Copula-based nonlinear quantile autoregression

Authors: Xiaohong Chen, Roger Koenker, Zhijie Xiao

Persistent link: http://hdl.handle.net/2345/2493

This work is posted on eScholarship@BC,
Boston College University Libraries.

Pre-print version of an article published in Econometrics Journal 12(1): S50-S67.
doi:10.1111/j.1368-423X.2008.00274.x.

These materials are made available for use in research, teaching and private study, pursuant to U.S. Copyright Law. The user must assume full responsibility for any use of the materials, including but not limited to, infringement of copyright and publication rights of reproduced materials. Any materials used for academic research or otherwise should be fully credited with the source. The publisher or original authors may retain copyright to the materials.

Pre-print version of an article published in Econometrics Journal 12(1): S50-S67. doi:10.1111/j.1368-423X.2008.00274.x.

Copula-Based Nonlinear Quantile Autoregression*

Xiaohong Chen[†], Roger Koenker[‡], and Zhijie Xiao[§]

August 9, 2008

Abstract

Parametric copulas are shown to be attractive devices for specifying quantile autoregressive models for nonlinear time-series. Estimation of local, quantile-specific copula-based time series models offers some salient advantages over classical global parametric approaches. Consistency and asymptotic normality of the proposed quantile estimators are established under mild conditions, allowing for global misspecification of parametric copulas and marginals, and without assuming any mixing rate condition. These results lead to a general framework for inference and model specification testing of extreme conditional value-at-risk for financial time series data.

Keywords: Quantile autoregression, Copula, Ergodic nonlinear Markov models

---

[*] Chen and Koenker gratefully acknowledge financial support from National Science Foundation grants SES-0631613 and SES-0544673, respectively.

[†] Cowles Foundation for Research in Economics, Yale University, Box 208281, New Haven, CT 06520, USA. Tel: 203-432-5852. Email: xiaohong.chen@yale.edu

[‡] Department of Economics, University of Illinois at Urbana-Champaign, Champaign, IL 61820, USA. Tel: 217-333-4558. Email: rkoenker@uiuc.edu

[§] Department of Economics, Boston College, Chestnut Hill, MA 02467, USA. Tel: 617-552-1709. Email: xiaoz@bc.edu. Tsinghua University, Beijing, 100084, China.

Pre-print version of an article published in Econometrics Journal 12(1): S50-S67. doi:10.1111/j.1368-423X.2008.00274.x.

1. Introduction

Estimation of models for conditional quantiles constitutes an essential ingredient in modern risk assessment. And yet, often, such quantile estimation and prediction rely heavily on unrealistic global distributional assumptions. In this paper we consider new estimation methods for conditional quantile functions that are motivated by parametric copula models, but retain some semi-parametric flexibility and thus, should deliver more robust and more accurate estimates, while also being well-suited to the evaluation of misspecification.

We employ parametric copula models to generate nonlinear-in-parameters quantile autoregression (QAR) models. Such models have several advantages over the linear QAR models previously considered in Koenker and Xiao (2006) since, by construction, the copula-based nonlinear QAR models are globally plausible with monotone conditional quantile functions over the entire support of the conditioning variables. Rather than imposing this global structure, however, we choose instead to estimate the implied conditional quantile function independently, thereby facilitating an analysis of potential misspecification of the global structure.

Copula-based Markov models provide a rich source of potential nonlinear dynamics describing temporal dependence (and tail dependence). They also permit us to carefully distinguish the temporal dependence from the specification of the marginal (stationary) distribution of the response. Stationarity of the processes considered implies that only one marginal distribution is required for the specification in addition to the choice of a copula. See, e.g., Chen and Fan (2006), Ibragimov (2006), Patton (2008) and the references therein for more detailed discussions about copula-based Markov models.

Choice of the parametric specification of the copula, C, and the marginal distribution F, is a challenging problem. In this paper, we restrict our attention to settings in which the choices of C and F could be globally misspecified, yet, they yield correct specification of a conditional quantile function at a particular quantile. This is obviously a weaker condition than the direct assertion that we have correctly specified C and F themselves, since each of the conditional quantile functions we consider are permitted to have their own vector of quantile-specific parameters. Indeed, this distinction between global parametric models and local, quantile-specific, ones is essential throughout the quantile regression literature, and facilitates inference for misspecification that arises from discrepancies in the quantile specific estimates of the model parameters (see Koenker (2005)). Moreover, we are able to derive the consistency and asymptotic normality of our quantile estimator under mild sufficient conditions. In particular, we only assume that the underlying copula-based Markov model is stationary ergodic, without requiring any mixing conditions, and our moment restrictions are only those necessary for the validity of a central limit theorem (even for independent and identically distributed data). Our results are relevant for estimation and inference about extreme conditional quantiles (or value-at-risk) for financial time series data, as such data typically display strong temporal dependence and tail dependence as well as heavy-tailed marginals.

Chen and Fan (2006) and Bouyé and Salmon (2008) have also suggested methods for estimating copula-based conditional quantile models. Both papers assume correct specification of the parametric copula dependence function $C(\cdot;\alpha)$ (without specifying the marginal distribution F). Chen and Fan (2006) first estimate the marginal F by a rescaled empirical marginal CDF, and then estimate the copula parameter α via maximum likelihood. Conditional quantile functions are then obtained by plugging in the estimated copula parameter and the empirical marginal CDF. This approach obviously relies heavily on the correct specification of the parametric copula function. Bouyé and Salmon (2008) propose to estimate several distinct, nonlinear quantile regression models implied by their copula specification. This is essentially the approach adopted here. Bouyé and Salmon (2008) refer to Chen and Fan (2006) for conditions and justifications of the asymptotic properties of their estimator. While Chen and Fan (2006) derive the asymptotic properties of their two-step estimator under the assumptions that the parametric copula is correctly specified and the time series is beta-mixing with fast enough decay rate, we obtain the asymptotic properties of the copula-based quantile estimator allowing for misspecified parametric copula and without any mixing condition.

The plan of the paper is as follows: We introduce the copula-based QAR model in Section 2. Assumptions and asymptotic properties of the proposed estimator are developed in Section 3. Section 4 briefly describes statistical inference and Section 5 concludes. For simplicity of illustration and without loss of generality, we focus our analysis on first order QAR processes in our analysis.

2. Copula-Based Quantile Autoregression Models

2.1. First-order strictly stationary Markov models

To motivate copula-based quantile autoregression models, we start with a first-order strictly stationary Markov process, $\{Y_t\}_{t=1}^n$, whose probabilistic properties are determined by the true joint distribution of Y_{t-1} and Y_t, say, $G^*(y_{t-1}, y_t)$.

Suppose that $G^*(y_{t-1}, y_t)$ has continuous marginal distribution function $F^*(\cdot)$, then by Sklar's Theorem, there exists an unique copula function $C^*(\cdot, \cdot)$ such that

$$G^*(y_{t-1}, y_t) \equiv C^*(F^*(y_{t-1}), F^*(y_t)),$$

where the copula function $C^*(\cdot, \cdot)$ is a bivariate probability distribution function with uniform marginals.

Differentiating $C^*(u, v)$ with respect to u, and evaluate at $u = F^*(x), v = F^*(y)$, we obtain the conditional distribution of Y_t given $Y_{t-1} = x$:

$$\Pr\left[Y_t < y | Y_{t-1} = x\right] = \left.\frac{\partial C^*(u, v)}{\partial u}\right|_{u=F^*(x), v=F^*(y)} \equiv C_1^*(F^*(x), F^*(y)).$$

For any $\tau \in (0, 1)$, solving

$$\tau = \Pr\left[Y_t < y | Y_{t-1} = x\right] \equiv C_1^*(F^*(x), F^*(y))$$

for y (in terms of τ), we obtain the τ-th conditional quantile function of Y_t given $Y_{t-1} = x$:

$$Q_{Y_t}(\tau | x) = F^{*-1}(C_1^{*-1}(\tau; F^*(x))),$$

where $F^{*-1}(\cdot)$ signifies the inverse of $F^*(\cdot)$ and $C_1^{*-1}(\cdot; u)$ is the partial inverse of $C_1^*(u, v)$ with respect to $v = F^*(y_t)$. Denote $h^*(x) \equiv C_1^{*-1}(\tau; F^*(x))$, so we may rewrite the τ-th conditional quantile function of Y_t given $Y_{t-1} = x$ as[1]

$$Q_{Y_t}(\tau | x) = F^{*-1}(h^*(x)) \equiv H^*(x).$$

In this paper, we will work with the class of copula-based, first-order, strictly stationary Markov models. We allow for most commonly used parametric copula functions, excluding the Fréchet-Hoeffding upper bound and lower bounds.

[1] As we can see from the definition, both h^* and H^* depend on τ. We suppress τ from h^* and H^* for notational simplicity.

Assumption DGP: $\{Y_t : t = 1, ..., n\}$ *is a sample from a stationary first-order Markov process generated from* $(F^*(\cdot), C^*(\cdot, \cdot))$, *where* $F^*(\cdot)$ *is the true invariant distribution and is absolutely continuous with respect to Lebesgue measure on the real line; the copula* $C^*(\cdot, \cdot)$ *for* (Y_{t-1}, Y_t)*is absolutely continuous with respect to Lebesgue measure on* $[0, 1]^2$, *and is neither the Fréchet-Hoeffding upper or lower bounds:* $\min \{F^*(Y_{t-1}), F^*(Y_t)\}$ *or* $\max \{F^*(Y_{t-1}) + F^*(Y_t) - 1, 0\}$.

Denote $f^*(\cdot)$ and $c^*(\cdot, \cdot)$ as the density functions corresponding to the marginal distribution $F^*(\cdot)$ and the copula function $C^*(\cdot, \cdot)$ respectively. Assumption DGP is equivalent to assuming that $\{Y_t : t = 1, ..., n\}$ *is a sample from a first-order stationary Markov process generated from* $(f^*(\cdot), g^*(\cdot|\cdot))$, where

$$g^*(\cdot|y_{t-1}) \equiv f^*(\cdot)c^*(F^*(y_{t-1}), F^*(\cdot))$$

is the true conditional density function of Y_t given $Y_{t-1} = y_{t-1}$.

2.1.1. The Autoregressive Transformation Model

As demonstrated in Chen and Fan (2006), all the copula-based first order Markov models can be expressed in terms of an autoregressive transformation model. Let $U_t = F^*(Y_t)$, then under assumption DGP, $\{U_t\}$ is a strictly stationary first-order Markov process with the joint distribution of U_t and U_{t-1} given by the copula $C^*(\cdot, \cdot)$. Let $\Lambda_1()$ be any increasing transformation, then there exist monotone increasing functions Λ_2 and σ such that,

$$\Lambda_1(F^*(Y_t)) = \Lambda_2(F^*(Y_{t-1})) + \sigma(F^*(Y_{t-1}))\varepsilon_t$$

or equivalently,

$$U_t = F^*(Y_t) = \Lambda_1^{-1} (\Lambda_2(U_{t-1}) + \sigma(U_{t-1})\varepsilon_t),$$

where the conditional density of ε_t given $U_{t-1} = F^*(Y_{t-1}) = u_{t-1}$ is

$$\begin{aligned} f_{\varepsilon|F^*(Y_{t-1})=u_{t-1}}(\varepsilon) &= c^*(u_{t-1}, \Lambda_1^{-1}(\Lambda_2(u_{t-1}) + \sigma(u_{t-1})\varepsilon))/D(u_{t-1}) \\ &= \frac{c^*(F^*(Y_{t-1}), \Lambda_1^{-1}(\Lambda_2(F^*(Y_{t-1})) + \sigma(F^*(Y_{t-1}))\varepsilon))}{D(F^*(Y_{t-1}))} \end{aligned}$$

where $D(u) = \frac{d\Lambda_1(\Lambda_2(u)+\sigma(u)\varepsilon)}{d\varepsilon}$, and satisfies the condition that

$$\Lambda_2(u_{t-1}) = E\left[\Lambda_1(U_t)\right| U_{t-1} = u_{t-1}] = \int_0^1 \Lambda_1(u) \times c^*(u_{t-1}, u)du.$$

In the special case that $\Lambda_1(u) = u$, we obtain $U_t = \Lambda_2(U_{t-1}) + \sigma(U_{t-1})\varepsilon_t$, i.e.

$$F^*(Y_t) = \Lambda_2(F^*(Y_{t-1})) + \sigma(F^*(Y_{t-1}))\varepsilon_t,$$

with

$$\Lambda_2(u_{t-1}) = E\left[\Lambda_1(U_t)\right| U_{t-1} = u_{t-1}] = \int_0^1 uc^*(u_{t-1}, u)du = 1 - \int_0^1 C_1^*(u_{t-1}, u)du.$$

2.2. Copula-based parametric quantile autoregression models

In practice, neither the true copula function $C^*(\cdot, \cdot)$ nor the true marginal distribution function $F^*(\cdot)$ of $\{Y_t\}$ is known. If we model both parametrically, by $C(\cdot, \cdot; \alpha)$ and $F(y; \beta)$, depending on unknown parameters α, β, then the τ-th conditional quantile function of Y_t, $Q_{Y_t}(\tau|x)$, becomes a function of the unknown parameters α and β, i.e.

$$Q_{Y_t}(\tau|x) = F^{-1}(C_1^{-1}(\tau; F(x, \beta), \alpha), \beta).$$

Denoting $\theta = (\alpha', \beta')'$ and $h(x, \alpha, \beta) \equiv C_1^{-1}(\tau; F(x, \beta), \alpha)$, we will write,

$$Q_{Y_t}(\tau|x) = F^{-1}(h(x, \alpha, \beta), \beta) \equiv H(x; \theta). \tag{2.1}$$

This copula formulation of the conditional quantile functions provides a rich source of potential nonlinear dynamics. By varying the choice of the copula specification we can induce a wide variety of nonlinear QAR(1) dependence, and the choice of the marginal, choice of F enables us to consider a wide range of possible tail behavior as well.

Copula-based models have been widely used in finance, especially in estimating conditional quantiles as required for Value-at-Risk (VaR) assessment, motivated by possible nonlinearity in financial time series dynamics. However, in many financial time series applications, the nature of the temporal dependence may vary over the quantiles of the conditional distribution. We would like to stress that although the conditional quantile function specification in the above representation assumes the parameters to be identical across quantiles, our estimation methods do not impose this restriction. Thus, we permit the estimated parameters to vary with τ and this provides an important diagnostic feature of the methodology.

The proposed QAR model is based on (2.1) but we permit different parameter values over τ, and write the vector of unknown parameters as $\theta(\tau) = (\alpha(\tau)', \beta(\tau)')'$. With $h(x, \alpha, \beta) \equiv C_1^{-1}(\tau; F(x, \beta), \alpha)$, we obtain the following nonlinear QAR model:

$$Q_{Y_t}(\tau|Y_{t-1}) = F^{-1}(h(Y_{t-1}, \alpha(\tau), \beta(\tau)), \beta(\tau)) \equiv H(Y_{t-1}, \theta(\tau)). \tag{2.2}$$

This nonlinear form of the QAR model can capture a wide range of systematic influences of conditioning variables on the conditional distribution of the response. Koenker and Xiao (2006) considered linear-in-parameter QAR processes in studying similar specifications. Maintaining a linear specification in the QAR model, however, requires rather strong regularity assumptions on the domain of the associated random variables imposed to ensure quantile monotonicity. Relaxing those assumptions implies that the conditional quantile functions are no longer linear.

From this point of view, copula-based models provide an important path toward extending linear QAR models to nonlinear quantile autoregression specifications.

The above analysis may be easily extended to k-th order nonlinear QAR models, but we will resist the temptation to tax the readers' patience with the notation required to accomplish this.

2.3. Examples

Example 1: Gaussian Copula

Let $\Phi_\alpha(\cdot,\cdot)$ be the distribution function of bivariate normal distribution with mean zeros, variances 1, and correlation coefficient α, and Φ be the CDF of a univariate standard normal. The bivariate Gaussian copula is given by

$$\begin{aligned} C(u,v;\alpha) &= \Phi_\alpha(\Phi^{-1}(u),\Phi^{-1}(v)) \\ &= \frac{1}{2\pi\sqrt{1-\alpha^2}}\int_{-\infty}^{\Phi^{-1}(u)}\int_{-\infty}^{\Phi^{-1}(v)}\exp\left\{-\frac{(s^2-2\alpha st+t^2)}{2(1-\alpha^2)}\right\}dsdt. \end{aligned}$$

Let $\{Y_t\}$ be a stationary Markov process of order 1 generated from a Gaussian copula $C^*(u,v)=\Phi_\alpha(\Phi^{-1}(u),\Phi^{-1}(v))$ and a marginal distribution $F^*(\cdot)$. Denote $U_t=F^*(Y_t)$, then the joint distribution of U_t and U_{t-1} is

$$C(u_{t-1},u_t;\alpha)=\Phi_\alpha(\Phi^{-1}(u_{t-1}),\Phi^{-1}(u_t)).$$

Differentiating $C(u_{t-1},u_t;\alpha)$ with respect to u_{t-1}, we obtain the conditional distribution of U_t given U_{t-1} :

$$C_1(u_{t-1},u_t;\alpha)=\Phi\left(\frac{\Phi^{-1}(u_t)-\alpha\Phi^{-1}(u_{t-1})}{\sqrt{1-\alpha^2}}\right).$$

For any $\tau\in[0,1]$, solving

$$\tau=C_1(u_{t-1},u_t;\alpha)=\Phi\left(\frac{\Phi^{-1}(u_t)-\alpha\Phi^{-1}(u_{t-1})}{\sqrt{1-\alpha^2}}\right)$$

for u_t, we obtain the τ-th conditional quantile function of U_t given $U_{t-1} = u_{t-1}$:

$$\begin{aligned} Q_{U_t}(\tau|u_{t-1}) &= \Phi\left(\alpha\Phi^{-1}(u_{t-1}) + \sqrt{1-\alpha^2}\Phi^{-1}(\tau)\right) \\ &= \Phi\left(\alpha\Phi^{-1}(F^*(y_{t-1})) + \sqrt{1-\alpha^2}\Phi^{-1}(\tau)\right) = h^*(\tau; F^*(y_{t-1}), \alpha). \end{aligned}$$

Let $Z_t = \Phi^{-1}(U_t) = \Phi^{-1}(F^*(Y_t))$. Then $\{Z_t\}$ is a Gaussian AR(1) process that can be represented by

$$Z_t = \alpha Z_{t-1} + \varepsilon_t$$

where $\varepsilon_t \sim N(0, (1-\alpha^2))$ and is independent of Z_{t-1}. We obtain the τ-th conditional quantile function of Z_t given Z_{t-1} :

$$Q_{Z_t}(\tau|Z_{t-1}) = b(\tau) + \alpha Z_{t-1}, \quad \text{with} \quad b(\tau) = \sqrt{1-\alpha^2}\Phi^{-1}(\tau),$$

a formulation that is the familiar linear AR(1) specification, which induces the simplest linear QAR model.

Example 2: Student-t copula

Let $\mathbf{t}_{\nu,\rho}(\cdot,\cdot)$ be the distribution function of bivariate Student-t distribution with mean zeros, variances 1, correlation coefficient ρ, and degrees of freedom ν. And let $t_\nu(\cdot)$ be the CDF of a univariate Student-t distribution with mean zero, variance 1, and degrees of freedom ν. The bivariate $t-$ copula is given by, with $\alpha = (\nu, \rho)$

$$\begin{aligned} &C(u, v; \alpha) = \mathbf{t}_{\nu,\rho}(t_\nu^{-1}(u), t_\nu^{-1}(v)) \\ = &\frac{1}{2\pi\sqrt{1-\rho^2}} \int_{-\infty}^{t_\nu^{-1}(u)} \int_{-\infty}^{t_\nu^{-1}(v)} \left\{1 + \frac{(s^2 - 2\rho st + t^2)}{\nu(1-\rho^2)}\right\}^{-(\nu+2)/2} dsdt. \end{aligned}$$

Let $\{Y_t\}$ be a stationary Markov process of order 1 generated from a standard bivariate t_ν-copula function $C^*(u, v) = \mathbf{t}_{\nu,\rho}(t_\nu^{-1}(u), t_\nu^{-1}(v))$ and a marginal distribution function $F^*(\cdot)$. Let $U_t = F^*(Y_t)$, then the τ-th conditional quantile function of U_t given U_{t-1} is given by

$$Q_{U_t}(\tau|\mathcal{F}_{t-1}) = t_\nu \left(\rho t_\nu^{-1}(F^*(Y_{t-1})) + \sigma(F^*(Y_{t-1}))t_{\nu+1}^{-1}(\tau)\right) = h^*(\tau; F^*(Y_{t-1}), \rho, \nu),$$

where

$$\sigma(F^*(Y_{t-1})) = \sqrt{\frac{\nu + \left[t_\nu^{-1}(F^*(Y_{t-1}))\right]^2}{\nu+1}(1-\rho^2)}.$$

Moreover, the transformed process $\{Z_t = t_\nu^{-1}(U_t) = t_\nu^{-1}(F^*(Y_t))\}$ is a Student-t process that can be represented by

$$Z_t = \rho Z_{t-1} + \sigma(Z_{t-1})e_t,$$

where $e_t \sim t_{\nu+1}$, and is independent of Y_{t-1},

$$\sigma(Z_{t-1}) = \sqrt{\frac{\nu + Z_{t-1}^2}{\nu+1}(1-\rho^2)}$$

is a known function of $Z_{t-1} = t_\nu^{-1}(F^*(Y_{t-1}))$. (If and only if the true marginal distribution F^* is also t_ν then $Z_t = t_\nu^{-1}(F^*(Y_t)) = Y_t$). The τ-th conditional quantile function of Z_t given Z_{t-1}, is then given by

$$Q_{Z_t}(\tau|\mathcal{F}_{t-1}) = \rho Z_{t-1} + \sigma(Z_{t-1})t_{\nu+1}^{-1}(\tau).$$

Let $\theta(\tau) = (\rho, \alpha(\tau), \beta(\tau))$, where

$$\alpha(\tau) = \frac{\nu(1-\rho^2)t_{\nu+1}^{-1}(\tau)^2}{1+\nu}, \quad \beta(\tau) = \frac{(1-\rho^2)t_{\nu+1}^{-1}(\tau)^2}{1+\nu}$$

we can rewrite the conditional quantile function as

$$Q_{Z_t}(\tau|\mathcal{F}_{t-1}) = \rho Z_{t-1} + \sqrt{\alpha(\tau) + \beta(\tau)Z_{t-1}^2}.$$

This example can be generalized to any first-order Markov models that are generated from an elliptical copula[2] and an elliptical marginal distribution of the same form. The conditional mean is linear and conditional variance is homoskedastic if and only if the copula is Gaussian with Gaussian marginal.

The Gaussian copula does not exhibit tail dependence, while the Student-t copula and other elliptical copula have symmetric tail dependence. For many financial applications, copulas that possess asymmetric tail dependence properties are more appropriate.

Example 3: Joe-Clayton copula

The Joe-Clayton copula is given by:

$$C(u,v;\alpha) = 1 - \{1 - [(1-\bar{u}^k)^{-\gamma} + (1-\overline{v}^k)^{-\gamma} - 1]^{-1/\gamma}\}^{1/k},$$

where $\bar{u} = 1-u$, $\alpha = (k,\gamma)'$ and $k \geq 1$, $\gamma > 0$. It is known that the lower tail dependence parameter for this family is $\lambda_L = 2^{-1/\gamma}$ and the upper tail dependence parameter is $\lambda_U = 2 - 2^{1/k}$. When $k = 1$, the Joe-Clayton copula reduces to the Clayton copula:

$$C(u,v;\alpha) = [u^{-\alpha} + v^{-\alpha} - 1]^{-1/\alpha}, \quad \text{where} \quad \alpha = \gamma > 0.$$

When $\gamma \to 0$, the Joe-Clayton copula approaches the Joe copula whose concordance ordering and upper tail dependence increase as k increases. For other properties of the Joe-Clayton copula, see Joe (1997) and Patton (2006). When coupled with heavy-tailed marginal distributions such as the Student's t distribution, this family of copulas can generate time series with clusters of extreme

[2] An elliptical copula is a copula generated from an elliptically symmetric bivariate distribution.

values and hence provide alternative models for economic and financial time series that exhibit such clusters.

For the Joe-Clayton copula, one can easily verify that

$$\begin{aligned} C_1(u_{t-1}, u_t; \alpha) &= (1 - u_{t-1})^{k-1}(1 - \bar{u}_{t-1}^k)^{-(\gamma+1)} \\ &\quad \times [(1 - \bar{u}_{t-1}^k)^{-\gamma} + (1 - \bar{u}_t^k)^{-\gamma} - 1]^{-(\gamma^{-1}+1)} \\ &\quad \times [1 - \{(1 - \bar{u}_{t-1}^k)^{-\gamma} + (1 - \bar{u}_t^k)^{-\gamma} - 1\}^{-1/\gamma}]^{k^{-1}-1}. \end{aligned}$$

For any $\tau \in [0, 1]$, solving $\tau = C_1(u_{t-1}, u_t; \alpha)$ for u_t, we obtain the τ-th conditional quantile function of U_t given u_{t-1} based on the Clayton copula:

$$Q_{U_t}(\tau|u_{t-1}) = [(\tau^{-\alpha/(1+\alpha)} - 1)u_{t-1}^{-\alpha} + 1]^{-1/\alpha}$$

Note that this expression and the similar expressions in the foregoing examples provide a convenient mechanism with which to simulate observations from the respective models. See Bouyé and Salmon (2008) for additional examples of copula-based conditional quantile functions.

3. Asymptotic Properties

In this section, we study estimation of the copula-based QAR model (2.2). The vector of parameters $\theta(\tau)$ and thus the conditional quantile of Y_t can be estimated by the following nonlinear quantile autoregression:

$$\min_{\theta \in \Theta} \sum_t \rho_\tau(Y_t - H(Y_{t-1}, \theta)), \tag{3.1}$$

where $\rho_\tau(u) \equiv u(\tau - I(u < 0))$ is the usual check function (Koenker and Bassett (1978)). We denote the solution as $\widehat{\theta}(\tau) \equiv \arg\min_{\theta \in \Theta} \sum_t \rho_\tau(Y_t - H(Y_{t-1}, \theta))$.

Then the τ-th conditional quantile of Y_t given Y_{t-1}, can be estimated by

$$\widehat{Q}_{Y_t}(\tau|Y_{t-1}=x)=H(x,\widehat{\theta}(\tau))\equiv F^{-1}\left(C_1^{-1}\left(\tau,F(x,\widehat{\beta}(\tau)),\widehat{\alpha}(\tau)\right),\widehat{\beta}(\tau)\right).$$

3.1. Consistency

To facilitate our analysis, we define

$$C_1(u,v;\alpha)\equiv\frac{\partial C(u,v;\alpha)}{\partial u};\ c(u,v;\alpha)\equiv\frac{\partial^2 C(u,v;\alpha)}{\partial u\partial v}.$$

Denote $C_1^{-1}(\tau,u;\alpha)$ as the inverse function of $C_1(u,v;\alpha)$ with respect to the argument v, and $H(x,\theta)\equiv F^{-1}(C_1^{-1}(\tau,F(x;\beta),\alpha);\beta)$.

We first introduce some simple regularity conditions to ensure consistency of our QAR estimator $\widehat{\theta}(\tau)$.

A1. The parameter space Θ is a compact subset in $\Re^k$.

A2. (i) $F(\cdot;\beta)$ and $F^{-1}(\cdot;\beta)$ (the inverse function of $F(\cdot;\beta)$) are continuous with respect to all their arguments; (ii) the copula function $C(u,v;\alpha)$ is second order differentiable with respect to u and v, and has copula density $c(u,v;\alpha)$; (iii) $C_1^{-1}(\tau,u;\alpha)$ (the inverse function of $C_1(u,v;\alpha)$ with respect to v) is continuous in α and u.

A3. (i) The true τ-th conditional quantile of Y_t given Y_{t-1}, $Q_{Y_t}(\tau|Y_{t-1})$, takes the form $H(Y_{t-1},\theta(\tau))\equiv F^{-1}(C_1^{-1}(\tau,F(Y_{t-1};\beta(\tau)),\alpha(\tau));\beta(\tau))$ for a $\theta(\tau)=(\alpha(\tau)',\beta(\tau)')\in\Theta$ for almost all Y_{t-1}; (ii) The true unknown conditional density of Y_t given Y_{t-1}, $g^*(\cdot|Y_{t-1})$, is bounded and continuous, and there exist $\epsilon_1>0,p>0$ such that $\Pr\left[g^*(Q_{Y_t}(\tau|Y_{t-1}))\geq\epsilon_1\right]\geq p$.

A4. For any $\epsilon > 0$, there exists a $\delta > 0$ such that, for any $\|\theta - \theta(\tau)\| > \epsilon$,

$$E\left\{\Pr\left[|H(Y_{t-1},\theta) - Q_{Y_t}(\tau|Y_{t-1})| > \delta \mid g^*(Q_{Y_t}(\tau|Y_{t-1})) \geq \epsilon_1\right]\right\} > 0.$$

A5. (i) $E\left(\sup_{\theta\in\Theta}|H(Y_{t-1},\theta)|\right) < \infty$; (ii) $\{Y_t\}$ is stationary, ergodic and satisfies assumption DGP.

Assumptions A1 - A4 and A5(i) are mild regularity conditions that are typically imposed even for parametric nonlinear quantile regression of Y_t given x_t with i.i.d. data $\{(Y_t, x_t)\}_{t=1}^n$. Thus they are natural conditions for our nonlinear Markov model (with $x_t = Y_{t-1}$). Assumption A5(ii) is a very mild condition on temporal dependence of $\{Y_t\}$. Although we do not assume the correct specification of the parametric functional forms of the copula $C(\cdot,\alpha)$ and the marginal distribution $F(\cdot,\beta)$, we assume that the parametric functional form of the conditional quantile $H(Y_{t-1},\theta(\tau))$ is correct at the τ-th quantile (assumption A3(i)). Hence, we do not need any beta-mixing decay rate condition on $\{Y_t\}$ that is assumed in Chen and Fan (2006). See Beare (2008) for temporal dependence properties of copula-based strictly stationary Markov processes.

Theorem 3.1. *(Consistency) For any fixed $\tau \in (0,1)$, under Assumptions A1 - A5, we have: $\widehat{\theta}(\tau) = \theta(\tau) + o_p(1)$.*

3.2. Normality

We introduce the following additional notation:

$$\dot{H}_\theta(x,\theta) \equiv \frac{\partial H(x;\theta)}{\partial\theta},\ \ddot{H}_{\theta\theta}(x,\theta) \equiv \frac{\partial^2 H(x;\theta)}{\partial\theta\partial\theta^\top}.$$

Given the consistency of $\widehat{\theta}(\tau)$, we only need to impose the following additional conditions in a shrinking neighborhood of $\theta(\tau)$. Denote $\Theta_0 = A_0 \times B_0 = \{\theta = (\alpha', \beta')' \in \Theta : \|\theta - \theta(\tau)\| = o_p(1)\}$. We assume:

A6. (i) $\dot{H}_\theta(Y_{t-1}, \theta)$ and $\ddot{H}_{\theta\theta}(Y_{t-1}, \theta)$ are well defined and measurable for all $\theta \in \Theta_0$ and for almost all Y_{t-1}; (ii) $E[\sup_{\theta\in\Theta_0} \left|\dot{H}_\theta(Y_{t-1}, \theta)\right|^2] < \infty$; (iii) $E\left(\sup_{\theta\in\Theta_0} \left|\ddot{H}_{\theta\theta}(Y_{t-1}, \theta)\right|\right) < \infty$; (iv) $V(\tau)$ and $\Omega(\tau)$ are finite non-singular, where

$$\begin{aligned} V(\tau) &\equiv E\left[g^*(Q_{Y_t}(\tau|Y_{t-1}))\dot{H}_\theta(Y_{t-1}, \theta(\tau))\dot{H}_\theta(Y_{t-1}, \theta(\tau))^\top\right], \\ \Omega(\tau) &\equiv E\left[\dot{H}_\theta(Y_{t-1}, \theta(\tau))\dot{H}_\theta(Y_{t-1}, \theta(\tau))^\top\right]. \end{aligned} \tag{3.2}$$

We impose assumption A6(i)(iii) for simplicity. We could replace assumption A6(i)(iii) by assuming that only $\dot{H}_\theta(Y_{t-1}, \theta)$ exists for $\theta \in \Theta_0$ and satisfies some milder regularity conditions such as those imposed in Huber (1967) and Pollard (1985) for i.i.d. data, and Hansen et al. (1995) for stationary ergodic data, without the need of the existence of $\ddot{H}_{\theta\theta}(Y_{t-1}, \theta)$ satisfying A6(iii).

Comparing our assumptions A1-A6 to the regularity conditions imposed in earlier papers (e.g. Weiss (1991), White (1994), Engle and Mangenelli (2004), and the references therein) on parametric nonlinear quantile time series models, we do not need any mixing nor near epoch dependence of mixing process conditions (see our A5(ii)), and our moment requirement is also much weaker than the existing ones (see our A5(i) and A6(ii)(iii)). Both these relaxations are important for financial applications that typically exhibit persistent temporal dependence and heavy-tailed marginals.

Pre-print version of an article published in Econometrics Journal 12(1): S50-S67. doi:10.1111/j.1368-423X.2008.00274.x.

Denote $f(\cdot;\beta)$ as the parametric density of $F(\cdot;\beta)$, and

$$h(x,\alpha,\beta)=C_1^{-1}(\tau;u,\alpha)\big|_{u=F(x,\beta)}=C_1^{-1}(\tau;F(x,\beta),\alpha)$$

with $C_{1u}^{-1}(\tau;u,\alpha)=\frac{\partial C_1^{-1}(\tau;u,\alpha)}{\partial u}$, $\dot{h}_\alpha(x,\alpha,\beta)=\frac{\partial h(x,\alpha,\beta)}{\partial\alpha}$ and $\dot{F}_\beta(x,\beta)=\frac{\partial F(x,\beta)}{\partial\beta}$. Then $V(\tau)$ and $\Omega(\tau)$ defined in (3.2) can be expressed as follows:

$$V(\tau)=\begin{bmatrix} V_{\alpha\alpha}(\tau) & V_{\alpha\beta}(\tau)\\ V_{\beta\alpha}(\tau) & V_{\beta\beta}(\tau)\end{bmatrix},\ \Omega(\tau)=\begin{bmatrix} \Omega_{\alpha\alpha}(\tau) & \Omega_{\alpha\beta}(\tau)\\ \Omega_{\beta\alpha}(\tau) & \Omega_{\beta\beta}(\tau)\end{bmatrix}, \tag{3.3}$$

where

$$\begin{aligned}
V_{\alpha\alpha}(\tau) &= E\left[\frac{g^*(Q_{Y_t}(\tau|Y_{t-1}))}{\{f(Q_{Y_t}(\tau|Y_{t-1}))\}^2}\dot{h}_\alpha(Y_{t-1};\theta(\tau))\dot{h}_\alpha(Y_{t-1};\theta(\tau))^\top\right];\\
V_{\alpha\beta}(\tau) &= E\left[\frac{g^*(Q_{Y_t}(\tau|Y_{t-1}))}{f(Q_{Y_t}(\tau|Y_{t-1}))}\dot{h}_\alpha(Y_{t-1};\theta(\tau))\frac{\partial F^{-1}(h(Y_{t-1};\theta(\tau)),\beta(\tau))}{\partial\beta}^\top\right]\\
&+ E\left[\frac{g^*(Q_{Y_t}(\tau|Y_{t-1}))}{\{f(Q_{Y_t}(\tau|Y_{t-1}))\}^2}\dot{h}_\alpha(Y_{t-1};\theta(\tau))C_{1u}^{-1}(\tau;F(Y_{t-1},\beta(\tau)),\alpha(\tau))\dot{F}_\beta(Y_{t-1},\beta(\tau))^\top\right];\\
V_{\beta\alpha}(\tau) &= V_{\alpha\beta}(\tau)^\top;
\end{aligned}$$

$$\begin{aligned}
V_{\beta\beta}(\tau) &= E\left[g^*(Q_{Y_t}(\tau|Y_{t-1}))\frac{\partial F^{-1}(h(Y_{t-1};\theta(\tau)),\beta(\tau))}{\partial\beta}\frac{\partial F^{-1}(h(Y_{t-1};\theta(\tau)),\beta(\tau))}{\partial\beta}^\top\right]\\
&+ 2E\left[\frac{g^*(Q_{Y_t}(\tau|Y_{t-1}))}{f(Q_{Y_t}(\tau|Y_{t-1}))}\frac{\partial F^{-1}(h(Y_{t-1};\theta(\tau)),\beta(\tau))}{\partial\beta}C_{1u}^{-1}(\tau;F(Y_{t-1},\beta(\tau)),\alpha(\tau))\dot{F}_\beta(Y_{t-1},\beta(\tau))^\top\right]\\
&+ E\left[\frac{g^*(Q_{Y_t}(\tau|Y_{t-1}))}{\{f(Q_{Y_t}(\tau|Y_{t-1}))\}^2}\left[C_{1u}^{-1}(\tau;F(Y_{t-1},\beta(\tau)),\alpha(\tau))\right]^2\cdot\dot{F}_\beta(Y_{t-1},\beta(\tau))\dot{F}_\beta(Y_{t-1},\beta(\tau))^\top\right].
\end{aligned}$$

$$\begin{aligned}
\Omega_{\alpha\alpha}(\tau) &= E\left[\frac{1}{\{f(Q_{Y_t}(\tau|Y_{t-1}))\}^2}\dot{h}_\alpha(Y_{t-1};\theta(\tau))\dot{h}_\alpha(Y_{t-1};\theta(\tau))^\top\right];\\
\Omega_{\alpha\beta}(\tau) &= E\left[\frac{1}{f(Q_{Y_t}(\tau|Y_{t-1}))}\dot{h}_\alpha(Y_{t-1};\theta(\tau))\frac{\partial F^{-1}(h(Y_{t-1};\theta(\tau)),\beta(\tau))}{\partial\beta}^\top\right]\\
&+ E\left[\frac{1}{\{f(Q_{Y_t}(\tau|Y_{t-1}))\}^2}\dot{h}_\alpha(Y_{t-1};\theta(\tau))C_{1u}^{-1}(\tau;F(Y_{t-1},\beta(\tau)),\alpha(\tau))\dot{F}_\beta(Y_{t-1},\beta(\tau))^\top\right];\\
\Omega_{\beta\alpha}(\tau) &= \Omega_{\alpha\beta}(\tau)^\top;
\end{aligned}$$

$$
\begin{aligned}
\Omega_{\beta\beta}(\tau) &= E\left[\frac{\partial F^{-1}(h(Y_{t-1};\theta(\tau)),\beta(\tau))}{\partial\beta}\frac{\partial F^{-1}(h(Y_{t-1};\theta(\tau)),\beta(\tau))}{\partial\beta}^{\top}\right] \\
&+ 2E\left[\frac{\partial F^{-1}(h(Y_{t-1};\theta(\tau)),\beta(\tau))}{\partial\beta}\frac{C_{1u}^{-1}(\tau;F(Y_{t-1},\beta(\tau)),\alpha(\tau))\cdot\dot{F}_{\beta}(Y_{t-1},\beta(\tau))^{\top}}{f\left(Q_{Y_t}(\tau|Y_{t-1})\right)}\right] \\
&+ E\left[\frac{[C_{1u}^{-1}(\tau;F(Y_{t-1},\beta(\tau)),\alpha(\tau))]^2}{\{f\left(Q_{Y_t}(\tau|Y_{t-1})\right)\}^2}\cdot\dot{F}_{\beta}(Y_{t-1},\beta(\tau))\dot{F}_{\beta}(Y_{t-1},\beta(\tau))^{\top}\right].
\end{aligned}
$$

Theorem 3.2. *For any fixed $\tau\in(0,1)$, under Assumptions A1 - A6 and $\theta(\tau)\in int(\Theta)$, we have:*

$$
\sqrt{n}\left(\widehat{\theta}(\tau)-\theta(\tau)\right)\Rightarrow N(0,\tau(1-\tau)V(\tau)^{-1}\Omega(\tau)V(\tau)^{-1}),
$$

with $V(\tau)$ and $\Omega(\tau)$ are given in (3.2) (or (3.3) equivalently).

Remark 1. *When the marginal distribution function of Y is completely known $F(y,\beta)=F(y)$, $V(\tau)$ and $\Omega(\tau)$ reduce to the following simplified forms:*

$$
\begin{aligned}
V(\tau) &= E\left\{\frac{g^*(Q_{Y_t}(\tau|Y_{t-1}))}{[f(Q_{Y_t}(\tau|Y_{t-1}))]^2}\dot{h}_{\alpha}(Y_{t-1};\alpha(\tau))\dot{h}_{\alpha}(Y_{t-1};\alpha(\tau))^{\top}\right\}, \\
\Omega(\tau) &= E\left[\frac{1}{[f(Q_{Y_t}(\tau|Y_{t-1}))]^2}\dot{h}_{\alpha}(Y_{t-1};\alpha(\tau))\dot{h}_{\alpha}(Y_{t-1};\alpha(\tau))^{\top}\right].
\end{aligned}
$$

Remark 2. *When both the copula function $C^*(u,v)=C(u,v;\alpha)$ and the marginal distribution $F^*(y)=F(y;\beta)$ are correctly specified, the parameters $\theta(\tau)$ define an explicit one-dimensional manifold in Θ, as illustrated in the examples of Section 2.3. To the extent that the estimated $\hat{\theta}(\tau)$ departs from this curve we can infer various forms of misspecification. See, for example, Koenker and Xiao (2002).*

Pre-print version of an article published in Econometrics Journal 12(1): S50-S67. doi:10.1111/j.1368-423X.2008.00274.x.

4. Inference

The asymptotic normality of the QAR estimate also facilitates inference. In order to standardize the QAR estimator and remove nuisance parameters from the limiting distribution, we need to estimate the asymptotic covariance Matrix. In particular, we need to estimate $\Omega(\tau)$ and $V(\tau)$. Let

$$\widehat{Q}_{Y_t}(\tau|Y_{t-1}) \equiv H(Y_{t-1}, \widehat{\theta}(\tau)),$$

and let $\widehat{f} = f(\cdot, \widehat{\beta})$ be the plug-in estimate of the parametric marginal density function. Then $\Omega(\tau)$ can be estimated by

$$\widehat{\Omega}_n(\tau) = \left[\begin{array}{cc} \widehat{\Omega}_{n,\alpha\alpha}(\tau) & \widehat{\Omega}_{n,\alpha\beta}(\tau) \\ \widehat{\Omega}_{n,\beta\alpha}(\tau) & \widehat{\Omega}_{n,\beta\beta}(\tau) \end{array} \right],$$

with

$$\begin{aligned} \widehat{\Omega}_{n,\alpha\alpha}(\tau) &= \frac{1}{n}\sum_{t=1}^{n} \frac{1}{\{\widehat{f}(\widehat{Q}_{Y_t}(\tau|Y_{t-1}))\}^2} \dot{h}_\alpha(Y_{t-1}; \widehat{\theta}(\tau)) \dot{h}_\alpha(Y_{t-1}; \widehat{\theta}(\tau))^\top; \\ \widehat{\Omega}_{n,\alpha\beta}(\tau) &= \frac{1}{n}\sum_{t=1}^{n} \frac{1}{\widehat{f}(\widehat{Q}_{Y_t}(\tau|Y_{t-1}))} \dot{h}_\alpha(Y_{t-1}; \widehat{\theta}(\tau)) \frac{\partial F^{-1}(h(Y_{t-1}; \widehat{\theta}(\tau)), \widehat{\beta}(\tau))}{\partial \beta}^\top \\ &+ \frac{1}{n}\sum_{t=1}^{n} \frac{\dot{h}_\alpha(Y_{t-1}; \widehat{\theta}(\tau))}{\{\widehat{f}(\widehat{Q}_{Y_t}(\tau|Y_{t-1}))\}^2} C_{1u}^{-1}(\tau; F(Y_{t-1}, \widehat{\beta}(\tau)), \widehat{\alpha}(\tau)) \dot{F}_\beta(Y_{t-1}, \widehat{\beta}(\tau))^\top; \\ \widehat{\Omega}_{n,\beta\alpha}(\tau) &= \widehat{\Omega}_{n,\alpha\beta}(\tau)^\top; \end{aligned}$$

$$\begin{aligned} \widehat{\Omega}_{n,\beta\beta}(\tau) &= \frac{1}{n}\sum_{t=1}^{n} \frac{\partial F^{-1}(h(Y_{t-1}; \widehat{\theta}(\tau)), \widehat{\beta}(\tau))}{\partial \beta} \frac{\partial F^{-1}(h(Y_{t-1}; \widehat{\theta}(\tau)), \widehat{\beta}(\tau))}{\partial \beta}^\top \\ &+ \frac{2}{n}\sum_{t=1}^{n} \frac{\partial F^{-1}(h(Y_{t-1}; \widehat{\theta}(\tau)), \widehat{\beta}(\tau))}{\partial \beta} \frac{C_{1u}^{-1}(\tau; F(Y_{t-1}, \widehat{\beta}(\tau)), \widehat{\alpha}(\tau)) \dot{F}_\beta(Y_{t-1}, \widehat{\beta}(\tau))^\top}{\widehat{f}(\widehat{Q}_{Y_t}(\tau|Y_{t-1}))} \\ &+ \frac{1}{n}\sum_{t=1}^{n} \frac{[C_{1u}^{-1}(\tau; F(Y_{t-1}, \widehat{\beta}(\tau)), \widehat{\alpha}(\tau))]^2}{\{\widehat{f}(\widehat{Q}_{Y_t}(\tau|Y_{t-1}))\}^2} \dot{F}_\beta(Y_{t-1}, \widehat{\beta}(\tau)) \dot{F}_\beta(Y_{t-1}, \widehat{\beta}(\tau))^\top. \end{aligned}$$

Next, the true (unknown) conditional density of Y_t given Y_{t-1}, $g^*(Q_{Y_t}(\tau|Y_{t-1}))$, can be estimated by the difference quotients,

$$\widehat{g}(\widehat{Q}_{Y_t}(\tau|Y_{t-1})) = (\tau_i - \tau_{i-1})/(\hat{Q}_{Y_t}(\tau_i|Y_{t-1}) - \hat{Q}_{Y_t}(\tau_{i-1}|Y_{t-1})),$$

for some appropriately chosen sequence of $\{\tau_i\}$'s. Then the matrix $V(\tau)$ can be estimated by

$$\widehat{V}_n(\tau) = \begin{bmatrix} \widehat{V}_{n,\alpha\alpha}(\tau) & \widehat{V}_{n,\alpha\beta}(\tau) \\ \widehat{V}_{n,\beta\alpha}(\tau) & \widehat{V}_{n,\beta\beta}(\tau) \end{bmatrix}$$

with

$$\begin{aligned}
\widehat{V}_{n,\alpha\alpha}(\tau) &= \frac{1}{n}\sum_{t=1}^{n} \frac{\widehat{g}(\widehat{Q}_{Y_t}(\tau|Y_{t-1}))}{\{\widehat{f}(\widehat{Q}_{Y_t}(\tau|Y_{t-1}))\}^2} \dot{h}_\alpha(Y_{t-1};\widehat{\theta}(\tau))\dot{h}_\alpha(Y_{t-1};\widehat{\theta}(\tau))^\top; \\
\widehat{V}_{n,\alpha\beta}(\tau) &= \frac{1}{n}\sum_{t=1}^{n} \frac{\widehat{g}(\widehat{Q}_{Y_t}(\tau|Y_{t-1}))}{\widehat{f}(\widehat{Q}_{Y_t}(\tau|Y_{t-1}))} \dot{h}_\alpha(Y_{t-1};\widehat{\theta}(\tau)) \frac{\partial F^{-1}(h(Y_{t-1};\widehat{\theta}(\tau)),\widehat{\beta}(\tau))}{\partial\beta}^\top \\
&+ \frac{1}{n}\sum_{t=1}^{n} \left[\frac{\widehat{g}(\widehat{Q}_{Y_t}(\tau|Y_{t-1}))}{\{\widehat{f}(\widehat{Q}_{Y_t}(\tau|Y_{t-1}))\}^2} \dot{h}_\alpha(Y_{t-1};\widehat{\theta}(\tau)) C_{1u}^{-1}(\tau;F(Y_{t-1},\widehat{\beta}(\tau)),\widehat{\alpha}(\tau))\dot{F}_\beta(Y_{t-1},\widehat{\beta}(\tau))^\top \right]; \\
\widehat{V}_{n,\beta\alpha}(\tau) &= \widehat{V}_{n,\alpha\beta}(\tau)^\top; \\
\widehat{V}_{n,\beta\beta}(\tau) &= \frac{1}{n}\sum_{t=1}^{n} \widehat{g}(\widehat{Q}_{Y_t}(\tau|Y_{t-1})) \frac{\partial F^{-1}(h(Y_{t-1};\widehat{\theta}(\tau)),\widehat{\beta}(\tau))}{\partial\beta} \frac{\partial F^{-1}(h(Y_{t-1};\widehat{\theta}(\tau)),\widehat{\beta}(\tau))}{\partial\beta}^\top \\
&+ \frac{2}{n}\sum_{t=1}^{n} \frac{\widehat{g}(\widehat{Q}_{Y_t}(\tau|Y_{t-1}))}{\widehat{f}(\widehat{Q}_{Y_t}(\tau|Y_{t-1}))} \frac{\partial F^{-1}(h(Y_{t-1};\widehat{\theta}(\tau)),\widehat{\beta}(\tau))}{\partial\beta} C_{1u}^{-1}(\tau;F(Y_{t-1},\widehat{\beta}(\tau)),\widehat{\alpha}(\tau))\dot{F}_\beta(Y_{t-1},\widehat{\beta}(\tau))^\top \\
&+ \frac{1}{n}\sum_{t=1}^{n} \frac{\widehat{g}(\widehat{Q}_{Y_t}(\tau|Y_{t-1}))}{\{\widehat{f}(\widehat{Q}_{Y_t}(\tau|Y_{t-1}))\}^2} [C_{1u}^{-1}(\tau;F(Y_{t-1},\widehat{\beta}(\tau)),\widehat{\alpha}(\tau))]^2 \dot{F}_\beta(Y_{t-1},\widehat{\beta}(\tau))\dot{F}_\beta(Y_{t-1},\widehat{\beta}(\tau))^\top.
\end{aligned}$$

Wald type tests can then be constructed immediately based on the standardized QAR estimators using $\widehat{\Omega}_n(\tau)$ and $\widehat{V}_n(\tau)$. The copula-based QAR models and related quantile regression estimation also provide important information about specification. Specification of, say, the copula function may be investigated based on parameter constancy over quantiles, along the lines of Koenker and Xiao (2006).

In addition, specification of conditional quantile models can be studied based on the quantile autoregression residuals. For example, if we want to test the hypothesis of a general form:

$$\mathrm{H}_0\colon\ R(\theta(\tau)) = 0$$

where $R(\theta)$ is an q-dimensional vector of smooth functions of θ, with derivatives to the second order, the asymptotic normality derived from the previous section facilitates the construction of a Wald statistic. Let

$$\dot{R}(\theta(\tau)) = \left[\frac{\partial R_1(\theta)}{\partial\theta}, \ldots, \frac{\partial R_q(\theta)}{\partial\theta}\right]\Bigg|_{\theta=\theta(\tau)},$$

denote a $p \times q$ matrix of derivatives of $R(\theta)$, we can construct the following regression Wald statistic

$$W_{n,\tau} \equiv nR(\widehat{\theta}(\tau))^\top \left[\tau(1-\tau)\dot{R}(\widehat{\theta}(\tau))^\top \widehat{V}_n(\tau)^{-1}\widehat{\Omega}_n(\tau)\widehat{V}_n(\tau)^{-1}\dot{R}(\widehat{\theta}(\tau))\right]^{-1} R(\widehat{\theta}(\tau)).$$

Under the hypothesis and our regularity conditions, we have

$$W_{n,\tau} \Rightarrow \chi_q^2$$

where χ_q^2 has a central chi-square distribution with q degrees of freedom.

5. Conclusion

There are many competing approaches to broadening the scope of nonlinear time series modeling. We have argued that parametric copulas offer an attractive framework for specifying nonlinear quantile autoregression models. In contrast to fully parametric methods like maximum likelihood that impose a global parametric structure, estimation of distinct copula-based QAR models retains considerable semiparametric flexibility by permitting local, quantile-specific parameters.

Pre-print version of an article published in Econometrics Journal 12(1): S50-S67. doi:10.1111/j.1368-423X.2008.00274.x.

There are many possible directions for future development. Inference and specification diagnostics is clearly a priority. Extensions to methods based on nonparametric estimation of the invariant distribution are possible. Finally, semi-parametric modeling of the copula itself as a sieve appears to be a feasible strategy for expanding the menu of currently available parametric copulas.

Pre-print version of an article published in Econometrics Journal 12(1): S50-S67. doi:10.1111/j.1368-423X.2008.00274.x.

6. Appendix: Mathematical Proofs

Proof of Theorem 3.1: We denote Y_{t-1} as x_t. Then $\widehat{\theta}(\tau) = \arg\min_{\theta\in\Theta} \sum_t \rho_\tau(Y_t - H(x_t,\theta))$ where $\rho_\tau(u) \equiv u(\tau - I(u<0))$. Define

$$\varepsilon_t \equiv Y_t - Q_{Y_t}(\tau|x_t) \equiv Y_t - H(x_t, \theta(\tau)).$$

Then $Q_{\varepsilon_t}(\tau|x_t) = 0$ and

$$Y_t = H(x_t, \theta(\tau)) + \varepsilon_t, \quad \Pr[\varepsilon_t \leq 0|x_t] = \tau.$$

Denote

$$\overline{H}(x_t,\theta) \equiv H(x_t,\theta) - H(x_t,\theta(\tau)) \text{ and } q_\tau(Y_t, x_t, \theta) \equiv \rho_\tau(\varepsilon_t - \overline{H}(x_t,\theta)) - \rho_\tau(\varepsilon_t),$$

and

$$Q_n(\theta) \equiv \frac{1}{n}\sum_{t=1}^{n} q_\tau(Y_t, x_t, \theta).$$

Then it is easy to see that

$$\widehat{\theta}(\tau) = \arg\min_{\theta\in\Theta} Q_n(\theta) \quad \text{and} \quad \theta(\tau) = \arg\min_{\theta\in\Theta} E\left[Q_n(\theta)\right].$$

We apply theorem 2.1 of Newey and McFadden (1994) to establish consistency. The compactness of Θ (assumption A1), continuity of $E\left[Q_n(\theta)\right]$ with respect to $\theta \in \Theta$ (assumptions A2 and A3) are directly assumed. It remans to verify uniform convergence ($\sup_{\theta\in\Theta} \left|Q_n(\theta) - E\left[Q_n(\theta)\right]\right| = o_p(1)$), and that $\theta(\tau)$ is the unique minimizer of $E\left[Q_n(\theta)\right]$.

Notice that under assumptions A2 and A3, $q_\tau(Y_t, x_t, \theta)$ is continuous in $\theta \in \Theta$ and measurable in (Y_t, x_t). Since

$$\sup_{\theta\in\Theta} |q_\tau(Y_t, x_t, \theta)| = \sup_{\theta\in\Theta} \left|\rho_\tau(\varepsilon_t - \overline{H}(x_t, \theta)) - \rho_\tau(\varepsilon_t)\right| \leq \sup_{\theta\in\Theta} |H(x_t, \theta) - H(x_t, \theta(\tau))|,$$

we have $E\left(\sup_{\theta\in\Theta} |q_\tau(Y_t, x_t, \theta)|\right) < \infty$ under assumption A5(i). These and compactness of Θ (assumption A1) and stationary ergodicity of $\{Y_t\}$ (assumption A5(ii)) together imply that all the conditions of proposition 7.1 of Hayashi (2000) hold. Thus, by apply the uniform law of large numbers for stationary ergodic processes (see, e.g., proposition 7.1 of Hayashi (2000)), we obtain the uniform convergence: $\sup_{\theta\in\Theta} |Q_n(\theta) - E[Q_n(\theta)]| = o_p(1)$.

Next we verify that $E[Q_n(\theta)]$ is uniquely minimized at $\theta(\tau)$. Recall that the true but unknown conditional density and distribution function of Y_t given x_t are $g^*(\cdot|x_t)$ and $G^*(\cdot|x_t)$ respectively, and use the following identity

$$\begin{aligned}\rho_\tau(u-v) - \rho_\tau(u) &= -v\psi_\tau(u) + (u-v)\{I(0 > u > v) - I(0 < u < v)\} \\ &= -v\psi_\tau(u) + \int_0^v \{I(u \leq s) - I(u < 0)\}ds, \qquad (6.1)\end{aligned}$$

where

$$\psi_\tau(u) \equiv \tau - I(u < 0), \quad \text{and by definition} \quad E[\psi_\tau(\varepsilon_t)|x_t] = 0.$$

we have, with simplified notation $\overline{H}_t = \overline{H}(x_t, \theta)$,

$$q_\tau(Y_t, x_t, \theta) = \rho_\tau(\varepsilon_t - \overline{H}_t) - \rho_\tau(\varepsilon_t) = -\overline{H}_t\psi_\tau(\varepsilon_t) + \int_0^{\overline{H}_t} \{I(\varepsilon_t \leq s) - I(\varepsilon_t < 0)\}ds.$$

thus $E\left[Q_n(\theta)\right] = E\left\{E[q_\tau(Y_t, x_t, \theta)|x_t]\right\}$ and

$$\begin{aligned}
& E[q_\tau(Y_t, x_t, \theta)|x_t] \\
= \; & E\left\{\int_0^{\overline{H}_t} \{I(\varepsilon_t \le s) - I(\varepsilon_t < 0)\} ds|x_t\right\} \\
= \; & 1\left(\overline{H}_t > 0\right) E\left\{\int_0^{\overline{H}_t} I(0 \le \varepsilon_t \le s) ds|x_t\right\} + 1\left(\overline{H}_t < 0\right) E\left\{\int_{\overline{H}_t}^0 I(s \le \varepsilon_t \le 0) ds|x_t\right\}.
\end{aligned}$$

Notice that under Assumptions A3,

$$\begin{aligned}
& 1\left(\overline{H}_t > 0\right) E\left\{\int_0^{\overline{H}_t} I(0 \le \varepsilon_t \le s) ds|x_t\right\} \\
= \; & 1\left(\overline{H}_t > 0\right) \int_0^{\overline{H}_t} \left[\int_{Q_{Y_t}(\tau|x_t)}^{s+Q_{Y_t}(\tau|x_t)} g^*(y|x_t) dy\right] ds \\
\ge \; & 1\left(\overline{H}_t > 0\right) 1\left(g^*(Q_{Y_t}(\tau|x_t) \ge \epsilon_1)\right) \int_0^{\overline{H}_t} \left[\int_{Q_{Y_t}(\tau|x_t)}^{s+Q_{Y_t}(\tau|x_t)} g^*(y|x_t) dy\right] ds \\
\ge \; & \frac{\epsilon_1}{2} 1\left(\overline{H}_t > 0\right) 1\left(g^*(Q_{Y_t}(\tau|x_t) \ge \epsilon_1)\right) \overline{H}_t^2,
\end{aligned}$$

and similar result can be obtained for the case $\overline{H}_t < 0$. Thus,

$$E\left[Q_n(\theta)\right] \ge \frac{\epsilon_1}{2} E\left[1\left(g^*(Q_{Y_t}(\tau|x_t) \ge \epsilon_1)\right) \overline{H}_t^2\right],$$

which, under Assumption A4, is strictly positive. Thus for any $\varepsilon > 0$, $\overline{Q}_n(\theta)$ is bounded away from zero, uniformly in θ for $\|\theta - \theta(\tau)\| \ge \varepsilon$. ■

Proof of Theorem 3.2: We obtain the asymptotic normality using Pollard's (1985) approach. In particular, we apply Pollard's (1985) theorem 2 except that we replace his i.i.d. assumption by our stationary ergodic data assumption A5(ii), (note that we could also apply theorem 7.1 of Newey and McFadden (1994)). Recall that $\widehat{\theta}(\tau) = \arg\min_{\theta \in \Theta} \frac{1}{n} \sum_t \rho_\tau(Y_t - H(x_t, \theta))$, and under our theorem 1,

$\widehat{\theta}(\tau) \in \Theta_0$ with probability approaching one. Note that $\psi_\tau(u) \equiv \tau - I(u < 0)$ is the right-hand derivative of $\rho_\tau(u) \equiv u(\tau - I(u < 0))$. ($\rho_\tau(u)$ is everywhere differentiable with respect to u except at $u = 0$). Under assumption A6(i), the derivative of $\rho_\tau(Y_t - H(x_t, \theta))$ with respect to $\theta \in \Theta_0$ exists (except at the point $Y_t = H(x_t, \theta)$), and is given by

$$\varphi_{t\tau}(\theta) \equiv [\tau - I(Y_t < H(x_t, \theta))]\, \dot{H}_\theta(x_t, \theta).$$

By the mean value theorem,

$$\rho_\tau(Y_t - H(x_t, \theta)) = \rho_\tau(Y_t - H(x_t, \theta(\tau))) + (\theta - \theta(\tau))^\top \varphi_{t\tau}(\theta(\tau)) + \|\theta - \theta(\tau)\|\, r_t(\theta)$$

with

$$r_t(\theta) \equiv \frac{(\theta - \theta(\tau))^\top [\varphi_{t\tau}(\overline{\theta}) - \varphi_{t\tau}(\theta(\tau))]}{\|\theta - \theta(\tau)\|},$$

where $\overline{\theta} \in \Theta_0$ is in between θ and $\theta(\tau)$. Likewise,

$$E[\rho_\tau(Y_t - H(x_t, \theta))] = E[\rho_\tau(Y_t - H(x_t, \theta(\tau)))] + (\theta - \theta(\tau))^\top E[\varphi_{t\tau}(\theta(\tau))] + \|\theta - \theta(\tau)\|\, E[r_t(\theta)].$$

Since $E[\tau - I(Y_t < H(x_t, \theta(\tau)))|x_t] = 0$ under assumption A3, we have, under assumptions A3, A5 and A6(i)(iv), that $E[\rho_\tau(Y_t - H(x_t, \theta))]$ has a second-order (i.e., $E[\varphi_{t\tau}(\theta)]$ has a first-order) derivative at $\theta(\tau)$ that is nonsingular, and is given by

$$-V(\tau) \equiv -E\left\{ g^*(H(x_t, \theta(\tau)))\dot{H}_\theta(x_t, \theta(\tau))\dot{H}_\theta(x_t, \theta(\tau))^\top \right\}.$$

Thus condition (i) of Pollard's (1985) theorem 2 is satisfied. Condition (ii) of Pollard's (1985) theorem 2 is directly assumed ($\theta(\tau) \in int(\Theta)$), and his condition (iii) holds due to our theorem 1 ($\left\|\widehat{\theta}(\tau) - \theta(\tau)\right\| = o_P(1)$). We shall replace his condition (iv) by a CLT for stationary ergodic martingale difference data. Since

$$\begin{aligned} E[\varphi_{t\tau}(\theta(\tau))|x_t] &= E\left\{ E\left(\tau - I(Y_t < H(x_t, \theta(\tau)))|x_t\right) \dot{H}_\theta(x_t, \theta(\tau)) \right\} = 0, \\ Var[\varphi_{t\tau}(\theta(\tau))|x_t] &= \tau(1-\tau)\dot{H}_\theta(x_t, \theta(\tau))\dot{H}_\theta(x_t, \theta(\tau))^\top. \end{aligned}$$

Under assumptions A3, A5(ii) and A6(iv), we can apply the CLT for strictly stationary ergodic martingale difference sequence (see, e.g., Hayashi (2000, page 106)), and obtain:

$$\frac{1}{\sqrt{n}}\sum_{t=1}^{n}\varphi_{t\tau}(\theta(\tau)) \Rightarrow N(0,\tau(1-\tau)\Omega(\tau))$$

with

$$\Omega(\tau) \equiv E\left\{\dot{H}_\theta(x_t,\theta(\tau))\dot{H}_\theta(x_t,\theta(\tau))^\top\right\}.$$

Thus it remains to verify that condition (v) (stochastic differentiability) of Pollard's (1985) theorem 2 holds:

$$\sup_{\theta\in\mathbf{U}_n}\frac{\left|\frac{1}{\sqrt{n}}\sum_t\left(r_t(\theta)-E[r_t(\theta)]\right)\right|}{1+\sqrt{n}\left\|\theta-\theta(\tau)\right\|} \to 0 \quad \text{in probability}$$

for each sequence of balls $\{\mathbf{U}_n\}$ that shrinks to $\theta(\tau)$ as $n\to\infty$. Since

$$r_t(\theta) \equiv \frac{(\theta-\theta(\tau))^\top[\varphi_{t\tau}(\overline{\theta})-\varphi_{t\tau}(\theta(\tau))]}{\|\theta-\theta(\tau)\|},$$

Pollard's (1985) condition (v) holds provided that

$$\sup_{\theta\in\mathbf{U}_n}\left|\frac{1}{n}\sum_t\frac{[\varphi_{t\tau}(\theta)-\varphi_{t\tau}(\theta(\tau))]-E[\varphi_{t\tau}(\theta)-\varphi_{t\tau}(\theta(\tau))]}{\|\theta-\theta(\tau)\|}\right| \to 0 \quad \text{in probability}$$

for each sequence of balls $\{\mathbf{U}_n\}$ that shrinks to $\theta(\tau)$ as $n\to\infty$.

Recall that $\varphi_{t\tau}(\theta)\equiv[\tau-I(Y_t<H(x_t,\theta))]\,\dot{H}_\theta(x_t,\theta)$, we have:

$$\begin{aligned}
& \varphi_{t\tau}(\theta)-\varphi_{t\tau}(\theta(\tau)) \\
= \; & \dot{H}_\theta(x_t,\theta)\left[I(Y_t<H(x_t,\theta(\tau)))-I(Y_t<H(x_t,\theta))\right] \\
& +\left\{\dot{H}_\theta(x_t,\theta)-\dot{H}_\theta(x_t,\theta(\tau))\right\}\left[\tau-I(Y_t<H(x_t,\theta(\tau)))\right] \\
\equiv \; & R_{1t}(\theta)+R_{2t}(\theta).
\end{aligned}$$

Under assumption A6(i)(iii) we have: for all $\mathbf{U}_n \subseteq \Theta_0$,

$$E\left(\sup_{\theta\in\mathbf{U}_n}\left|\frac{R_{2t}(\theta)}{\|\theta-\theta(\tau)\|}\right|\right) \leq E\left(\sup_{\overline{\theta}\in\Theta_0}\left|\ddot{H}_{\theta\theta}(x_t,\overline{\theta})\right|\right) < \infty.$$

By assumption A3,

$$E[R_{2t}(\theta)] = E\left(\left\{\dot{H}_\theta(x_t,\theta)-\dot{H}_\theta(x_t,\theta(\tau))\right\}E\{\tau - I(Y_t < H(x_t,\theta(\tau)))|x_t\}\right) = 0.$$

Thus, under assumptions A5(ii) and A6(i)(iii), by the uniform law of large numbers for stationary ergodic processes, since $\mathbf{U}_n \subseteq \Theta_0 \subset \Theta$ we obtain:

$$\sup_{\theta\in\mathbf{U}_n}\left|\frac{1}{n}\sum_t \frac{R_{2t}(\theta)-E[R_{2t}(\theta)]}{\|\theta-\theta(\tau)\|}\right| = o_P(1)$$

for each sequence of balls $\{\mathbf{U}_n\}$ that shrinks to $\theta(\tau)$ as $n \to \infty$. Consequently, Pollard's (1985) condition (v) holds provided that

$$\sup_{\theta\in\mathbf{U}_n}\left|\frac{1}{n}\sum_t \frac{R_{1t}(\theta)-E[R_{1t}(\theta)]}{\|\theta-\theta(\tau)\|}\right| = o_P(1) \tag{6.2}$$

for each sequence of balls $\{\mathbf{U}_n\}$ that shrinks to $\theta(\tau)$ as $n \to \infty$.

For any positive sequence of decreasing numbers $\{\varepsilon_n\}$, denote $\mathbf{U}_n \equiv \{\theta \in \Theta_0 : \theta \neq \theta(\tau), \|\theta-\theta(\tau)\| < \varepsilon_n\}$. Then, under assumption A.6(i)(ii), we have:

$$\begin{aligned} & E\left(\sup_{\theta\in\mathbf{U}_n}\left|\frac{R_{1t}(\theta)}{\|\theta-\theta(\tau)\|}\right|\right) \\ \leq\ & E\left(\sup_{\theta\in\Theta_0}\left\|\dot{H}_\theta(x_t,\theta)\right\| \times E\left[\sup_{\theta\in\mathbf{U}_n}\frac{|I(Y_t<H(x_t,\theta(\tau)))-I(Y_t<H(x_t,\theta))|}{\|\theta-\theta(\tau)\|}|x_t\right]\right) \end{aligned}$$

For all $\theta \in \Theta_0$, under assumption A6(i)(iii), we have

$$H(x_t,\theta) = H(x_t,\theta(\tau)) + \dot{H}_\theta(x_t,\theta(\tau))^\top(\theta-\theta(\tau)) + \frac{(\theta-\theta(\tau))^\top\ddot{H}_{\theta\theta}(x_t,\overline{\theta})(\theta-\theta(\tau))}{2}$$

Pre-print version of an article published in Econometrics Journal 12(1): S50-S67. doi:10.1111/j.1368-423X.2008.00274.x.

with $E\left(\sup_{\overline{\theta}\in\Theta_0}\left|\ddot{H}_{\theta\theta}(x_t,\overline{\theta})\right|\right)<\infty$. Therefore, under assumptions A3 and A6(i)(iii), conditioning on x_t, there exists a small $\epsilon(x_t)>0$ such that for all $\theta\in\Theta_0$ with $\|\theta-\theta(\tau)\|<\epsilon(x_t)$, we have that $Y_t-H(x_t,\theta(\tau))$ and $Y_t-H(x_t,\theta)$ are of the *same sign*. Hence, under assumptions A3 and A6(i)(ii), conditioning on x_t and for any $\varepsilon_n\leq\epsilon(x_t)$ with $\varepsilon_n\searrow 0$, we have:

$$
\begin{aligned}
& E\left(\sup_{\theta\in\mathbf{U}_n}\frac{|I(Y_t<H(x_t,\theta(\tau)))-I(Y_t<H(x_t,\theta))|}{\|\theta-\theta(\tau)\|}|x_t\right)\\
\leq\; & E\left(\sup_{\theta\in\mathbf{U}_n:\|\theta-\theta(\tau)\|<\epsilon(x_t)}1\{\overline{H}_t>0\}\frac{I(Y_t<H(x_t,\theta))-I(Y_t<H(x_t,\theta(\tau)))}{\|\theta-\theta(\tau)\|}|x_t\right)\\
+\; & E\left(\sup_{\theta\in\mathbf{U}_n:\|\theta-\theta(\tau)\|<\epsilon(x_t)}1\{\overline{H}_t<0\}\frac{I(Y_t<H(x_t,\theta(\tau)))-I(Y_t<H(x_t,\theta))}{\|\theta-\theta(\tau)\|}|x_t\right)\\
\leq\; & \text{const}.g^*(H(x_t,\theta(\tau)))\times\sup_{\theta\in\Theta_0}\left\|\dot{H}_\theta(x_t,\theta)\right\|;
\end{aligned}
$$

hence for $\varepsilon_n\searrow 0$,

$$
E\left(\sup_{\theta\in\mathbf{U}_n}\left|\frac{R_{1t}(\theta)}{\|\theta-\theta(\tau)\|}\right|\right)\leq\text{const}.E\left(\{\sup_{\theta\in\Theta_0}\left\|\dot{H}_\theta(x_t,\theta)\right\|\}^2\times g^*(H(x_t,\theta(\tau)))\right)<\infty.
$$

This and the uniform law of large numbers for stationary ergodic processes now imply that (6.2) holds. Therefore Pollard's (1985) theorem 2 is applicable and we obtain the desired normality result: $\sqrt{n}(\widehat{\theta}(\tau)-\theta(\tau))\Rightarrow N(0,V(\tau)^{-1}\tau(1-\tau)\Omega(\tau)V(\tau)^{-1})$. ■

References

[1] Beare, B., 2008, Copulas and temporal dependence, Manuscript, University of California, San Diego.

[2] Bouyé, E., Salmon, M., 2008, Dynamic copula quantile regressions and tail area dynamic dependence in forex markets, Manuscript, Financial Econometrics Research Centre, Warwick Business School, UK.

[3] Chen, X. and Y. Fan, 2006, Estimation of Copula-Based Semiparametric Time Series Models, Journal of Econometrics, V.130, 307-335.

[4] Engle, R. and S. Mangenelli, 2004, CAViaR: Conditional Autoregressive Value at Risk by Regression Quantiles, Journal of Business and Economic Statistics, V. 22, 367-381.

[5] Hansen, L.P., J. Heaton and E. Luttmer, 1995, Econometric Evaluation of Asset Pricing Models, The Review of Financial Studies, V8, 237-274.

[6] Hayashi, F., 2000, Econometrics. Princeton University Press.

[7] Huber, P., The behavior of maximum likelihood estimates under nonstandard conditions, in Le Cam and Neyman (eds), Proceedings of Berkeley Symposium on Mathematical Statistics and Probability, Volume I, University of California Press, 1967.

[8] Ibragimov, R., 2006, Copulas-based characterizations and higher-order Markov processes. Manuscript, Harvard University.

[9] Koenker, R., 2005, Quantile Regression, Econometric Society Monographs no. 38, Cambridge University Press.

[10] Koenker, R. and G. Bassett, 1978, Regression Quantiles, Econometrica, V46, 33-49.

[11] Koenker, R. and Z. Xiao, 2006, Quantile Autoregression, Journal of the American Statistical Association, 2006, Vol. 101, 980-990.

[12] Koenker, R. and Z. Xiao, 2002, Inference on the Quantile Regression Process, Econometrica, 81, 1583-1612.

[13] 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.

[14] Patton, A. 2006, Modelling asymmetric exchange rate dependence, International Economic Review, 47, 527-556.

[15] Patton, A. 2008, Copula models for financial time series, forthcoming in Handbook of Financial Time Series, Springer-Verlag.

[16] Pollard, D., 1985, New ways to prove central limit theorems, Econometric Theory, 1, 295-313.

[17] Weiss, A., 1991, Estimating Nonlinear Dynamic Models Using Least Absolute Errors Estimation, Econometric Theory, 7, 46-68.

[18] White, H., 1994, Estimation, inference and specification analysis, Econometric Society Monographs no. 22, Cambridge University Press.