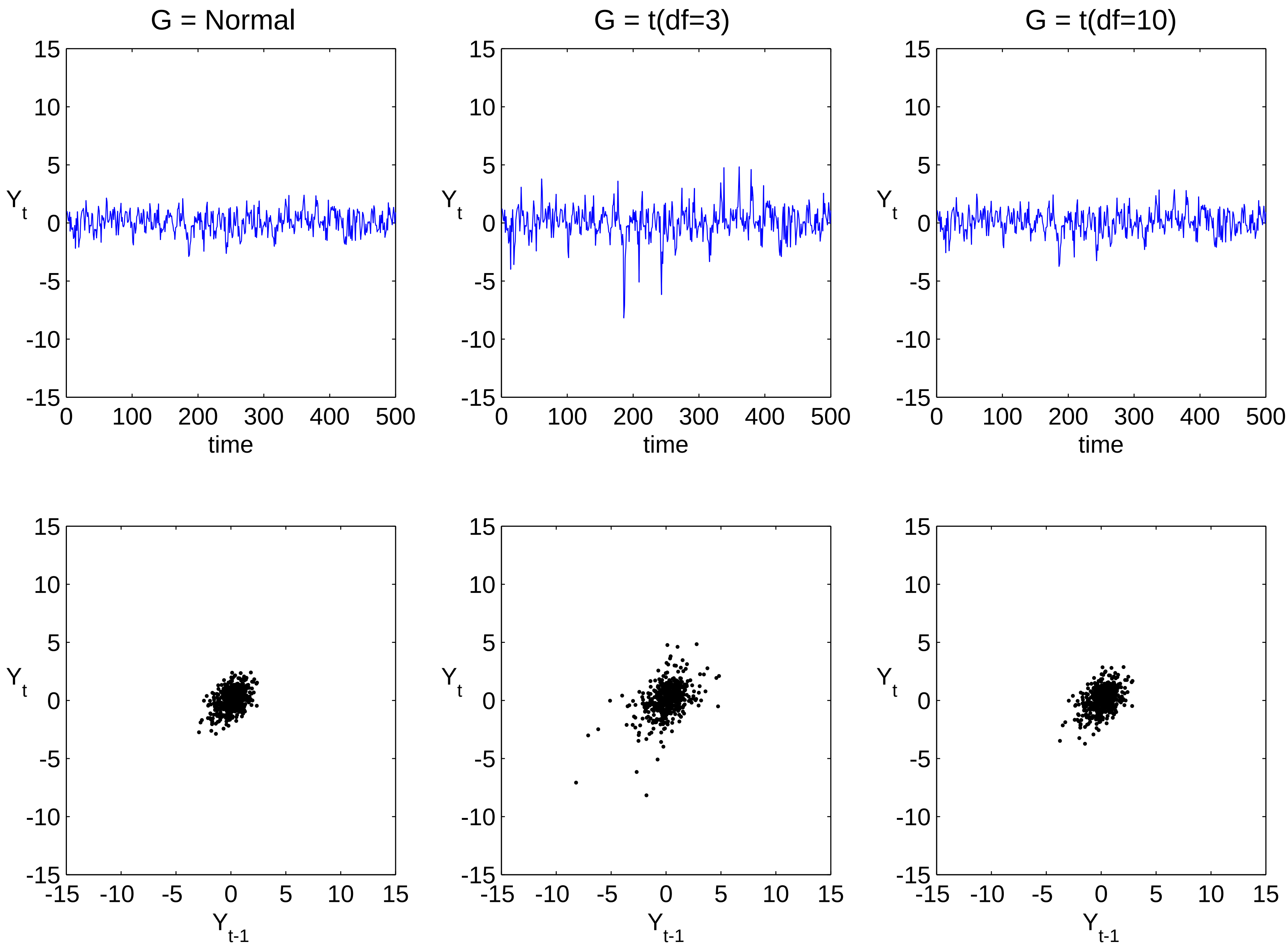

Figure 1: Gaussian Copula

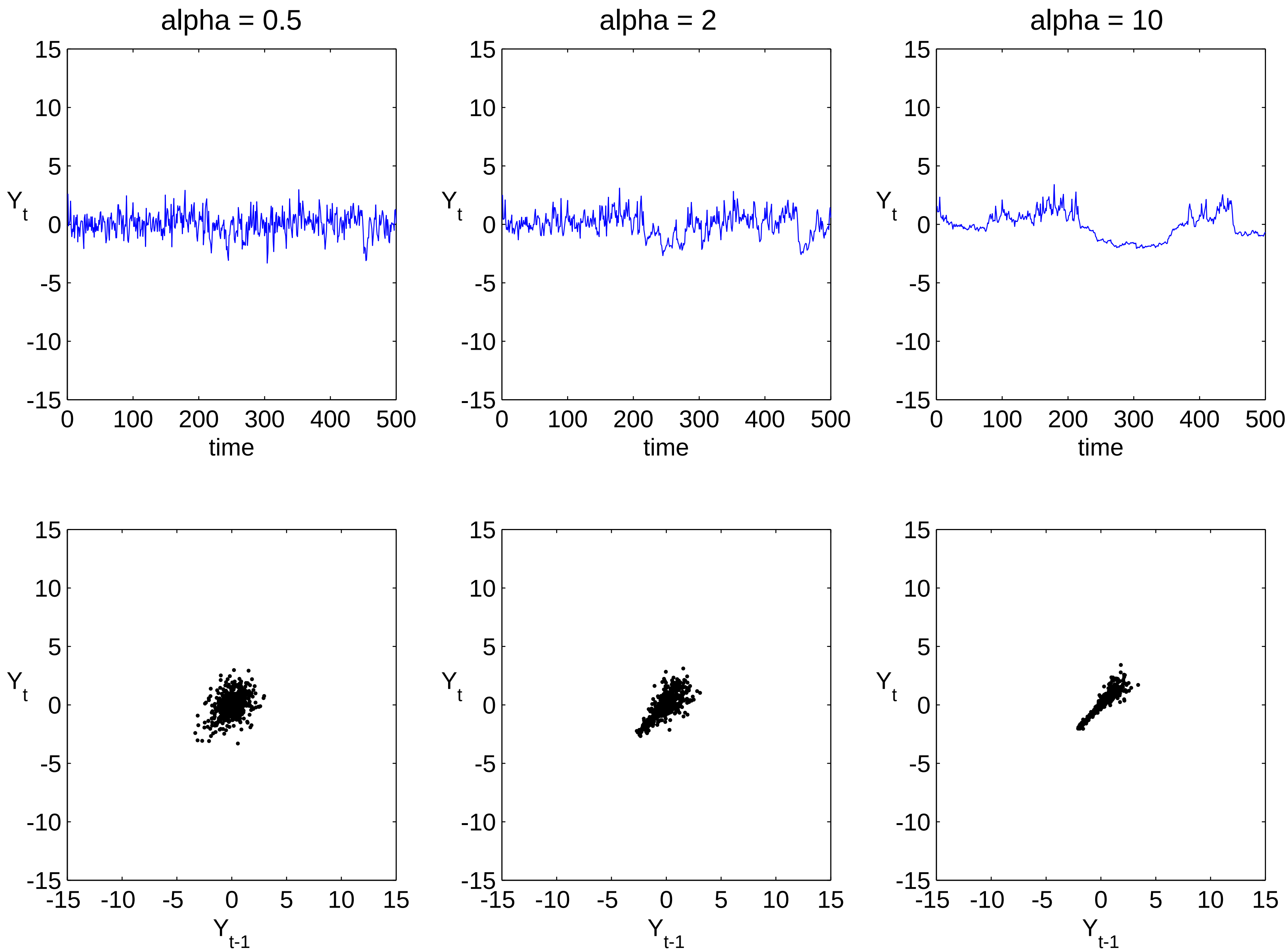

Figure 2a: Clayton Copula, G = normal

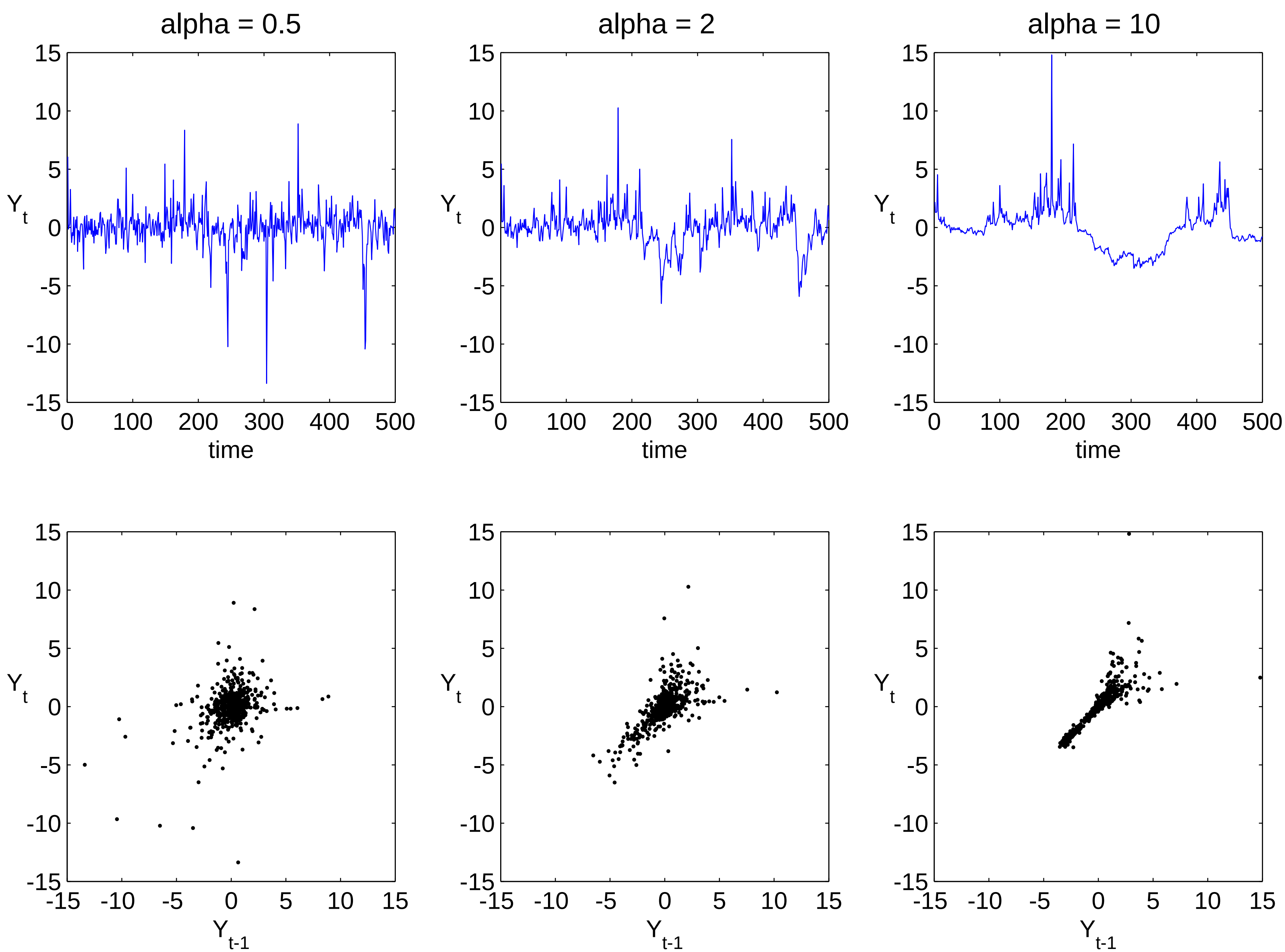

Figure 2b: Clayton Copula, G = t (df=3)