Original Article

Journal of Time Series Analysis

First version received March 2010

Published online in Wiley Online Library: 29 October 2010

(wileyonlinelibrary.com) DOI: 10.1111/j.1467-9892.2010.00696.x

Threshold quantile autoregressive models

Antonio F. Galvao Jr.[a], Gabriel Montes-Rojas[b] and Jose Olmo[c,*,†]

This article studies estimation and asymptotic properties of Threshold Quantile Autoregressive processes. In particular, we show the consistency of the threshold and slope parameter estimators for each quantile and regime, and derive the asymptotic normality of the slope parameter estimators. A Monte Carlo experiment shows that the standard ordinary least squares estimation method is not able to detect important nonlinearities produced in the quantile process. The empirical study concentrates on modelling the dynamics of the conditional distribution of unemployment growth after the second world war. The results show evidence of important heterogeneity associated with unemployment and strong asymmetric persistence of unemployment growth in the higher quantiles.

Keywords: Nonlinear models; quantile regression; threshold models.
JEL classifications: C14; C22; C32; C50.

1. INTRODUCTION

During the last 20 years, linear time series models have proven to be rather limited and inadequate for describing the dynamics of macroeconomic and financial time series. These data are plagued with nonlinear phenomena such as asymmetries in the occurrence and persistence of negative and positive shocks, time-irreversibility, different tail behaviour of the distribution of the data, and asymmetric effects in the mean process due to the heteroscedasticity in the data. These phenomena have sparked a growing interest in applied and theoretical econometrics for developing time series processes modelling these nonlinearities in various ways.

Among the models that have enjoyed greater popularity are the threshold autoregressive models (TAR) of Tong (1978, 1983, 1990), Tong and Lim (1980) and Tsay (1989). Extensions of these models to accommodate stylized facts in macroeconomics and finance are abundant in the literature. Thus, for modelling financial time series, Nelson (1991), Glosten *et al.* (1993) and Zakoian (1994) propose threshold models capable of describing the asymmetric dynamics produced by feedback and leverage effects observed in the conditional volatility process. On a related setting, Clarida *et al.* (2006) and Sarno *et al.* (2006) make use of threshold processes to model the asymmetric dynamics of exchange rates and of the term structure of interest rates respectively. An alternative to threshold models for modelling nonlinearities in financial and economic data are Smooth Transition Models (STAR). These models are characterized by an infinite number of regimes determined by a state variable that changes smoothly from one state to the other, see for example Chan and Tong (1986), Terasvirta (1994) and the references therein.

All these models are aimed to describe the dynamics of the conditional mean process. Methods based on mean regression are, however, inadequate to capture nonlinearities in the data produced by general forms of heterogeneity when different regimes coincide in the mean process but not in the entire quantile process. The idea that time series may have a different behaviour across quantiles has attracted considerable attention in the theoretical and empirical literatures (see for instance Koenker and Xiao (2006) and the references therein. Koenker and Xiao (2006) propose quantile autoregression models in which the autoregressive coefficients can be expressed as a monotone function of a single scalar random variable. A natural extension of this model would consider the existence of different regimes depending on the quantile of the series to be modelled. Thus, following this direction, Cai and Stander (2008) and Cai (2010) introduced Quantile Self-Exciting Threshold AutoRegressive (Q-SETAR) models within a Bayesian framework. Both papers focus on self-exciting processes, although they do not derive the asymptotic properties of the model parameter estimators. Related papers dealing with the asymptotic theory of relevant estimators for the quantile problem are Caner (2002), which develops the limit law for the least absolute deviation (LAD) estimator of the threshold parameter in linear regression, and more recently Kato (2009), which extends the scope of convexity arguments to the case where estimators are obtained as stochastic processes and applies this technique to LAD inference for threshold regression.

The aim of this article is to extend the studies of Koenker and Xiao (2006), Cai and Stander (2008) and Cai (2010) on threshold quantile autoregressive processes by studying estimation and asymptotic properties of the parameters of interest in the model. We show in a Monte Carlo experiment that standard ordinary least squares (OLS) estimation methods are not capable of detecting nonlinearities produced in the quantile process. The study of nonlinear dynamics of the conditional distribution of time series is a relevant and rather unexplored area in applied macroeconometrics and time series. The empirical section illustrates the usefulness of

[a]*University of Iowa*
[b]*City University London*
[c]*Centro Universitario de la Defensa and City University London*
[*]*Correspondence to: Jose Olmo, Centro Universitario de la Defensa de Zaragoza, Academia General Militar, Ctra, de Huesca s/n. 50.090 Zaragoza, Spain.*
[†]*E-mail: jolmo@unizar.es*

this nonlinear methodology for modelling the dynamics of unemployment growth in the US after the second world war. This empirical study extends the work of Koenker and Xiao (2006) by focusing on unemployment growth instead of unemployment rates and by contemplating the possibility of nonlinearities in the quantile process. Our study shows that the conditional distribution of future unemployment growth, and therefore the corresponding quantile process, respond asymmetrically and nonlinearly to current growth rates of unemployment. A model with three regimes seems the best suited to describe the dynamics of the quantile process associated with the time series. For example, the 90% conditional quantile of the series of unemployment growth shows an explosive behaviour when current unemployment growth levels are beyond 4.64%, and most notably beyond 6.92%. The other quantiles are also defined by three regimes; in this case determined by threshold values around −5% and −1% respectively. In all cases, the upper regime is modelled by an AR(1) model with positive and statistically significant coefficient indicating persistence in unemployment growth. Other findings common to all quantiles are the mean reverting character of the middle regime and the lack of memory observed in the lower regime. The latter observation provides evidence of a random behaviour in unemployment growth when unemployment decreases and the economy expands rapidly. These findings, especially those for the upper regime, suggest the convenience of using economic policies to stimulate labour markets when unemployment is set to grow so as to avoid levels in unemployment that can produce persistence in unemployment growth and trigger periods of economic and social distress.

The article is structured as follows. Section 2 discusses the family of threshold quantile autoregressive models and their statistical properties, in particular, estimation procedures, consistency and asymptotic normality of the estimators. Section 3 presents a small Monte Carlo experiment that shows the performance of quantile regression against OLS methods for estimating nonlinear quantile processes and also pays attention to the monotonicity assumption usually imposed in quantile regression models. Section 4 discusses an application to study nonlinearities in the dynamics of the quantile process of unemployment growth, and Section 5 concludes. Proofs are gathered in a mathematical appendix.

2. STATISTICAL MODEL, ESTIMATION AND INFERENCE

2.1. Statistical model

For simplicity in the exposition the following sections are restricted to studying nonlinear processes with at most two regimes. Note, however, that the methodology introduced here can be easily extended to more regimes. The empirical application in Section 4 will show this by implementing a T-QAR model that makes allowance for three regimes. To illustrate this methodology we start with a simple T-QAR model defined by a self-exciting threshold autoregressive (SETAR) structure of order one.

Let the τth conditional quantile function of y_t be written as

$$Q_{y_t}(\tau \mid \Im_t) = \begin{cases} \theta_{01}(\tau) + \theta_{11}(\tau) y_{t-1}, & y_{t-1} \le \gamma(\tau), \\ \theta_{02}(\tau) + \theta_{12}(\tau) y_{t-1}, & y_{t-1} > \gamma(\tau), \end{cases} \tag{1}$$

with $\Im_t$ denoting the σ-field generated by $\{y_s, s \le t-1\}$. The model allows for different values of the threshold parameter $\gamma(\cdot)$ across different quantiles and assumes monotonicity of the conditional quantiles. Koenker and Xiao (2006) present a discussion of the monotonicity of conditional quantile models in linear time series. Koenker and Xiao (2006) remark that this crossing problem appears more acute in the autoregressive case than in ordinary regression applications since the support of the design space, that is the set of covariates x_t that occur with positive probability, is determined within the model. In our simulation and application sections, we will compute the conditional quantile functions and evaluate the model with respect to the monotonicity assumption. The results show that monotonicity applies. However, although it exceeds the scope of this article, there are several methods to ensure monotonic fits for conditional quantile functions. For simple linear quantile autoregressive models Koenker and Xiao (2006) argue that what really matters is that one can find a linear reparameterization of the model that does exhibit comonotonicity over some relevant region of the covariate space. To solve crossing problems, He (1997), for example, proposed to impose a location-scale regression model, which naturally satisfies monotonicity. Koenker and Ng (2005) developed a computational method for quantile regression that imposes the noncrossing constraints in simultaneous fitting of a finite number of quantile curves. Mammen (1991) proposed two-step estimators, with mean estimation in the first step followed by isotonization in the second. Dette and Volgushev (2008) proposed to obtain monotonic quantile curves by applying an integral transform to a local polynomial estimate of the conditional distribution function, and derived pointwise limit theory for this estimator. More recently, Chernozhukov *et al.* (2010) introduced a natural monotonization of the empirical curves by sampling from the estimated nonmonotone model, and then taking the resulting conditional quantile curves which by construction are monotone in the probability index. This construction of the monotone curves may be seen as a bootstrap and as a sorting or monotone rearrangement of the original nonmonotone curve. Moreover, they show that the rearranged curve is closer to the true quantile curve in finite samples than the original curve is, and derive functional limit distribution theory for the rearranged curve to perform simultaneous inference on the entire quantile function.

Now we generalize the above class of models within the two-regime setting, and use a more compact notation. Consider a nonlinear process $\{y_t\}$ with two possible regimes defined by the real functions $q_t(\tau)$ and $\gamma(\tau)$. For a fixed τ, the first function is the threshold variable and the second the threshold value. To simplify the model, we assume that the threshold variable is the same across the quantile process, and denote it q_t. Thus, the T-QAR counterpart of SETAR specifications is given by $q_t = q(x_t)$ with $x_t = (1, y_{t-1}, y_{t-2}, \ldots, y_{t-p_y})'$. The conditional quantile of y_t is then modelled by $Q_{y_t}(\tau \mid \Im_t) = h(\tau, \gamma(\tau))$, with $h(\cdot,\cdot)$ a piecewise linear process defined by

$$h(\tau, \gamma(\tau)) = x_t(\gamma(\tau))'\theta(\tau), \tag{2}$$

where $\theta(\cdot) = (\theta_{01},\theta_{02},\theta_{11},\theta_{12},\ldots,\theta_{p_y1},\theta_{p_y2})'$ is a 2(1 + p_y) dimensional vector of parameters, and $x(\gamma(\cdot)) = x \otimes (1\{q \leq \gamma(\cdot)\},1\{q > \gamma(\cdot)\})'$, with $1\{\cdot\}$ an indicator function and $\otimes$ the Kronecker product.

2.2. Estimation and asymptotic properties

For models with a known threshold parameter $\gamma_0 \equiv \gamma_0(\tau)$, standard quantile regression (QR) estimation procedures can be applied. These methods consist on finding

$$\hat{\theta}_{\gamma_0}(\tau) = \arg\min_{\theta} \sum_{t=1}^{T} \rho_\tau(y_t - x_t(\gamma_0)'\theta), \tag{3}$$

where $\rho_\tau(u) = u(\tau - I(u < 0))$ as in Koenker and Bassett (1978). A simple modification of Kato (2009) study for LAD models can be used to prove that under some regularity conditions and for fixed γ, there exists a unique solution to the minimization problem (3). Koenker and Xiao (2006) show the asymptotic properties of the standard linear quantile autoregressive estimators. To generalize Koenker and Xiao (2006) to the regime switching framework we need the following set of assumptions.

Assumptions:

A1: (y_t,q_t) is strictly stationary, ergodic and ρ-mixing, with ρ-mixing coefficients satisfying $\sum_{m-1}^{\infty} \rho_m^{1/2} < \infty$;

A2: $\tau \in \mathcal{T} = [c, 1-c]$ with $c \in (0,1/2)$. $\gamma(\tau)$ lies in a compact set $\mathcal{G} \subset \mathbb{R}$ for every $\tau \in \mathcal{T}$, and $\theta(\tau) \in \text{int}\ \Theta$, with Θ compact and convex;

A3: $E(||x_t||^{2+\epsilon}) < \infty$ with $\varepsilon > 0$, $||\cdot||$ the usual Euclidean norm, and $\max_t ||x_t|| = O(\sqrt{T})$;

A4: Let $F_t(\cdot \mid \Im_t) = F_t(\cdot)$ denote the conditional distribution function of y_t given $\Im_t$. $F_t(\cdot)$ has a continuous Lebesgue density, f_t, with $0 < f_t(u) < \infty$ on $\mathcal{U} = \{u : 0 < F_t(u) < 1\}$ and f_t is uniformly integrable on $\mathcal{U}$;

A5: For all $\tau \in \mathcal{T}$, $(\theta_0(\tau),\gamma_0(\tau)) = \arg\min_{(\theta,\gamma)} E[\rho_\tau(y_t - x_t(\gamma)'\theta)]$ exists and is unique.

A6:

$$\hat{\Omega}_0(\gamma, \gamma^*) = \frac{1}{T}\sum_{t=1}^{T} x_t(\gamma)x_t'(\gamma^*) \quad \text{and} \quad \hat{\Omega}_1(\tau, \gamma) = \frac{1}{T}\sum_{t=1}^{T} \hat{f}_{t-1}(F_{t-1}^{-1}(\tau))x_t(\gamma)x_t(\gamma)'$$

converge almost surely to $\Omega_0(\gamma, \gamma^*) = E[x_t(\gamma)x_t'(\gamma^*)]$ and $\Omega_1(\tau, \gamma) = E[f_{t-1}(F_{t-1}^{-1}(\tau))x_t(\gamma)x_t(\gamma)']$ respectively, uniformly over $\gamma(\tau) \in \mathcal{G}$ and all $\tau \in \mathcal{T}$.

These assumptions are common in the QR and regime switching literatures. A1 guarantees that the process is stationary and that the series satisfies Hansen's (2000) Ass 1.1. Depending on the specific type of model this assumption can be replaced by milder conditions on the values that the parameters in θ can take. For instance Koenker and Xiao (2006) allow for unit root in some quantiles under the condition that by integrating over the quantiles the process is stationary. A2 imposes that $\gamma(\cdot)$ lies on a compact set. This assumption was used by Hansen (1996) and Kato (2009). A3 and A4 are common in the QR literature. A5 guarantees that for each threshold parameter value and quantile the QR problem has a unique solution. A6 assumes the consistency of the estimators of the variance parameters $\Omega_0(\gamma,\gamma)$ and $\Omega_1(\tau,\gamma)$ uniformly over the parameter space $\mathcal{G} \times \mathcal{T}$.

Standard asymptotic theory for quantile regression shows that if γ_0 is known, then we have the following result;

Lemma 1. *Given assumptions A1–A6, and for γ_0 known with $\tau \in \mathcal{T}$ fixed,*

$$\sqrt{T}(\hat{\theta}_{\gamma_0}(\tau) - \theta_{\gamma_0}(\tau)) \xrightarrow{d} \text{N}(0, \Sigma(\tau, \gamma_0)),$$

where $\Sigma(\tau,\gamma_0) = \tau(1-\tau)\Omega_1(\tau,\gamma_0)^{-1}\Omega_0(\gamma_0,\gamma_0)\Omega_1(\tau,\gamma_0)^{-1}$.

Proof. The proof of this result uses standard QR asymptotic theory. The proof is a simple extension of that in Koenker and Xiao (2006) for the asymptotic normality in quantile autoregressive models. □

The most interesting case is, however, when the threshold value is not known and must be estimated. In the standard OLS framework, concerned with nonlinearities in the mean process, there are two possible scenarios determined by the continuity or discontinuity of the threshold model. Each method implies a different methodology and properties of estimators. Thus, for the continuous case the corresponding estimator $\hat{\gamma}$ is a $\sqrt{T}$ consistent estimator of γ, which converges in distribution, see Chan and Tsay (1998), to a normal random variable with mean zero and variance that depends on the correlation of $\hat{\gamma}$ with the vector of estimators $\hat{\theta}$. On the other hand, if the threshold model is discontinuous, see Chan (1993), the estimator $\hat{\gamma}$ converges at a faster rate (T consistent), and is therefore asymptotically independent of the vector of estimators of θ. In this case, two-stage estimation procedures can be used to consistently estimate the parameters of the model, see Hansen (1997), Hansen (2000) for details on the method. Caner (2002) extends the results in Chan (1993) and Hansen (2000) to the LAD estimation of a threshold model.

Similarly, we can study the consistency and asymptotic behaviour of the parameter estimators of $(\theta_0(\tau),\ \gamma_0(\tau))$ in the QR framework. The continuous case is a straightforward extension of Chan and Tsay (1998) using QR methods. The discontinuous case, defined now by the discontinuity of the quantile process on the threshold variable for certain τ, is, however, non-standard and deserves further attention. The following paragraphs derive the consistency and asymptotic distribution of the parameter estimators of the Q-TAR model in the discontinuous case. We need first to formalize this by imposing an assumption on discontinuity of the quantile process at $q(x_t) = \gamma(\cdot)$. This is shown in A7.

A7: For all $\tau \in \mathcal{T}$, let $\theta_1(\tau) = (\theta_{01}(\tau), \theta_{11}(\tau), \ldots, \theta_{p_y 1}(\tau))'$ and $\theta_2(\tau) = (\theta_{02}(\tau), \theta_{12}(\tau), \ldots, \theta_{p_y 2}(\tau))'$ be the autoregressive vectors for the two regimes. Then $(\theta_1(\tau) - \theta_2(\tau))' x_* \neq 0$, with x_* a realization of x_t and such that $q(x_*) = \gamma(\tau)$.[1]

The estimators of the autoregressive model and the threshold parameter are given by

$$(\hat{\theta}(\tau), \hat{\gamma}(\tau)) = \arg\min_{(\theta,\gamma)} \sum_{t=1}^{T} \rho_\tau\big(y_t - x_t(\gamma)'\theta\big). \tag{4}$$

As in the OLS context, this minimization problem can be carried out in two-stages. For fixed τ, consider a grid of γ values in the real line, and for each value estimate model (3) and save $\hat{\theta}_\gamma(\tau)$. Then, minimize

$$\hat{\gamma}(\tau) = \arg\min_{\gamma} \sum_{t=1}^{T} \rho_\tau(y_t - x_t(\gamma)'\hat{\theta}_\gamma(\tau)). \tag{5}$$

Lemma 2 shows the consistency of these estimators.

Lemma 2. *Given assumptions A1–A7; for fixed $\tau \in \mathcal{T}$,*

$$(\hat{\theta}(\tau), \hat{\gamma}(\tau)) = (\theta_0(\tau), \gamma_0(\tau)) + o_p(1).$$

Proof. The proof is given in the Appendix. □

Now we discuss the asymptotic normality of the parameter estimator $\hat{\theta}(\tau)$. To show this result for a fixed pair (τ,γ) with $\tau \in \mathcal{T}$ and $\gamma \in \mathcal{G}$, we first derive the weak convergence of the related process $\hat{\theta}_\gamma(\tau)$ to a bivariate Gaussian process indexed by τ and γ. The asymptotic normality of the estimator $\hat{\theta}(\tau)$, defined by fixing τ and γ in the process $\hat{\theta}_\gamma(\tau)$, follows from considering these parameters as constants in the Gaussian process representation of the asymptotic distribution of the latter process. In contrast to Hansen (1996) that uses the OLS score function for nonlinearities in the mean process, we use a different approach based on the Bahadur representation of the QR model. Thus, in Lemma 3 we first state the Bahadur representation as an intermediate result. It is important to note that weak convergence and Bahadur representation of a two parameter process have been studied already in the quantile regression literature, and similar results can be found in Su and Xiao (2008) and Qu (2008). Lemma 3 shows that $\hat{\theta}_\gamma(\tau)$ has a Bahadur representation uniformly in both τ and γ.

Lemma 3. *Suppose assumptions A1–A7 hold. Then*

$$\sup_{\tau\in\mathcal{T}} \sup_{\gamma\in\mathcal{G}} \|\sqrt{T}(\hat{\theta}_\gamma(\tau) - \theta_\gamma(\tau)) - \Omega_1^{-1}(\tau,\gamma) S_T(\tau,\gamma)\| = o_p(1), \tag{6}$$

where $S_T(\tau,\gamma) = \frac{1}{\sqrt{T}} \sum_{t=1}^{T} x_t(\gamma)\psi_\tau(y_t - F_{t-1}^{-1}(\tau))$ is the score function and $\psi_\tau(u) = \tau - I(u < 0)$ is the influence function of quantile regression models.

Proof. The proof is given in the Appendix. □

Theorem 1 shows the asymptotic distribution of the bivariate process $\hat{\theta}_\gamma(\tau)$. This result also lies the foundations for developing nonlinearity tests for the quantile process, see Galvao *et al.* (2010).

Theorem 1. *Given assumptions A1–A7,*

$$\sqrt{T}(\hat{\theta}_\gamma(\tau) - \theta_\gamma(\tau)) \Rightarrow B(\tau,\gamma), \tag{7}$$

with $B(\tau,\gamma)$ a bivariate Gaussian process with mean zero and covariance kernel defined by

$$K((\tau_i,\gamma_i),(\tau_j,\gamma_j)) = E(B(\tau_i,\gamma_i)B(\tau_j,\gamma_j)) = (\tau_i \wedge \tau_j - \tau_i\tau_j)\Omega_1(\tau_i,\gamma_i)^{-1}\Omega_0(\gamma_i,\gamma_j)\Omega_1(\tau_j,\gamma_j)^{-1},$$

with $\tau_i, \tau_j \in \mathcal{T}$ and $\gamma_i, \gamma_j \in \mathcal{G}$.

Proof. The proof is given in the Appendix. □

By fixing the values of τ and γ, it is immediate to derive the asymptotic distribution of the estimator $\hat{\theta}(\tau)$.

LEMMA 4. *Given assumptions A1–A7, for a fixed pair* $(\tau, \gamma) \in \mathcal{T} \times \mathcal{G}$,

$$\sqrt{T}(\hat{\theta}(\tau) - \theta(\tau)) \xrightarrow{d} N(0, \Sigma(\tau, \gamma)). \quad (8)$$

PROOF. The proof follows from Theorem 1 and noting that $K((\tau,\gamma),(\tau,\gamma)) = \Sigma(\tau,\gamma)$. □

It remains to study the asymptotic distribution of the threshold parameter estimator $\hat{\gamma}(\tau)$. This is an unresolved problem that depends on finding the rate of convergence of the estimator. Heuristically, and following arguments from standard threshold models for the mean process, see Chan (1993), it seems clear that the rate of convergence is T. The proof for the quantile process relies on empirical processes techniques for dependent data as in Arcones and Yu (1994), Doukhan *et al.* (1995), and Dehling *et al.* (2002). Bootstrap methods cannot be used either for approximating the finite-sample distribution of the standardized estimator of $\gamma(\tau)$ if the rate of convergence is not known. A feasible and unexplored solution is the use of subsampling methods, see Politis *et al.* (1999). The study of this methodology is, however, beyond the scope of this article. The consistency of the threshold estimator derived in Lemma 2 along with the rest of above results are sufficient for the subsequent Monte Carlo and empirical analysis.

3. MONTE CARLO EXPERIMENTS

We present Monte Carlo experiments to analyse some finite sample properties of the T-QAR method, in particular identification of the regimes and validity of the monotonicity assumption across quantiles. First, we study the identification of the proposed model compared with a simple linear SETAR model with two regimes sharing the same first two conditional statistical moments. Secondly, we evaluate the model by the monotonicity requirement on the conditional quantile functions.

3.1. Estimation

This experiment shows that QR estimation methods allowing for threshold nonlinearity perform better than standard OLS methods when the source of heterogeneity in the time series is not given by nonlinearities in the conditional mean process. To illustrate this, consider a baseline linear location-scale two-regime switching self-exciting threshold autoregressive (SETAR) process, as a particular T-QAR model:

$$y_t = \begin{cases} 0.5 + 0.5y_{t-1} - y_{t-1}u_t, & y_{t-1} \leq \gamma_0, \\ 0.5 + 0.5y_{t-1} + y_{t-1}u_t, & y_{t-1} > \gamma_0, \end{cases} \quad (9)$$

where $u_t \sim$ iid N(0,1) and $\gamma_0 = 0$.

Note that although this model is nonlinear, its first conditional moments are identical across regimes, that is,

$$E[y_t \,|\, y_{t-1} \leq 0] = E[y_t \,|\, y_{t-1} > 0] = 0.5 + 0.5y_{t-1} \quad (10)$$

and

$$\text{var}[y_t \,|\, y_{t-1} \leq 0] = \text{var}[y_t \,|\, y_{t-1} > 0] = y_{t-1}^2. \quad (11)$$

The heteroscedastic residual sequence $a_t = (1 - 2I(y_{t-1} > 0))y_{t-1}u_t$ from (9) is different across regimes but a_t^2 is identical for all t. However, the conditional τ-quantile process is

$$Q_{y_t}(\tau \,|\, y_{t-1} \leq 0) = 0.5 + (0.5 - F_u^{-1}(\tau))y_{t-1}, \quad (12)$$

and

$$Q_{y_t}(\tau \,|\, y_{t-1} > 0) = 0.5 + (0.5 + F_u^{-1}(\tau))y_{t-1}. \quad (13)$$

Unless $F_u^{-1}(\tau) = 0$, with F_u the distribution function of u, both regimes have different conditional quantiles and are clearly identified. If the distribution of u is symmetric $F_u^{-1}(1/2) = 0$, the conditional quantile processes coincide at the median, but they will differ in the upper and lower extreme quantiles.

We evaluate this model by conducting a Monte Carlo experiment with $T = 500$, where T is the sample size, and $R = 500$ where R is the number of repeated experiments. Figure 1 plots the root mean square error (RMSE) of estimating γ_0. Let $\hat{\theta}_{\gamma_0}(\tau)$ denote the QR estimate for fixed τ and γ_0, and $\bar{\theta}(\gamma_0)$ be the OLS estimate for fixed γ_0. To estimate γ, we consider a compact grid with the empirical percentiles of y_t from the 10th quantile to the 90th quantile. Moreover, we consider $\tau \in [0.10, 0.15, \ldots, 0.85, 0.90]$. The following estimators of γ_0 are considered:

$$\text{Solid line}: \hat{\gamma}(\tau) = \arg\min_{\gamma} \sum_{t=1}^{T} \rho_\tau(y_t - x_t(\gamma)'\hat{\theta}_\gamma(\tau))$$

$$\text{Dashed line}: \bar{\gamma}(\tau) = \arg\min_{\gamma} \sum_{t=1}^{T} (y_t - x_t(\gamma)'\hat{\theta}_\gamma(\tau))^2$$

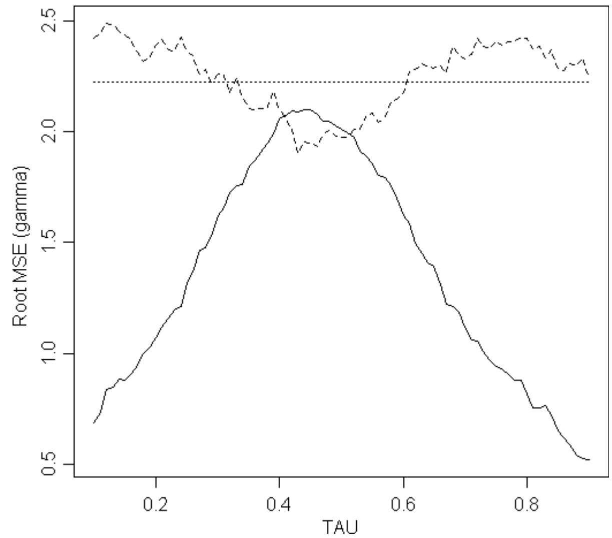

Figure 1. RMSE in estimating γ

$$\text{Dotted line}: \bar{\gamma} = \arg\min_{\gamma} \sum_{t=1}^{T} (y_t - x_t(\gamma)'\bar{\theta}(\gamma))^2$$

Here the regressor variable $x_t(\gamma)$ is defined as in Section 2. $\hat{\gamma}(\tau)$ is our preferred estimator, while $\bar{\gamma}$ is the least squares SETAR estimator for the threshold parameter and $\bar{\gamma}(\tau)$ is an hybrid estimator that runs QR for each pair (τ,γ) but minimizes the sum of square errors. The latter is used to show that it is the joint minimization of the QR problem what identifies[2] γ_0. As it can be seen from Figure 1, only $\hat{\gamma}(\tau)$ (solid line) shows small RMSE, which decreases as $|\tau - 0.5|$ increases. This increase is because model (9) becomes linear for the median, hence the difficulty of the estimation method for identifying γ_0. The plot also reveals that the least squares estimator, $\bar{\gamma}$ (dotted line) cannot identify the presence of two regimes. The reason for this is that OLS methods for estimation of the threshold parameter are based on minimization of the residual variance, which is the same for the simple linear model and the threshold nonlinear model. From (11), one can see that the second conditional moment is identical for the two regimes, such that the OLS estimator is not able to differentiate between the two regimes, and as a result the RMSE remains constant across quantiles. The hybrid method, $\bar{\gamma}(\tau)$ (dashed line), shows similar results. In particular, the method works slightly better than OLS to identify the regimes about the median and underperforms it for the other quantiles.

3.2. Quantile monotonicity

It is important to evaluate the model by the monotonicity requirement on the conditional quantile functions. To study the behaviour of the conditional quantiles, we evaluate the constructed predicted quantile functions from our proposed T-QAR with two regimes. We consider the quantile grid {0.05,0.1,0.15,0.20,…,0.9,0.95}, and evaluate $D_{\tau_1,\tau_2}(\Im_t) = \hat{Q}_{y_t}(\tau_2 \mid \Im_t) - \hat{Q}_{y_t}(\tau_1 \mid \Im_t)$ for consecutive quantiles $\tau_2 > \tau_1$. If monotonicity is satisfied D_{τ_1,τ_2} should be non-negative. Then, we compute the proportion of violations of the monotonicity assumption at the sample values of $\{y_t\}$, and take the average over 1000 Monte Carlo replications (Table 1). We consider two different DGPs. First, a linear model of the form $y_t = 0.5 + 0.5y_{t-1} + u_t$ (DGP$_1$); second a two regime model of the form $y_t = 0.5I(y_{t-1} \leq 0) - 0.5I(y_{t-1} > 0) + 0.5y_{t-1}I(y_{t-1} \leq 0) - 0.5y_{t-1}I(y_{t-1} > 0) + u_t$ (DGP$_2$), where $u_t \sim$ i.i.d. N(0,1). The simulation experiments show that the number of crossings is small. For the linear model, DGP$_1$, the proportion of violations is consistently below 5%, and there is a uniform behaviour across quantiles. In DGP$_2$ the proportion of crossings has an average 10%, with slightly higher values for low than for high quantiles.

Table 1. Monotonicity

	$D_{0.1,0.05}$	$D_{0.15,0.1}$	$D_{0.20,0.1}$	$D_{0.25,0.1}$	$D_{0.30,0.1}$	$D_{0.35,0.1}$	$D_{0.4,0.35}$	$D_{0.45,0.35}$	$D_{0.5,0.35}$
DGP$_1$	0.044	0.047	0.049	0.051	0.051	0.052	0.054	0.047	0.049
DGP$_2$	0.120	0.118	0.122	0.122	0.122	0.125	0.116	0.111	0.104
	$D_{0.55,0.35}$	$D_{0.6,0.35}$	$D_{0.65,0.35}$	$D_{0.7,0.35}$	$D_{0.75,0.35}$	$D_{0.8,0.35}$	$D_{0.85,0.35}$	$D_{0.9,0.35}$	$D_{0.95,0.35}$
DGP$_1$	0.049	0.049	0.050	0.048	0.048	0.037	0.040	0.043	0.041
DGP$_2$	0.096	0.087	0.086	0.083	0.076	0.076	0.078	0.077	0.075

Note: Monte Carlo experiments based on 1000 replications.

4. EMPIRICAL APPLICATION

Most macroeconomic series are affected by booms and busts in economic activity producing in turn expansionary periods followed by periods of economic contraction. It is well known in empirical macroeconometrics that these phenomena are reflected in an asymmetric behaviour in the dynamics of the corresponding time series. Empirical studies analysing these effects are, for example Beaudry and Koop (1993) that study the dynamics in US GDP of negative and positive shocks and uncover stronger effects in an economic activity from negative shocks than from positive shocks. Sarno *et al.* (2006) using regime switching processes observe asymmetries in the relation between interest rates and exchange rates between two countries that depend on the magnitude of the fluctuation of the exchange rate series. These authors observe that for small fluctuations these macroeconomic variables can depart from equilibrium conditions, however, as the fluctuations become more important exchange rates adjust so as to eliminate profit opportunities and restore equilibrium conditions in assets and exchange rate markets. A related study exploring asymmetries in the dynamics of exchange rates using TAR models is Coakley and Fuertes (2006). For series of unemployment Coakley *et al.* (2002) and Koenker and Xiao (2006) carry out thorough studies to measure the persistence in the dynamics on unemployment rates for different countries. Whereas the first authors focus on Europe, mainly UK and Germany, the second authors study unemployment rates in the US.

Our analysis extends the latter two studies by contemplating the possibility of nonlinearities in the quantile process and by analysing unemployment growth instead of unemployment rates. We postulate that the heterogeneity found in the unemployment growth series, that is growth in the number of people unemployed, can be due to asymmetric dynamics in the quantile process that, for some quantiles of the process, can potentially depend on different manners on previous values of the variable and lead us, therefore, to propose different TAR processes for different quantiles. The choice of this series, in contrast to the US unemployment rate series, is due to its clear stationary character. Many empirical studies in the unit root literature have investigated unemployment rate data. Nelson and Plosser (1982) in their seminal article studied the unit root property for annual unemployment rates and found evidence of stationarity. For the data period analysed in our study, however, evidence on stationarity of the unemployment rate series is mixed in contrast to unemployment growth series. More importantly, unemployment growth is an important variable from the policymaker perspective, as it clearly anticipates periods of social calm or distress. Figure 2 reports the monthly unemployment growth series from February 1948 to June 2007.

Autoregressive processes of order one are usually sufficient to describe the dynamics of macroeconomic stationary series as unemployment growth and unemployment rates. Furthermore, in nonlinear settings where the number of parameters increases with the number of regimes it is very convenient to choose parsimonious models that require a low number of parameters. For these reasons the first model under study in this application is the following T-QAR model with two regimes and piece-wise AR(1) processes:

$$Q_{y_t}(\tau \mid \Im_t) = \begin{cases} \theta_{01}(\tau) + \theta_{11}(\tau) y_{t-1}, & y_{t-1} \leq \gamma(\tau), \\ \theta_{02}(\tau) + \theta_{12}(\tau) y_{t-1}, & y_{t-1} > \gamma(\tau). \end{cases} \tag{14}$$

Table 2 reports the OLS estimates for this model and the results of standard tests for the linearity of the process. Hansen (1996) supremum and average Wald-test statistics cannot reject the linearity of the mean process at the 15% significance level suggesting a linear process for the conditional mean. In particular, the linear AR(1) version of process (14) yields an estimate of the autoregressive parameter of 0.116 with 0.037 standard deviation. The estimates from the QR method reported in Table 3 show, on the other hand, certain nonlinearities in the quantile process. The hypothesis tests for the linearity of the quantile process are developed in Galvao *et al.* (2008). As in the regime switching and structural break literatures two different types of tests are studied, supremum and average Wald tests. For nonstandard cases given by an unknown threshold parameter these tests are implemented using simulation and resampling techniques. We report *p*-values obtained from the *p*-value transformation method introduced in Hansen (1996) but adapted to quantile processes, see Galvao *et al.* (2008) for more details on the methods.

When the quantile increases toward unity we find statistical evidence of two regimes in the autoregressive specification above. A closer look to the upper quantiles shows a strong positive autocorrelation when unemployment growth is beyond the 1.969 threshold. The results for the median and lower quantiles are more puzzling. Thus, for the 0.05 and 0.10 quantiles it is the lower regime and not the upper regime the one to exhibit a positive autocorrelation with past values of unemployment growth. The threshold estimate in this case is very negative.

The large variability found in the estimates of the threshold value across quantiles can be due to the existence of more than two regimes in the data. To assess the adequacy of the model, we analyse graphically the number of crossings of the different quantile processes. Bassett and Koenker (1982) suggest that the presence of a significant amount of crossings between quantile curves can be a clear indication of a bad specification of the model. Figure 3 reports the quantile curves corresponding to Table 3 and Figure 4 plots the estimated conditional quantile at each time point. The figures suggest a correct specification of the model since there are no crossings among quantiles. Nevertheless, to analyse the fit of more involved models we also investigate the process with three regimes:

$$Q_{y_t}(\tau \mid \Im_t) = \begin{cases} \theta_{01}(\tau) + \theta_{11}(\tau) y_{t-1}, & y_{t-1} \leq \gamma_1(\tau), \\ \theta_{02}(\tau) + \theta_{12}(\tau) y_{t-1}, & \gamma_1(\tau) < y_{t-1} \leq \gamma_2(\tau), \\ \theta_{03}(\tau) + \theta_{13}(\tau) y_{t-1}, & y_{t-1} > \gamma_2(\tau), \end{cases} \tag{15}$$

with $\gamma_1(\tau) \leq \gamma_2(\tau)$ for each τ. The estimates from the OLS and QR methods are reported in Tables 4 and 5 respectively. Figure 5 illustrates these results.

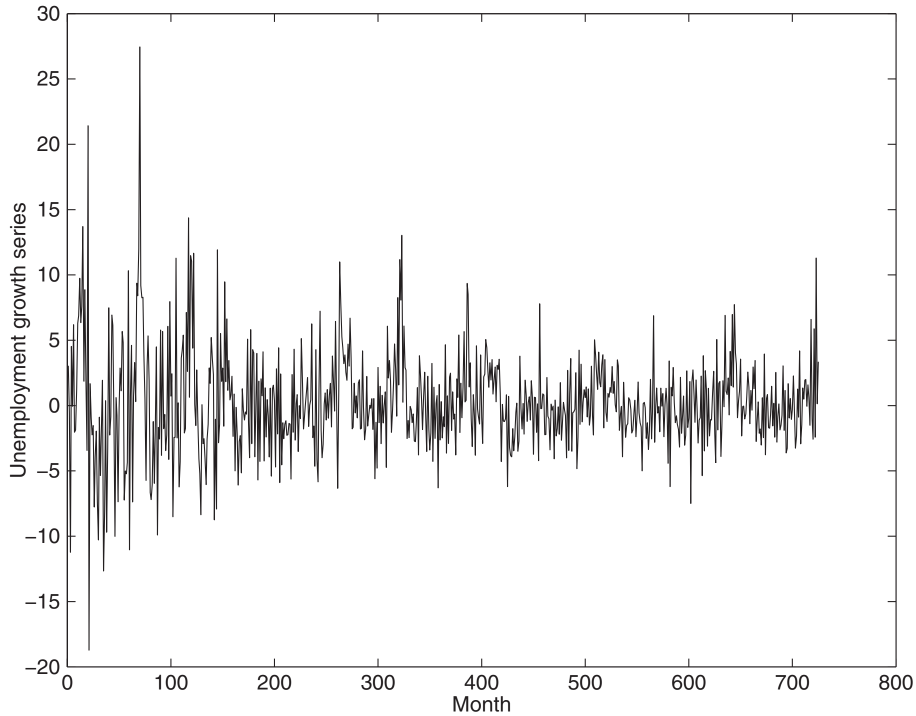

Figure 2. Unemployment growth series (in %). Monthly unemployment growth for US spanning the period February 1948 to June 2007. Source: freelunch.com

The *p*-value obtained from the supremum and average Wald tests is zero in both cases and for the entire quantile process (these *p*-values are not shown in the table). The counterpart tests using Hansen (1996) *p*-value methods also reject the null hypothesis of linearity of the mean process. These results show strong evidence of nonlinearity given by three regimes not only in the mean process but also on the entire quantile function. The analysis in Figures 6 and 7 reveals no crossings for $\tau = \{0.10, 0.25, 0.50, 0.75\}$ quantile curves. For the highest considered quantile, $\tau = 0.90$, the curve corresponding to the second regime crosses the rest of quantile curves and seems at odds with the rest of lines. This result can be indicative of a bad fit of the model for extreme upper quantiles. An explanation for this estimation problem is the absence of relevant information to estimate the model parameters corresponding to this quantile. The unreported curve for the 0.95 case is very similar to 0.90. On the other hand, for the lowest quantiles, $\tau = 0.05$ inclusive, the three regime process seems well suited to describe the nonlinearities in the data.

More specific comments on the results follow. The lower regime is common across quantiles and is defined by very extreme negative threshold values ranging between −5.395 and −2.906. The intercept and slope of (15) corresponding to this regime are not statistically significant indicating no memory on unemployment growth for this regime. The middle regime determined by an upper threshold value ranging between −0.859 and −0.473 shows, on the other hand, a negative autocorrelation in unemployment growth that increases in magnitude with the quantile. The autoregressive parameter corresponding to the middle regression equation is negative across quantiles. This fact along with the negative sign of the observations in this regime imply that periods of negative growth in unemployment are followed by periods of positive growth. This is consistent across quantiles. Finally, the estimates from the upper regime uncover a strongly statistically significant positive autocorrelation between unemployment growth levels that

Table 2. OLS estimates

Linear model						
Regime 1		Regime 2				
θ_{01}	θ_{11}	θ_{02}	θ_{12}			
0.229 (0.147)	0.116 (0.037)	0.229 (0.147)	0.116 (0.037)			
Threshold model						
Regime 1		Regime 2		Threshold	Wald-test (sup)	Wald-test (ave)
θ_{01}	θ_{11}	θ_{02}	θ_{12}	γ	*p*-value	*p*-value
0.723 (0.423)	0.119 (0.110)	−0.278 (0.235)	0.242 (0.113)	−1.348	0.267	0.148

Note: Standard deviations in parenthesis.

Table 3. QR estimates

	Regime 1		Regime 2		Threshold	Wald-test (sup)	Wald-test (ave)
τ	$\theta_{01}(\tau)$	$\theta_{11}(\tau)$	$\theta_{02}(\tau)$	$\theta_{12}(\tau)$	$\gamma(\tau)$	*p*-value	*p*-value
0.05	−0.564	0.607	−2.149	0.050	−2.043	0.089	0.320
	(0.891)	(0.113)	(0.554)	(0.156)			
0.10	−0.624	0.595	−3.942	0.025	−1.983	0.019	0.189
	(0.775)	(0.237)	(0.259)	(0.105)			
0.25	−1.327	−0.0097	−4.147	0.380	3.915	0.622	0.344
	(0.166)	(0.064)	(1.697)	(0.238)			
0.50	0.944	0.034	0.210	0.196	−2.897	0.041	0.006
	(0.963)	(0.158)	(0.179)	(0.067)			
0.75	1.928	0.019	−0.587	0.770	1.969	0.028	0.002
	(0.235)	(0.073)	(1.161)	(0.268)			
0.90	4.529	0.087	1.053	0.876	1.969	0.027	0.003
	(0.373)	(0.088)	(1.327)	(0.279)			
0.95	3.222	0.058	1.373	0.947	1.969	0.005	0.000
	(0.741)	(0.092)	(0.579)	(0.422)			
Kolmogorov–Smirnov						0.040	0.008

Note: Standard deviations in parenthesis.

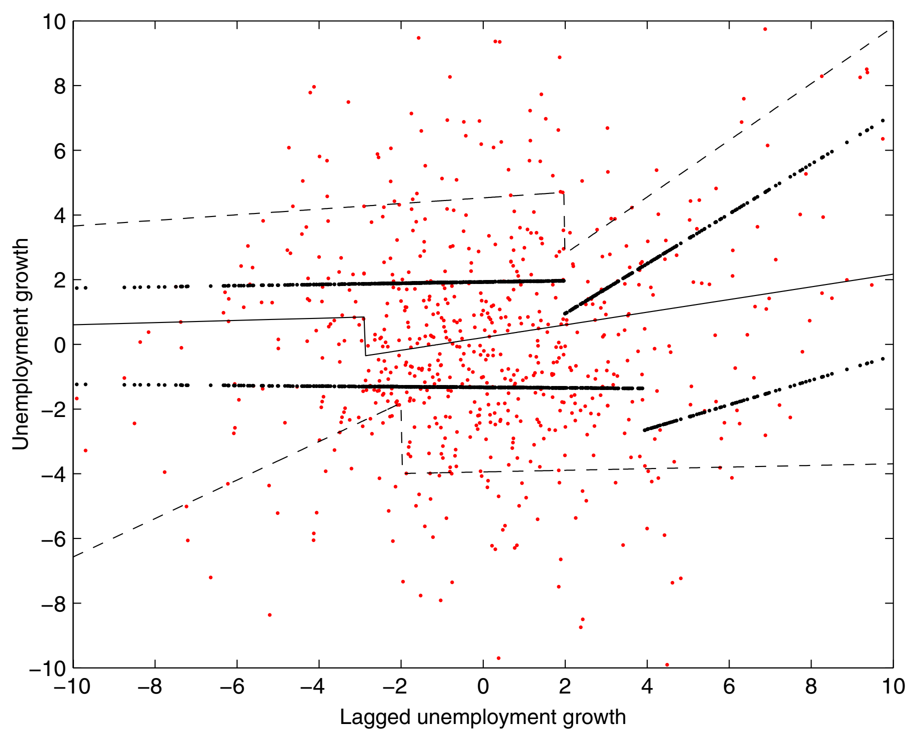

Figure 3. Conditional quantile curves for two regimes

shows evidence of strong persistence for most of the quantiles. It is worth noting that for the extreme 0.90 and 0.95 quantiles the regimes are given by thresholds 4.64% and 6.93% respectively. The standard error of the estimates is greater than for middle quantiles due to the absence of information to estimate the parameters in these quantiles. There is also evidence of persistence that occurs for extreme levels of unemployment growth.

The results of this analysis show nonlinearities in the dynamics of unemployment growth characterized by three regimes defined by two threshold values given by values ranging between [−5.395, −0.473] and [−2.906, −0.859], for the lower and upper threshold respectively. The characteristics of unemployment growth in each regime are completely different; whereas data in the lower regime shows no linear persistence, for the middle and upper regimes it shows signs of predictability. More specifically, the results suggest that once unemployment growth becomes close to zero or positive it starts to be positively correlated with previous levels of unemployment growth. The middle regime indicates a mean-reverting character of unemployment whereas the upper regime a persistent behaviour of the variable. From a statistical point of view, these results suggest that only the occurrence of large shocks of negative sign can make possible for the country to escape from ramping levels of unemployment once the upper threshold is reached. These shocks can take the form of some type of policy intervention or any other unexpected shock affecting positively the economy.

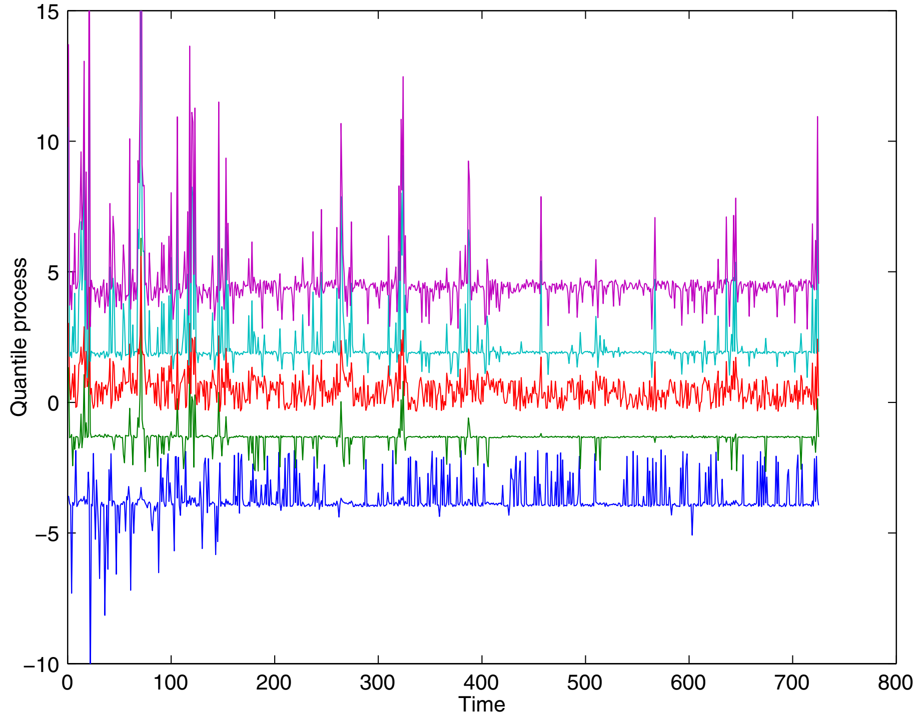

Figure 4. Conditional quantile time series for two regimes

Table 4. OLS estimates

Regime 1		Regime 2		Regime 3		Threshold		Wald-test (sup)	Wald-test (ave)
θ_{01}	θ_{11}	θ_{02}	θ_{12}	θ_{03}	θ_{13}	γ_1	γ_2	*p*-value	*p*-value
−0.514 (1.719)	0.019 (0.218)	−4.334 (0.164)	−0.574 (0.101)	1.811 (0.216)	0.302 (0.077)	−5.395	−0.859	0.00	0.00

Note: Standard deviations in parenthesis.

Table 5. QR estimates

	Regime 1		Regime 2		Regime 3		Threshold	
τ	$\theta_{01}(\tau)$	$\theta_{11}(\tau)$	$\theta_{02}(\tau)$	$\theta_{12}(\tau)$	$\theta_{03}(\tau)$	$\theta_{13}(\tau)$	$\gamma_1(\tau)$	$\gamma_2(\tau)$
0.05	−6.145 (6.302)	0.159 (0.622)	−7.176 (0.441)	−0.485 (0.072)	−0.165 (0.186)	0.007 (0.054)	−5.395	−0.473
0.10	−0.812 (4.999)	0.580 (0.571)	−5.681 (0.423)	−0.347 (0.198)	0.103 (0.222)	0.003 (0.065)	−5.395	−0.473
0.25	−1.402 (8.699)	0.194 (1.256)	−3.831 (0.207)	−0.239 (0.096)	0.598 (0.206)	0.092 (0.063)	−5.395	−0.473
0.50	−1.224 (1.582)	−0.155 (0.157)	−3.276 (0.168)	−0.347 (0.064)	1.030 (0.179)	0.297 (0.063)	−5.395	−0.859
0.75	2.442 (0.733)	0.088 (0.111)	−3.279 (0.170)	−0.727 (0.078)	2.111 (0.215)	0.582 (0.086)	−2.906	−0.859
0.90	4.139 (0.279)	−0.035 (0.100)	28.706 (7.355)	−4.236 (1.079)	−15.516 (18.493)	3.534 (1.910)	4.643	6.926
0.95	6.191 (0.494)	0.023 (0.172)	31.384 (6.792)	−4.387 (1.033)	−9.913 (17.440)	3.112 (1.723)	4.643	6.926

Note: Standard deviations in parenthesis.

5. CONCLUSIONS

This article provides useful insights on estimation procedures to identify threshold parameters in TAR models. Moreover, it shows that the threshold quantile autoregressive (T-QAR) framework has similar asymptotic properties to those found in Chan (1993) study for OLS based models. Simulation results show that the gains in RMSE are considerable when QR methods are used. These results are

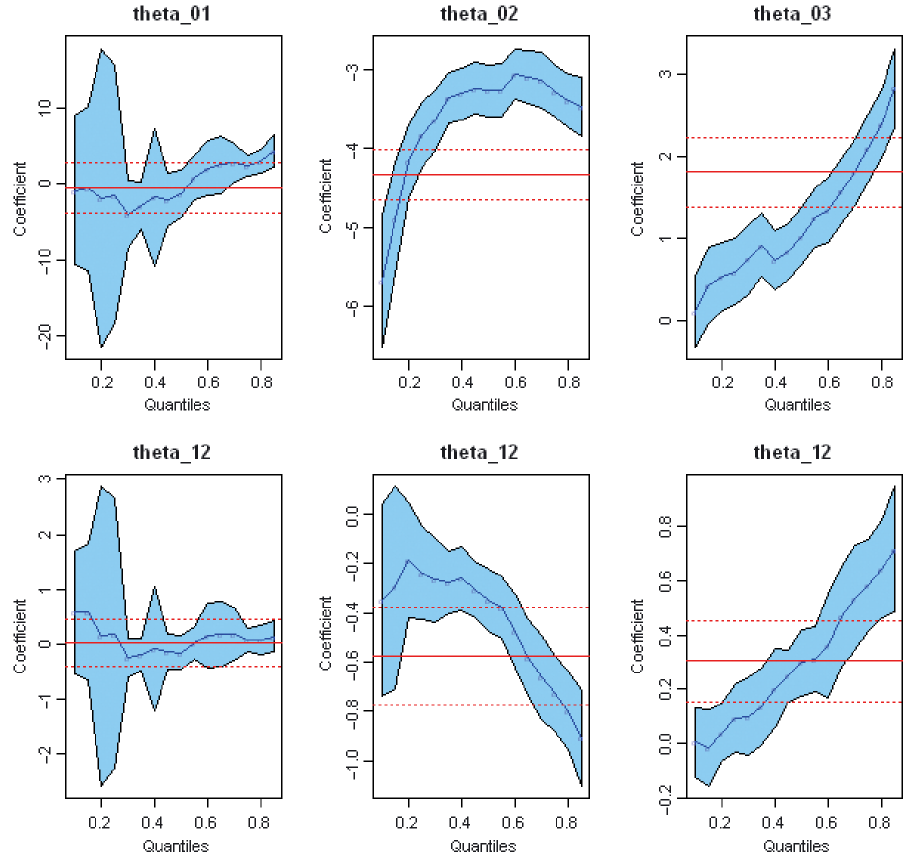

Figure 5. Parameter estimates: OLS and QR for three regimes

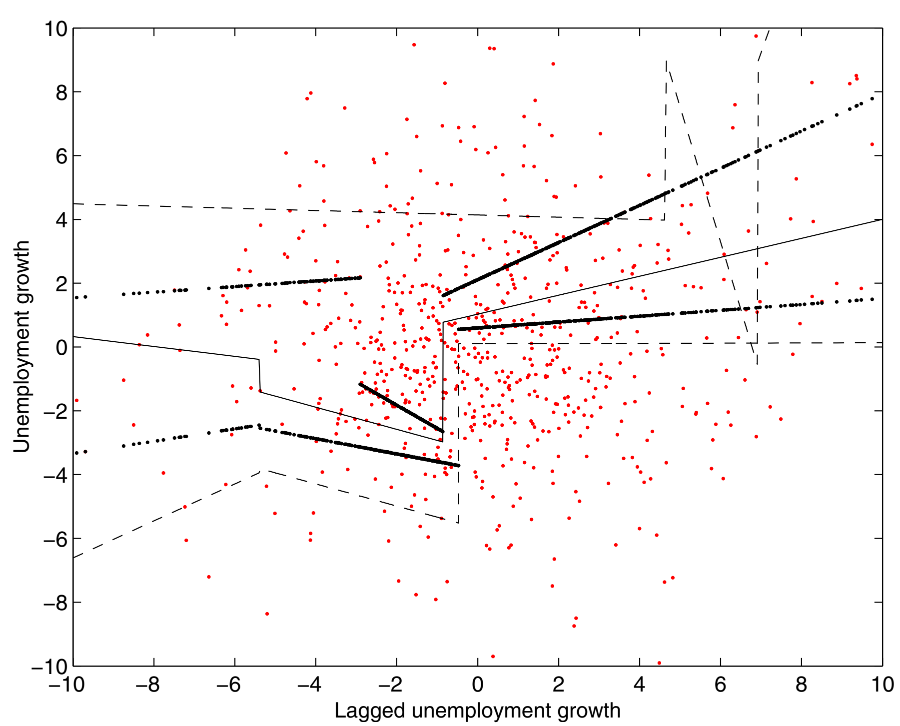

Figure 6. Conditional quantile curves for three regimes

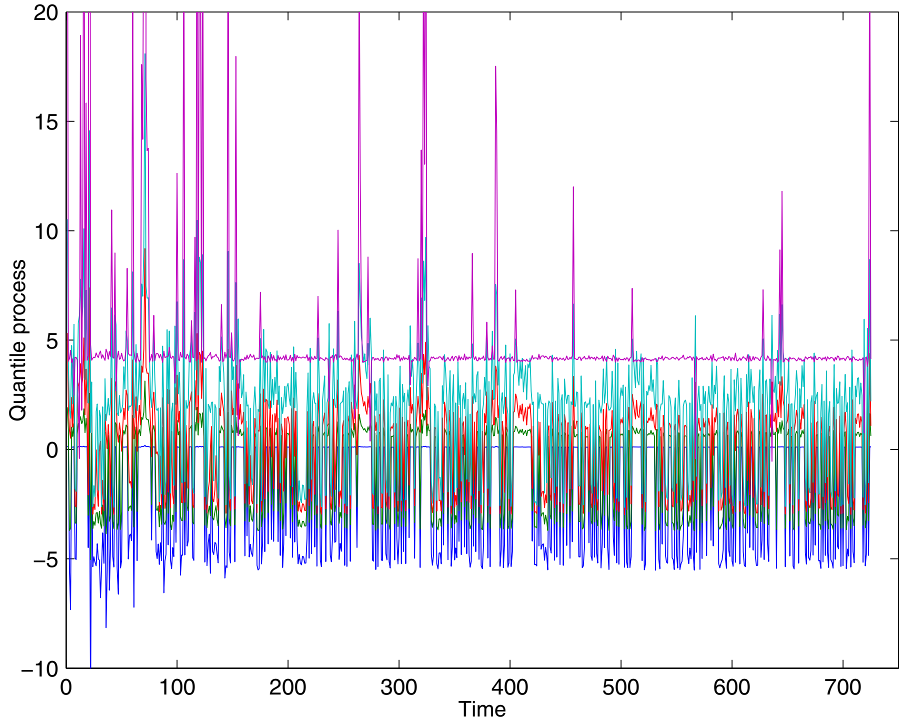

Figure 7. Conditional quantile time series for three regimes

exploited in an empirical application extending a related study on unemployment by Koenker and Xiao (2006) and Coakley *et al.* (2002). In our study, we find statistical evidence of the existence of three regimes in the conditional quantile process of unemployment growth. More specifically, our results suggest that strong contractions in the level of unemployment do not have an effect on future levels of unemployment. On the other hand, once unemployment growth becomes close to zero or positive it starts to be positively correlated with previous levels of unemployment growth producing strong persistence on unemployment growth. This is more relevant for extreme quantiles. From a policy perspective these results indicate that so as to revert dynamics given by an increasing unemployment growth there must be external shocks to the economy of opposite sign that bring unemployment levels below certain thresholds.

These models have several other potential applications left for future research but that highlight the usefulness of this methodology. For instance, the application of T-QAR models for modelling the dynamics of exchange rates. These time series are usually modelled by nonlinear processes that assume the presence of some fluctuation bands. Once these bands are exceeded non-arbitrage conditions preventing profit opportunities with zero cost make these time series to be mean-reverting. It can be interesting to analyse the quantile processes corresponding to these models, and see whether these bands are different across quantiles, and potential economic implications.

APPENDIX

The proofs of the results in this atricle use the notation for empirical processes as in Dehling *et al.* (2002). Define $\mathbb{E}_T = T^{-1}\sum_{i=1}^{T}\delta_{X_i}$, where δ_x assigns mass 1 at x and zero elsewhere, such that for any class $\Im$ of measurable function $f:\chi\to\mathbb{R}$. We use the following empirical process notation for $W=(y,x)$: $f\mapsto\mathbb{E}_T[f(W)] = \frac{1}{T}\sum_{t=1}^{T}f(W_t)$, where $Ef=\int_\chi f(x)P(dx)$. The first result of the article is to estimate the parameters of interest, $\mu=(\theta,\gamma)$, through quantile regression. This is the same as maximizing $M_T(\mu)=\mathbb{E}_T m_\mu$, where $m_\mu(x)\equiv-(\rho_\tau(y_t-x_t(\gamma)'\theta(\tau))-\rho_\tau(y_t-x_t(\gamma_0)'\theta_0(\tau)))$. First, we prove Lemma 2.

Proof. We need to establish the conditions for the argmax theorem, which will yield the consistency. First, observe that, for each $\tau\in\mathcal{T}$, $(\hat\theta(\tau),\hat\gamma(\tau))$ minimizes $Q_T(\theta,\gamma)\equiv\mathbb{E}_T(\rho_\tau(y_t-x_t(\gamma)'\theta(\tau))-\rho_\tau(y_t-x_t(\gamma_0)'\theta_0(\tau))$. Define $Q(\theta,\gamma)\equiv E(\rho_\tau(y_t-x_t(\gamma)'\theta(\tau))-\rho_\tau(y_t-x_t(\gamma_0)'\theta_0(\tau))$. By assumptions A4 and A5 $Q(\theta,\gamma)$ is uniquely minimized at $(\theta_0(\tau),\gamma_0(\tau))$ for each $\tau\in\mathcal{T}$.

Now, for simplicity, define α and β as the subsets of parameters in θ that correspond to the regimes $q_{t-1}\le\gamma$ and $q_{t-1}>\gamma$ respectively, where the dependence on τ is omitted to simplify notation. We will use the notation $x_t=(1,y_{t-1},y_{t-2},\ldots,y_{t-p_y})$, $\theta=(\alpha,\beta)$ and $\mu=(\theta,\gamma)$. For the particular SETAR(1) case, $q_{t-1}=y_{t-1}$.

Fix a compact set $B\subset\Theta\times\mathcal{G}$. We now verify that $\Im_B\equiv\{m_\mu:\mu\in B\}$ is Glivenko-Cantelli, where using assumption A7, we note that

$$m_\mu = -[\rho_\tau(\xi(\tau) - x_t'\alpha + x_t'\alpha_0) - \rho_\tau(\xi(\tau))]1(q_{t-1} < \gamma \wedge \gamma_0) - [\rho_\tau(\xi(\tau) - x_t'\beta + x_t'\alpha_0) - \rho_\tau(\xi(\tau))]1(\gamma < q_{t-1} \le \gamma_0)$$
$$- [\rho_\tau(\xi(\tau) - x_t'\alpha + x_t'\beta_0) - \rho_\tau(\xi(\tau))]1(\gamma_0 < q_{t-1} \le \gamma) - [\rho_\tau(\xi(\tau) - x_t'\beta + x_t'\beta_0) - \rho_\tau(\xi(\tau))]1(q_{t-1} > \gamma \vee \gamma_0),$$

where $\xi(\tau) = y_t - x_t(\gamma)'\theta_0(\tau)$.

In addition, for the first term, $\{\rho_\tau(\xi(\tau) - x_t'\alpha - x_t'\alpha_0) - \rho_\tau(\xi(\tau)) : \mu \in B\}$ and $1\{q_{t-1} < \gamma \wedge \gamma_0 : \mu \in B\}$ are separately Glivenko-Cantelli classes. For the first part, let $g(\tau, \mu) \equiv \rho_\tau(\xi(\tau) - x_t'\alpha + x_t'\alpha_0) - \rho_\tau(\xi(\tau))$. Define $g_T(\tau, \mu) \equiv \mathbb{E}g(\tau, \mu)$ and $g_\infty(\tau,\mu) \equiv Eg(\tau,\mu)$. By Knight's identity $\rho_\tau(u - v) - \rho_\tau(u) = -(\tau - I(u < 0))v + \int_0^v [I\{u \le s\} - I\{u \le 0\}]ds$, we have, by setting $u = y - x'\alpha_0$ and $v = x'(\alpha - \alpha_0)$, that

$$\rho_\tau(y_t - x_t'\alpha) - \rho_\tau(y_t - x_t'\alpha_0) = -(\tau - I(y < x'\alpha))x'(\alpha - \alpha_0) + \int_0^{x'(\alpha-\alpha_0)} [I\{y \le x'\alpha_0 + s\} - I\{y \le x'\alpha_0\}]ds.$$

Thus, it follows that for any $\tau \in (0,1)$ and $\alpha \in \mathbb{R}^p$, by A2 and A3 $|g_\infty(\tau,\mu)| \le 2\cdot E|x'(\alpha - \alpha_0)| \le 2\cdot\|x\|\cdot\|\alpha - \alpha_0\| < \infty$. We can also show that for any compact set B, $|g_T(\tau,\mu)| = |g_\infty(\tau,\mu)| + o_p(1)$ uniformly in $(\tau, \mu) \in \mathcal{T} \times B$. To do so, note that the statement is true pointwise by law of large numbers. The uniform convergence follows because $|g_T(\tau',\alpha') - g_T(\tau'',\alpha'')| \le C_1|\tau' - \tau''| + C_2\|\alpha' - \alpha''\|$, for some bounded constants C_1 and C_2. Hence the empirical process $(\tau,\mu) \mapsto g_T(\tau,\mu)$ is stochastically equicontinuous, which implies the uniform convergence.

The second part, $1\{q_{t-1} < \gamma \wedge \gamma_0 : \mu \in B\}$, being Glivenko-Cantelli follows directly from assumption A1–A4 and monotonicity of the indicator function such that it is a VC subgraph class as in Arcones and Yu (1994), and Dehling *et al.* (2002). Thus, from a simple estimation of covering numbers there are positive constants such that the entropy with bracketing number is finite. Since by A2–A4 the product of the components is integrable, the product of the two classes is also Glivenko-Cantelli[3]. Similar arguments reveal that the remaining terms of the sum are also Glivenko-Cantelli and the same theorem yields that $\Im_B$ is Glivenko-Cantelli.

Finally, existency of a solution $(\hat\theta, \hat\gamma)$ can be shown by using similar arguments as in Kosorok (2008, Ch.14), hence, by an argmax theorem, as Lemma B1 in Chernozhukov and Hansen (2006), $(\hat\theta(\tau), \hat\gamma(\tau)) \xrightarrow{p} (\theta_0(\tau), \gamma_0(\tau))$. □

Now we show Lemma 3. Under assumptions A1–A7, we want to show that the estimator $\hat\theta_\gamma(\tau)$ has a Bahadur representation uniformly in both τ and γ. Define $\Omega_1(\tau, \gamma) = T^{-1}\sum_{n=1}^T f_t(x_t(\gamma)'\theta_0)x_t(\gamma)x_t(\gamma)'$ and $S_T(\tau, \gamma) = T^{-1/2}\sum_{n=1}^T \psi_\tau(y_t - x_t(\gamma)'\theta_0)x_t(\gamma)$, where $\psi_\tau(u) = \tau - I(u < 0)$ is the influence function of the quantile regression model. As in Lemma 2, we define α and β as the subsets of parameters in θ that correspond to the regimes $q_{t-1} \le \gamma$ and $q_{t-1} > \gamma$ respectively, such that $\theta_{0\tau} = (\alpha_{0\tau}', \beta_{0\tau}')'$ signifies the true parameter vector. In addition, define the following weighted quantities,

$$\hat\Delta_{\tau\gamma} = \begin{pmatrix} \sqrt{T}(\hat\alpha_{\tau\gamma} - \alpha_{0\tau}) \\ \sqrt{T}(\hat\beta_{\tau\gamma} - \beta_{0\tau}), \end{pmatrix}, \quad \Delta_\tau = \begin{pmatrix} \sqrt{T}(\alpha_{\tau\gamma} - \alpha_{0\tau}) \\ \sqrt{T}(\beta_{\tau\gamma} - \beta_{0\tau}), \end{pmatrix}.$$

Let $y_t^* - x_t(\gamma)'\theta_{0\tau}$ and $y_{t,\gamma}^*(\Delta_\tau) = y_t^* - x_t(\gamma)'\Delta_\tau/\sqrt{T} = y_t - x_t(\gamma)'\theta(\tau)$. It follows from (4) that $\hat\Delta_{\tau\gamma} = \min_{\Delta_\tau}\sum_{t=1}^T \rho_\tau(y_{t,\gamma}^*(\Delta_\tau))$. Set

$$V_n(\tau, \gamma, \Delta_\tau) = T^{-1/2}\sum_{t=1}^T \psi_\tau(y_{t,\gamma}^*(\Delta_\tau))x_t(\gamma) = T^{-1/2}\sum_{t=1}^T \psi_\tau(y_t - x_t(\gamma)'(\theta_{0\tau} + T^{-1/2}\Delta_\tau))x_t(\gamma),$$
$$\bar V_n(\tau, \gamma, \Delta_\tau) = T^{-1/2}\sum_{t=1}^T E[\psi_\tau(y_t - x_t(\gamma)'(\theta_{0\tau} + T^{-1/2}\Delta_\tau))x_t(\gamma)].$$

Noting that $-\Delta_\tau' V_n(\tau, \gamma, \lambda\Delta_\tau)$ is an increasing function of $\lambda \ge 1$ and using A6 for the uniform consistency of $\hat\Omega_1(\tau, \gamma)$, Lemma 3 then follows from an application of Lem A.4 of Koenker and Zhao (1996). This is proven by noting that from assumptions A1–A7 and similar arguments as in Lem 2 and 3 of Su and Xiao (2008), $\sup_{\tau\in\mathcal{T}}\sup_{\gamma\in\mathcal{G}}\sup_{\|\Delta\|\le M}\|\bar V_n(\tau, \gamma, \Delta) - \bar V_n(\tau, \gamma, 0) + \Omega_1(\tau, \gamma)\Delta\| = o_p(1)$ and $\sup_{\tau\in\mathcal{T}}\sup_{\gamma\in\mathcal{G}}\|V_n(\tau, \gamma, \hat\Delta_{\tau\gamma})\| = o_p(1)$, and from the following auxiliary Lemma.

Lemma A1. *Suppose Assumptions A1–A7 hold. Then,*

$$\sup_{\tau\in\mathcal{T}}\sup_{\gamma\in\mathcal{G}}\sup_{\|\Delta\|\le M}\|V_n(\tau, \gamma, \Delta) - V_n(\tau, \gamma, 0) - [\bar V_n(\tau, \gamma, \Delta) - \bar V_n(\tau, \gamma, 0)]\| = o_p(1).$$

Proof. The detailed proof is omitted but available from the authors upon request. □

The next auxiliary lemma shows Donskerness of the $S_T(\tau,\gamma)$ process.

Lemma A2. *Suppose assumptions A1–A7 hold. Then, the process $S_T(\tau,\gamma)$ is Donsker.*

PROOF. Let

$$S_T(\tau,\gamma,\theta(\tau)) = T^{-1/2}\sum_{t=1}^{T} x_{t\gamma}[1(y_t \le x_{t\gamma}'\theta(\tau)) - \tau] = T^{-1/2}\sum_{t=1}^{T} x_t(\gamma)[1(F_t(y_t) \le \tau) - \tau],$$

where $F_t(\cdot)$ is the conditional distribution function of y_t, and the last equality follows because Assumption A4 implies $F(\cdot)$ is absolute continuous and strictly increasing almost everywhere. Define $u_t = F_t(y_t)$, then u_t has a standard uniform distribution. Hence,

$$\begin{aligned} S_T(\tau,\gamma,\theta(\tau)) &= T^{-1/2}\sum_{t=1}^{T} x_t(\gamma)[1(u_t \le \tau) - \tau] \\ &= T^{-1/2}\sum_{t=1}^{T}(x_t 1(y_{t-1} \le \gamma), \quad x_t 1(y_{t-1} > \gamma))[1(u_t \le \tau) - \tau]. \end{aligned}$$

Note that the functional class $\mathcal{V} = \{1\{y_{t-1} \le \gamma\}, \gamma \in \mathcal{G}\}$ is a VC subgraph class and hence also Donsker class, with envelope 2 [see e.g. Arcones and Yu (1994) and Dehling *et al.* (2002)]. By a similar arument $\mathcal{H} = \{1\{u_t \le \tau\}, \tau \in \mathcal{T}\}$ is Donsker. Note that $\mathcal{J} = \{\tau \mapsto \tau\}$ also belongs to a bounded Donsker class. Hence the functional class $\mathcal{H} - \mathcal{J}$ is also Donsker with envelope equal 2. The product of $\mathcal{V}$ with x also forms a Donsker class with a square integrable envelope $2 \cdot \max_{j \in 1,\ldots,d}|x_j|$. Hence, since the envelope is square integrable, $S_T(\tau,\gamma)$ is Donsker. □

Now we prove Theorem 1 and Lemma 4.

PROOF. Let us fix $\gamma \in \mathcal{G}$, for certain $\tau \in \mathcal{T}$ given. Using the above notation this implies that $\hat\theta(\tau) = \hat\theta_\gamma(\tau)$ and $\theta(\tau) = \theta_\gamma(\tau)$. Now, from the Bahadur representation of the QR model

$$\sqrt{T}\left(\hat\theta(\tau) - \theta(\tau)\right) = \hat\Omega_1^{-1}(\tau,\gamma)S_T(\tau,\gamma) + o_p(1),$$

where

$$S_T(\tau,\gamma) = \frac{1}{\sqrt{T}}\sum_{t=1}^{T} x_t(\gamma)\psi_\tau(y_t - F_t^{-1}(\tau))$$

is the score function and $\psi_\tau(u) = \tau - I(u < 0)$ is the influence function in the quantile regression models. By the law of iterated expectations $E[x_t(\gamma)\psi_\tau(y_t - F_t^{-1}(\tau)) \mid \Im_t] = 0$. Now, by the central limit theorem, Slutsky's theorem, and A1–A7,

$$\frac{1}{\sqrt{T}}\hat\Omega_1^{-1}(\tau,\gamma(\tau))\sum_{t=1}^{T} x_t(\gamma)\psi_\tau(y_t - F_t^{-1}(\tau)) \Rightarrow \mathrm{N}(0,\Sigma(\tau,\gamma)), \tag{16}$$

with $\Sigma(\tau,\gamma) = \tau(1-\tau)\Omega_1(\tau,\gamma)^{-1}\Omega_0(\gamma,\gamma)\Omega_1(\tau,\gamma)^{-1}$. This proves Lemma 4. We can extend this result to the corresponding functional process indexed by τ and γ, with τ and γ dense in $\mathcal{T}$ and $\mathcal{G}$ respectively. By Lemma A2, this is possible given that the class of functions we are interested belongs to the Donsker class. Then, this process converges in distribution in the Skorohod space $D(\mathcal{T},\mathcal{G})$, equipped with the uniform norm, to a bivariate Gaussian process with zero mean and covariance kernel

$$K((\tau_i,\gamma_i),(\tau_j,\gamma_j)) = (\tau_i \wedge \tau_j - \tau_i\tau_j)\Omega_1(\tau_i,\gamma_i)^{-1}\Omega_0(\gamma_i,\gamma_j)\Omega_1(\tau_j,\gamma_j)^{-1},$$

for every $i, j = 1,\ldots,n$ with $\tau_i,\tau_j \in \mathcal{T}$ and $\gamma_i,\gamma_j \in \mathcal{G}$. Finally, we obtain

$$\sqrt{T}(\hat\theta_\gamma(\tau) - \theta_\gamma(\tau)) = \Omega_1^{-1}(\tau,\gamma)S_T(\tau,\gamma) + o_p(1),$$

that converges in distribution to the mean zero bivariate Gaussian process $B(\tau,\gamma)$ with covariance kernel $K((\tau_i,\gamma_i),(\tau_j,\gamma_j))$. □

Acknowledgements

The authors express their appreciation to Anil Bera, Jesus Gonzalo, Bruce Hansen, Kengo Kato, Roger Koenker, Zhongjun Qu, and participants in the seminars at Cass Business School, City University London, University of Wisconsin-Milwaukee, Royal Economic Society 2009 Conference, 2009 North American Summer Meeting of the Econometric Society, 2009 Far East and South Asia Meeting of the Econometric Society and 2009 European Meeting of the Econometric Society for helpful comments and discussions. The authors also thank an anonymous referee for useful suggestions and comments that led to a considerably improved version of the article. Any remaining errors are the authors' own.

NOTES

1. For the quantile τ of the Q-TAR process with $q(x_t) = y_{t-1}$ this condition is equal to $\theta_{01}(\tau) + \theta_{11}(\tau)\gamma(\tau) \neq \theta_{02}(\tau) + \theta_{12}(\tau)\gamma(\tau)$.
2. For sake of space and interest we omit the proof of consistency of this estimator.

3. See e.g. cor 9.26 in Kosorok (2008). It is important to note that the 'preservation of products' result does not hold in general for Donsker classes, but it holds for integrable Glivenko-Cantelli classes.

REFERENCES

Arcones, M. A. and Yu, B. (1994) Central limit theorems for empirical and U-processes of stationary mixing sequences. *Journal of Theoretical Probability* **7**, 47–71.
Bassett, G. W. and Koenker, R. (1982) An empirical quantile function for linear models with iid errors. *Journal of the American Statistical Association* **77**, 407–15.
Beaudry, P. and Koop, G. (1993) Do recessions permanently change output? *Journal of Monetary Economics* **31**, 149–63.
Cai, Y. (2010) Forecasting for quantile self exciting threshold autoregressive time series models. *Biometrika* **97**, 199–208.
Cai, Y. and Stander, J. (2008) Quantile-self exciting threshold time series models. *Journal of Time Series Analysis* **29**, 187–202.
Caner, M. (2002) A note on least absolute deviation estimation of a threshold model. *Econometric Theory* **18**, 800–14.
Chan, K. S. (1993) Consistency and limiting distribution of the least squares estimator of a threshold autoregressive model. *Annals of Statistics* **21**, 520–33.
Chan, K. S. and Tong, H. (1986) On estimating thresholds in autoregressive models. *Journal of Time Series Analysis* **7**, 179–94.
Chan, K. S. and Tsay, R. (1998) Limiting properties of the conditional least squares estimator of a continuous TAR model. *Biometrika* **85**, 413–26.
Chernozhukov, V., Fernandez-Val, I. and Galichon, A. (2010) Quantile and probability curves without crossing. *Econometrica* **78**, 1093–125.
Chernozhukov, V. and Hansen, C. (2006) Instrumental quantile regression inference for structural and treatment effects models. *Journal of Econometrics* **132**, 491–525.
Clarida, R., Sarno, L., Taylor, M. and Valente, G. (2006) The role of asymmetries and regime shifts in the term structure of interest rates. *Journal of Business* **79**, 1193–224.
Coakley, J. and Fuertes, A.-M. (2006) Testing for sign and amplitude asymmetries using threshold autoregressions. *Journal of Economic Dynamics and Control* **30**, 623–54.
Coakley, J., Fuertes, A.-M. and Zoega, G. (2002) Evaluating the persistence and structuralist theories of unemployment. *Studies in Nonlinear Dynamics and Econometrics* **5**, 1–22.
Dehling, H., Mikosch, T. and Sorensen, M. (2002) *Empirical Process Techniques for Dependent Data*. Boston: Birkhauser Press.
Dette, H. and Volgushev, S. (2008) Non-crossing nonparametric estimates of quantile curves. *Journal of the Royal Statistical Society, Series* B **70**, 609–27.
Doukhan, P., Massart, P. and Rio, E. (1995) Invariance principles for absolutely regular empirical processes. *Annales de l Institut Henri Poincaré* **31**, 393–427.
Galvao, A. F., Kato, K., Montes-Rojas, G. and Olmo, J. (2010) A uniform test for linearity in quantile regression, *Mimeo*.
Galvao, A. F., Montes-Rojas, G. and Olmo, J. (2008) *Threshold quantile autoregressive models*. 09/05 Discussion Paper Series Economics Department, City University London.
Glosten, L. R., Jagannatthan, R. and Runkle, D. (1993) On the relation between the expected value and the volatility of the nominal excess returns on stocks. *Journal of Finance* **48**, 1779–801.
Hansen, B. E. (1996) Inference when a nuisance parameter is not identified under the null hypothesis. *Econometrica* **64**, 413–30.
Hansen, B. E. (1997) Inference in TAR models. *Studies in Nonlinear Dynamics and Econometrics* **2**, 1–14.
Hansen, B. E. (2000) Sample splitting and threshold estimation. *Econometrica* **68**, 575–603.
He, X. (1997) Quantile curves without crossing. *The American Statistician* **51**, 186–92.
Kato, K. (2009) Asymptotics for argmin processes: convexity arguments. *Journal of Multivariate Analysis* **100**, 1816–29.
Koenker, R. and Bassett, G. W. (1978) Regression quantiles. *Econometrica* **46**, 33–49.
Koenker, R. and Ng, P. (2005) Inequality constrained quantile regression. *Sankhya* **67**, 418–40.
Koenker, R. and Xiao, Z. (2006) Quantile autoregression. *Journal of the American Statistical Association* **101**, 980–90.
Koenker, R. and Zhao, Q. (1996) Conditional quantile estimation and inference for ARCH models. *Econometric Theory* **12**, 793–813.
Kosorok, M. R. (2008) *Introduction to Empirical Processes and Semiparametric Inference*. New York: Springer-Verlag Press.
Mammen, E. (1991) Nonparametric regression under qualitative smoothness assumptions. *Annals of Statistics* **19**, 741–59.
Nelson, D. (1991) Conditional heteroskedasticity in asset returns: a new approach. *Econometrica* **59**, 347–70.
Nelson, D. and Plosser, C. (1982) Trends and random walks in macroeconomic time series: some evidence and implications. *Journal of Monetary Economics* **10**, 139–62.
Politis, D. N., Romano, J. P. and Wolf, M. (1999) *Subsampling*. New York: Springer-Verlag Press.
Qu, Z. (2008) Testing for structural change in regression quantiles. *Journal of Econometrics* **146**, 170–84.
Sarno, L., Valente, G. and Leon, H. (2006) Nonlinearity in deviations from uncovered interest parity: an explanation of the forward bias puzzle. *Review of Finance* **10**, 443–82.
Su, L. and Xiao, Z. (2008) Testing for parameter stability in quantile regression models. *Statistics and Probability Letters*, **78**, 2768–75.
Terasvirta, T. (1994) Specification, estimation, and evaluation of smooth transition autoregressive models. *Journal of the American Statistical Association* **89**, 208–18.
Tong, H. (1978) On a threshold model. In *Pattern Recognition and Signal Processing, NATO ASI Series E: Applied Sc. No. 29,* (ed. C. H. Chen). The Netherlands: Sijthoff & Noordhoff, pp. 575–86.
Tong, H. (1983) *Threshold Models in Non-linear Time Series Analysis.* Lecture Notes in Statistics, No.21, New York: Springer-Verlag.
Tong, H. (1990) *Non-Linear Time Series: A Dynamical System Approach*. Oxford: Oxford University Press.
Tong, H. and Lim, K. (1980) Threshold autoregression, limit cycles and cyclical data. *Journal of Royal Statistical Society* **4**, 245–92.
Tsay, R. (1989) Testing and modeling threshold autoregressive processes. *Journal of the American Statistical Association* **18**, 931–55.
Zakoian, J. (1994) Threshold heteroscedastic models. *Journal of Economic Dynamics and Control* **18**, 931–55.