Quantile autoregression

Authors: Roger Koenker, Zhijie Xiao

Post-print version of an article published in Journal of the American Statistical Association 101(475): 980-990. doi:10.1198/016214506000000672.

Post-print version of an article published in Journal of the American Statistical Association 101(475): 980-990. doi:10.1198/016214506000000672.

QUANTILE AUTOREGRESSION

ROGER KOENKER AND ZHIJIE XIAO

Abstract. We consider quantile autoregression (QAR) models in which the autoregressive coefficients can be expressed as monotone functions of a single, scalar random variable. The models can capture systematic influences of conditioning variables on the location, scale and shape of the conditional distribution of the response, and therefore constitute a significant extension of classical constant coefficient linear time series models in which the effect of conditioning is confined to a location shift. The models may be interpreted as a special case of the general random coefficient autoregression model with strongly dependent coefficients. Statistical properties of the proposed model and associated estimators are studied. The limiting distributions of the autoregression quantile process are derived. Quantile autoregression inference methods are also investigated. Empirical applications of the model to the U.S. unemployment rate and U.S. gasoline prices highlight the potential of the model.

1. Introduction

Constant coefficient linear time series models have played an enormously successful role in statistics, and gradually various forms of random coefficient time series models have also emerged as viable competitors in particular fields of application. One variant of the latter class of models, although perhaps not immediately recognizable as such, is the linear quantile regression model. This model has received considerable attention in the theoretical literature, and can be easily estimated with the quantile regression methods proposed in Koenker and Bassett (1978). Curiously, however, all of the theoretical work dealing with this model (that we are aware of) focuses exclusively on the iid innovation case that restricts the autoregressive coefficients to be independent of the specified quantiles. In this paper we seek to relax this restriction and consider

Corresponding author: Roger Koenker, Department of Economics, University of Illinois, Champaign, Il, 61820. Email: rkoenker@uiuc.edu.
Version July 8, 2005. This research was partially supported by NSF grant SES-02-40781. The authors would like to thank the Co-Editor, Associate Editor, two referees, and Steve Portnoy and Peter Phillips for valuable comments and discussions regarding this work.

Post-print version of an article published in Journal of the American Statistical Association 101(475): 980-990. doi:10.1198/016214506000000672.

linear quantile autoregression models whose autoregressive (slope) parameters may vary with quantiles $\tau \in [0, 1]$. We hope that these models might expand the modeling options for time series that display asymmetric dynamics or local persistency.

Considerable recent research effort has been devoted to modifications of traditional constant coefficient dynamic models to incorporate a variety of heterogeneous innovation effects. An important motivation for such modifications is the introduction of asymmetries into model dynamics. It is widely acknowledged that many important economic variables may display asymmetric adjustment paths (e.g. Neftci (1984), Enders and Granger (1998)). The observation that firms are more apt to increase than to reduce prices is a key feature of many macroeconomic models. Beaudry and Koop (1993) have argued that positive shocks to U.S. GDP are more persistent than negative shocks, indicating asymmetric business cycle dynamics over different quantiles of the innovation process. In addition, while it is generally recognized that output fluctuations are persistent, less persistent results are also found at longer horizons (Beaudry and Koop (1993)), suggesting some form of "local persistency." See, *inter alia,* Delong and Summers (1986), Hamilton (1989), Evans and Wachtel (1993), Bradley and Jansen (1997), Hess and Iwata (1997), and Kuan and Huang (2001). A related development is the growing literature on threshold autoregression (TAR) see e.g. Balke and Fomby (1997); Tsay (1997); Gonzalez and Gonzalo (1998); Hansen (2000); and Caner and Hansen (2001).

We believe that quantile regression methods can provide an alternative way to study asymmetric dynamics and local persistency in time series. We propose a new quantile autoregression (QAR) model in which autoregressive coefficients may take distinct values over different quantiles of the innovation process. We show that some forms of the model can exhibit unit-root-like tendencies or even temporarily explosive behavior, but occasional episodes of mean reversion are sufficient to insure stationarity. The models lead to interesting new hypotheses and inference apparatus for time series.

The paper is organized as follows: We introduce the model and study some basic statistical properties of the QAR process in Section 2. Section 3 develops the limiting distribution of the QAR estimator. Section 4 considers some restrictions imposed on the model by the monotonicity requirement on the conditional quantile functions. Statistical inference, including testing for asymmetric dynamics, is explored in Section

Post-print version of an article published in Journal of the American Statistical Association 101(475): 980-990. doi:10.1198/016214506000000672.

5. Section 6 reports a Monte Carlo experiment on the sampling performance of the proposed inference procedure. An empirical application to U.S. unemployment rate time series is given in Section 7. Proofs appear in the Appendix.

2. The Model

There is a substantial theoretical literature, including Weiss (1987), Knight (1989), Koul and Saleh(1995), Koul and Mukherjee(1994), Hercé (1996), Hasan and Koenker (1997), Jurečková and Hallin (1999) dealing with the linear quantile autoregression model. In this model the τ-th conditional quantile function of the response y_t is expressed as a linear function of lagged values of the response. The current paper wish to study estimation and inference in a more general class of quantile autoregressive (QAR) models in which all of the autoregressive coefficients are allowed to be τ-dependent, and therefore are capable of altering the location, scale and shape of the conditional densities.

2.1. **The Model.**

Let $\{U_t\}$ be a sequence of *iid* standard uniform random variables, and consider the pth order autoregressive process,

$$y_t = \theta_0(U_t) + \theta_1(U_t)y_{t-1} + \cdots + \theta_p(U_t)y_{t-p}, \tag{1}$$

where the θ_j's are unknown functions $[0,1] \to \mathbb{R}$ that we will want to estimate. Provided that the right hand side of (1) is monotone increasing in U_t, it follows that the τth conditional quantile function of y_t can be written as,

$$Q_{y_t}(\tau|y_{t-1}, ..., y_{t-p}) = \theta_0(\tau) + \theta_1(\tau)y_{t-1} + ... + \theta_p(\tau)y_{t-p}, \tag{2}$$

or somewhat more compactly as,

$$Q_{y_t}(\tau|\mathcal{F}_{t-1}) = x_t^\top \theta(\tau). \tag{3}$$

where $x_t = (1, y_{t-1}, ..., y_{t-p})^\top$, and $\mathcal{F}_t$ is the σ-field generated by $\{y_s, s \leq t\}$. The transition from (1) to (2) is an immediate consequence of the fact that for any monotone increasing function g and standard uniform random variable, U, we have

$$Q_{g(U)}(\tau) = g(Q_U(\tau)) = g(\tau),$$

where $Q_U(\tau) = \tau$ is the quantile function of U. In the above model, the autoregressive coefficients may be τ-dependent and thus can vary over the quantiles. The conditioning variables not only shift the location of the distribution of y_t, but also

Post-print version of an article published in Journal of the American Statistical Association 101(475): 980-990. doi:10.1198/016214506000000672.

may alter the scale and shape of the conditional distribution. We will refer to this model as the QAR(p) model.

We will argue that QAR models can play a useful role in expanding the modeling territory between classical stationary linear time series models and their unit root alternatives. To illustrate this in the QAR(1) case, consider the model

$$Q_{y_t}(\tau|\mathcal{F}_{t-1}) = \theta_0(\tau) + \theta_1(\tau)y_{t-1}, \tag{4}$$

with $\theta_0(\tau) = \sigma\Phi^{-1}(\tau)$, and $\theta_1(\tau) = \min\{\gamma_0 + \gamma_1\tau, 1\}$ for $\gamma_0 \in (0,1)$ and $\gamma_1 > 0$. In this model if $U_t > (1-\gamma_0)/\gamma_1$ the model generates the y_t according to the standard Gaussian unit root model, but for smaller realizations of U_t we have a mean reversion tendency. Thus, the model exhibits a form of asymmetric persistence in the sense that sequences of strongly positive innovations tend to reinforce its unit root like behavior, while occasional negative realizations induce mean reversion and thus undermine the persistence of the process. The classical Gaussian AR(1) model is obtained by setting $\theta_1(\tau)$ to a constant.

The formulation in (4) reveals that the model may be interpreted as somewhat special form of random coefficient autoregressive (RCAR) model. Such models arise naturally in many time series applications. Discussions of the role of RCAR models can be found in, *inter alia*, Nicholls and Quinn (1982), Tjøstheim(1986), Pourahmadi (1986), Brandt (1986), Karlsen(1990), and Tong (1990). In contrast to most of the literature on RCAR models, in which the coefficients are typically assumed to be stochastically independent of one another, the QAR model has coefficients that are functionally dependent.

Monotonicity of the conditional quantile functions imposes some discipline on the forms taken by the θ functions. This discipline essentially requires that the function $Q_{y_t}(\tau|y_{t-1},...,y_{t-p})$ is monotone in τ in some relevant region Υ of $(y_{t-1},...,y_{t-p})$-space. The correspondance between the random coefficient formulation of the QAR model (1) and the conditional quantile function formulation (2) presupposes the monotonicity of the latter in τ. In the region Υ where this monotonicity holds (1) can be regarded as a valid mechanism for simulating from the QAR model (2). Of course, model (1) can, even in the absence of this monotonicity, be taken as a valid data generating mechanism, however the link to the strictly linear conditional quantile model is no longer valid. At points where the monotonicity is violated the conditional quantile

Post-print version of an article published in Journal of the American Statistical Association 101(475): 980-990. doi:10.1198/016214506000000672.

functions corresponding to the model described by (1) have linear "kinks". Attempting to fit such piecewise linear models with linear specifications can be hazardous. We will return to this issue in the discussion of Section 4. In the next section we briefly describe some essential features of the QAR model.

2.2. **Properties of the QAR Process.** The QAR(p) model (1) can be reformulated in more conventional random coefficient notation as,

$$y_t = \mu_0 + \alpha_{1,t} y_{t-1} + \cdots + \alpha_{p,t} y_{t-p} + u_t \tag{5}$$

where $\mu_0 = E\theta_0(U_t)$, $u_t = \theta_0(U_t) - \mu_0$, and $\alpha_{j,t} = \theta_j(U_t)$, for $j = 1, ..., p$. Thus, $\{u_t\}$ is an iid sequence of random variables with distribution function $F(\cdot) = \theta_0^{-1}(\cdot + \mu_0)$, and the $\alpha_{j,t}$ coefficients are functions of this u_t innovation random variable. The QAR(p) process (5) can be expressed as an p-dimensional vector autoregression process of order 1:

$$Y_t = \Gamma + A_t Y_{t-1} + V_t$$

with

$$\Gamma = \begin{bmatrix} \mu_0 \\ 0_{p-1} \end{bmatrix}, \; A_t = \begin{bmatrix} A_{p-1,t} & \alpha_{p,t} \\ I_{p-1} & 0_{p-1} \end{bmatrix}, \; V_t = \begin{bmatrix} u_t \\ 0_{p-1} \end{bmatrix},$$

where $A_{p-1,t} = [\ \alpha_{1,t}, \ \ \ldots, \ \ \alpha_{p-1,t}\ ]$, $Y_t = [y_t, \cdots, y_{t-p+1}]^\top$, and 0_{p-1} is the $(p-1)$-dimensional vector of zeros. In the Appendix, we show that under regularity conditions given in the following Theorem, an $\mathcal{F}_t$-measurable solution for (5) can be found.

To formalize the foregoing discussion and facilitate later asymptotic analysis, we introduce the following conditions.

A.1: $\{u_t\}$ are *iid* random variables with mean 0 and variance $\sigma^2 < \infty$. The distribution function of u_t, F, has a continuous density f with $f(u) > 0$ on $\mathcal{U} = \{u : 0 < F(u) < 1\}$.

A.2: Let $E(A_t \otimes A_t) = \Omega_A$, the eigenvalues of Ω_A have moduli less than unity.

A.3: Denote the conditional distribution function $\Pr[y_t < \cdot | \mathcal{F}_{t-1}]$ as $F_{t-1}(\cdot)$ and its derivative as $f_{t-1}(\cdot)$, f_{t-1} is uniformly integrable on $\mathcal{U}$.

Post-print version of an article published in Journal of the American Statistical Association 101(475): 980-990. doi:10.1198/016214506000000672.

Theorem 2.1. *Under assumptions A.1 and A.2, the time series y_t given by (5) is covariance stationary and satisfies a central limit theorem*

$$\frac{1}{\sqrt{n}}\sum_{t=1}^{n}(y_t-\mu_y) \Rightarrow N\left(0,\omega_y^2\right),$$

where $\mu_y = \mu_0/(1-\sum_{j=1}^{p}\mu_j)$, $\omega_y^2 = \lim n^{-1}E[\sum_{t=1}^{n}(y_t-\mu_y)]^2$, and $\mu_j = E(\alpha_{j,t})$, $j = 1, ..., p$.

To illustrate some important features of the QAR process, we consider the simplest case of QAR(1) process,

$$y_t = \alpha_t y_{t-1} + u_t, \tag{6}$$

where $\alpha_t = \theta_1(U_t)$ and $u_t = \theta_0(U_t)$ corresponding to (4), whose properties are summarized in the following corollary.

Corollary 2.1. *If y_t is determined by (6), and $\omega_\alpha^2 = E(\alpha_t)^2 < 1$, under assumption A.1, y_t is covariance stationary and satisfies a central limit theorem*

$$\frac{1}{\sqrt{n}}\sum_{t=1}^{n} y_t \Rightarrow N\left(0,\omega_y^2\right),$$

where $\omega_y^2 = \sigma^2(1+\mu_\alpha)/((1-\mu_\alpha)(1-\omega_\alpha^2))$ with $\mu_\alpha = E(\alpha_t) < 1$.

In the example given in Section 2.1, $\alpha_t = \theta_1(U_t) = \min\{\gamma_0+\gamma_1 U_t, 1\} \leq 1$, and $\Pr(|\alpha_t| < 1) > 0$, the condition of Corollary 2.1 holds and the process y_t is globally stationary but can still display local (and asymmetric) persistency in the presence of certain type of shocks (positive shocks in the example). Corollary 2.1 also indicates that even with $\alpha_t > 1$ over some range of quantiles, as long as $\omega_\alpha^2 = E(\alpha_t)^2 < 1$, y_t can still be covariance stationary in the long run. Thus, a quantile autoregressive process may allow for some (transient) forms of explosive behavior while maintaining stationarity in the long run.

Under the assumptions in Corollary 2.1, by recursively substituting in (6), we can see that

$$y_t = \sum_{j=0}^{\infty}\beta_{t,j}u_{t-j}, \text{ where } \beta_{t,0} = 1, \text{ and } \beta_{t,j} = \prod_{i=0}^{j-1}\alpha_{t-i}, \text{ for } j \geq 1, \tag{7}$$

is a stationary $\mathcal{F}_t$-measurable solution to (6). In addition, if $\sum_{j=0}^{\infty}\beta_{t,j}v_{t-j}$ converges in L^p, then y_t has a finite p-th order moment. The $\mathcal{F}_t$-measurable solution of (6) gives

Post-print version of an article published in Journal of the American Statistical Association 101(475): 980-990. doi:10.1198/016214506000000672.

a doubly stochastic $MA(\infty)$ representation of y_t. In particular, the impulse response of y_t to a shock u_{t-j} is stochastic and is given by $\beta_{t,j}$. On the other hand, although the impulse response of the quantile autoregressive process is stochastic, it does converge (to zero) in mean square (and thus in probability) as $j \to \infty$, corroborating the stationarity of y_t. If we denote the autocovariance function of y_t by $\gamma_y(h)$, it is easy to verify that $\gamma_y(h) = \mu_\alpha^{|h|}\sigma_y^2$ where $\sigma_y^2 = \sigma^2/(1-\omega_\alpha^2)$.

Remark 2.1. *Comparing to the QAR(1) process y_t, if we consider a conventional AR(1) process with autoregressive coefficient μ_α and denote the corresponding process by $\underline{y}_t$, the long-run variance of y_t (given by ω_y^2) is (as expected) larger than that of $\underline{y}_t$. The additional variance the QAR process y_t comes from the variation of α_t. In fact, ω_y^2 can be decomposed into the summation of the long-run variance of $\underline{y}_t$ and an additional term that is determined by the variance of α_t:*

$$\omega_y^2 = \underline{\omega}_y^2 + \frac{\sigma^2}{(1-\mu_\alpha)^2(1-\omega_\alpha^2)} Var(\alpha_t),$$

where $\underline{\omega}_y^2 = \sigma^2/(1-\mu_\alpha)^2$ is the long-run variance of $\underline{y}_t$.

We consider estimation and related inference on the QAR model in the next two sections.

3. Estimation

Estimation of the quantile autoregressive model (3) involves solving the problem

$$\min_{\theta \in \mathbb{R}^{p+1}} \sum_{t=1}^{n} \rho_\tau(y_t - x_t^\top \theta), \tag{8}$$

where $\rho_\tau(u) = u(\tau - I(u<0))$ as in Koenker and Bassett (1978). Solutions, $\widehat{\theta}(\tau)$, are called autoregression quantiles. Given $\widehat{\theta}(\tau)$, the τ-th conditional quantile function of y_t, conditional on x_t, could be estimated by,

$$\hat{Q}_{y_t}(\tau|x_t) = x_t^\top \widehat{\theta}(\tau),$$

and the conditional density of y_t can be estimated by the difference quotients,

$$\hat{f}_{y_t}(\tau|x_{t-1}) = (\tau_i - \tau_{i-1})/(\hat{Q}_{y_t}(\tau_i|x_{t-1}) - \hat{Q}_{y_t}(\tau_{i-1}|x_{t-1})),$$

for some appropriately chosen sequence of τ's.

Post-print version of an article published in Journal of the American Statistical Association 101(475): 980-990. doi:10.1198/016214506000000672.

If we denote $E(y_t)$ as μ_y, $E(y_t y_{t-j})$ as γ_j, and let $\Omega_0 = E(x_t x_t^\top) = \lim n^{-1} \sum_{t=1}^n x_t x_t^\top$, then

$$\Omega_0 = \begin{bmatrix} 1 & \boldsymbol{\mu}_y^\top \\ \boldsymbol{\mu}_y & \Omega_y \end{bmatrix}$$

where $\boldsymbol{\mu}_y = \mu_y \cdot 1_{p\times 1}$, and

$$\Omega_y = \begin{bmatrix} \gamma_0 & \cdots & \gamma_{p-1} \\ \vdots & \ddots & \vdots \\ \gamma_{p-1} & \cdots & \gamma_0 \end{bmatrix}.$$

In the special case of QAR(1) model (6), $\Omega_0 = E(x_t x_t^\top) = diag[1, \gamma_0]$, $\gamma_0 = E[y_t^2]$. Let $\Omega_1 = \lim n^{-1} \sum_{t=1}^n f_{t-1}[F_{t-1}^{-1}(\tau)] x_t x_t^\top$, and define $\Sigma = \Omega_1^{-1}\Omega_0\Omega_1^{-1}$. The asymptotic distribution of $\widehat{\theta}(\tau)$ is summarized in the following Theorem.

Theorem 3.1. *Under assumptions A.1 - A.3,*

$$\Sigma^{-1/2}\sqrt{n}(\widehat{\theta}(\tau) - \theta(\tau)) \Rightarrow B_k(\tau),$$

where $B_k(\tau)$ represents a k-dimensional standard Brownian Bridge, $k = p + 1$.

By definition, for any fixed τ, $B_k(\tau)$ is $\mathcal{N}(0, \tau(1-\tau)I_k)$. In the important special case with constant coefficients, $\Omega_1 = f[F^{-1}(\tau)]\Omega_0$, where $f(\cdot)$ and $F(\cdot)$ are the density and distribution functions of u_t, respectively. We state this result in the following corollary.

Corollary 3.1. *Under assumptions A.1 - A.3, if the coefficients α_{jt} are constants, then*

$$f[F^{-1}(\tau)]\Omega_0^{1/2}\sqrt{n}(\widehat{\theta}(\tau) - \theta(\tau)) \Rightarrow B_k(\tau).$$

An alternative form of the model that is widely used in economic applications is the augmented Dickey-Fuller (ADF) regression

$$y_t = \mu_0 + \delta_{0,t} y_{t-1} + \sum_{j=1}^{p-1} \delta_{j,t} \Delta y_{t-j} + u_t, \tag{9}$$

where, corresponding to (5),

$$\delta_{0,t} = \sum_{s=1}^{p} \alpha_{s,t}, \ \delta_{j,t} = -\sum_{s=j+1}^{p} \alpha_{s,t}, \ j = 1, \cdots, p-1.$$

Post-print version of an article published in Journal of the American Statistical Association 101(475): 980-990. doi:10.1198/016214506000000672.

In the above transformed model, $\delta_{0,t}$ is the critical parameter corresponding the largest autoregressive root. Let $z_t = (1, y_{t-1}, \Delta y_{t-1}, ..., \Delta y_{t-p+1})^\top$, we may write the quantile regression counterpart of (9) as

$$Q_{y_t}(\tau|\mathcal{F}_{t-1}) = z_t^\top \delta(\tau), \tag{10}$$

where

$$\delta(\tau) = (\alpha_0(\tau), \delta_0(\tau), \delta_1(\tau), \cdots, \delta_{p-1}(\tau))^\top.$$

The limiting distributions of the quantile regression estimators $\widehat{\delta}(\tau)$ can be obtained from our previous analysis. If we define

$$J = \begin{bmatrix} 1 & 0 & 0 & \cdots & 0 \\ 0 & 1 & 1 & \cdots & 1 \\ 0 & 0 & -1 & & -1 \\ & & & \ddots & \\ 0 & 0 & 0 & \cdots & -1 \end{bmatrix}, \quad \text{and} \quad \Delta = J\Sigma J,$$

then we have, under assumptions A.1 - A.3,

$$\Delta^{-1/2}\sqrt{n}(\widehat{\delta}(\tau) - \delta(\tau)) \Rightarrow B_k(\tau).$$

If we focus our attention on the largest autoregressive root $\delta_{0,t}$ in the ADF type regression (9) and consider the special case that $\delta_{j,t} =$ constant for $j = 1, ..., p-1$, then, a result similar to Corollary 2.1 can be obtained.

Corollary 3.2. *Under assumptions A.1-A.3, if $\delta_{j,t} =$ constant for $j = 1, ..., p-1$, and $\delta_{0,t} \leq 1$ and $|\delta_{0,t}| < 1$ with positive probability, then the time series y_t given by (9) is covariance stationary and satisfies a central limit theorem.*

4. Quantile Monotonicity

As in other linear quantile regression applications, linear QAR models should be cautiously interpreted as useful local approximations to more complex nonlinear global models. If we take the linear form of the model too literally then obviously at some point, or points, there will be "crossings" of the conditional quantile functions – unless these functions are precisely parallel in which case we are back to the pure location shift form of the model. This crossing problem appears more acute in

Post-print version of an article published in Journal of the American Statistical Association 101(475): 980-990. doi:10.1198/016214506000000672.

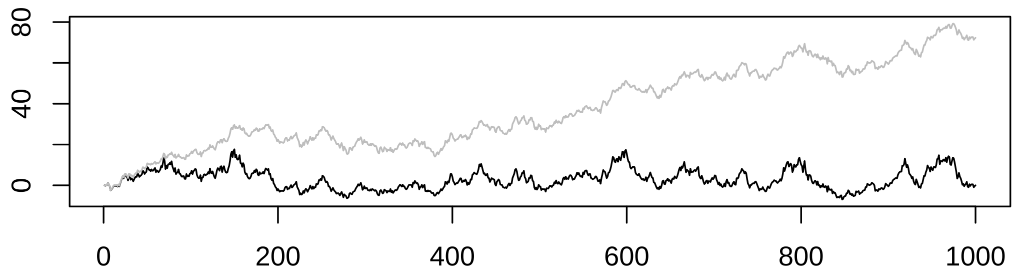

FIGURE 1. QAR and Unit Root Time-Series: The figure contrasts two time series generated by the same sequence of innovations. The grey sample path is a random walk with standard Gaussian innovations; the black sample path illustrates a QAR series generated by the same innovations with random AR(1) coefficient $.85 + .25\Phi(u_t)$. The latter series although exhibiting explosive behavior in the upper tail is stationary as described in the text.

the autoregressive case than in ordinary regression applications since the support of the design space, i.e. the set of x_t that occur with positive probability, is determined within the model. Nevertheless, we may still regard the linear models specified above as valid local approximations over a region of interest.

It should be stressed that the *estimated* conditional quantile functions,

$$\hat{Q}_y(\tau|x) = x^\top \widehat{\theta}(\tau),$$

are guaranteed to be monotone at the mean design point, $x = \bar{x}$, as shown in Bassett and Koenker (1982), for linear quantile regression models. In our random coefficient view of the QAR model,

$$y_t = x_t^\top \theta(U_t),$$

we express the observable random variable y_t as a linear function of conditioning covariates. But rather than assuming that the coordinates of the vector θ are independent random variables we adopt a diametrically opposite viewpoint – that they are perfectly functionally dependent, all driven by a single random uniform variable. If the functions $(\theta_0, ..., \theta_p)$ are all monotonically increasing then the coordinates of

Post-print version of an article published in Journal of the American Statistical Association 101(475): 980-990. doi:10.1198/016214506000000672.

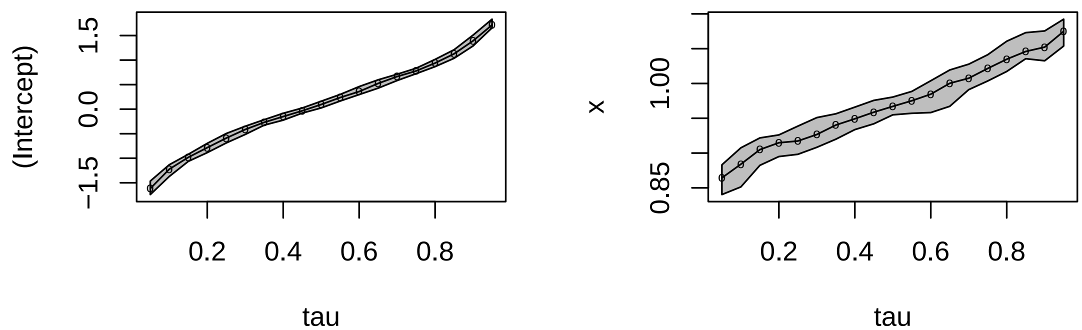

Figure 2. Estimating the QAR model: The figure illustrates estimates of the QAR(1) model based on the black time series of the previous figure. The left panel represents the intercept estimate at 19 equally spaced quantiles, the right panel represents the AR(1) slope estimate at the same quantiles. The shaded region is a .90 confidence band. Note that the slope estimate quite accurate reproduces the linear form of the QAR(1) coefficient used to generate the data.

the random vector α_t are said to be comonotonic in the sense of Schmeidler (1986).[1] This is often the case, but there are important cases for which this monotonicity fails. What then?

What really matters is that we can find a linear reparameterization of the model that does exhibit comonotonicity over some relevant region of covariate space. Since for any nonsingular matrix A we can write,

$$Q_y(\tau|x) = x^\top A^{-1} A\theta(\tau),$$

we can choose $p+1$ linearly independent design points $\{x_s : s = 1, ..., p+1\}$ where $Q_y(\tau|x_s)$ is monotone in τ, then choosing the matrix A so that Ax_s is the sth unit basis vector for $\mathbb{R}^{p+1}$ we have

$$Q_y(\tau|x_s) = \gamma_s(\tau),$$

[1]Random variables X and Y on a probability space $(\Omega, \mathcal{A}, P)$ are said to be comonotonic if there are monotone functions, g and h and a random variable Z on $(\Omega, \mathcal{A}, P)$ such that $X = g(Z)$ and $Y = h(Z)$.

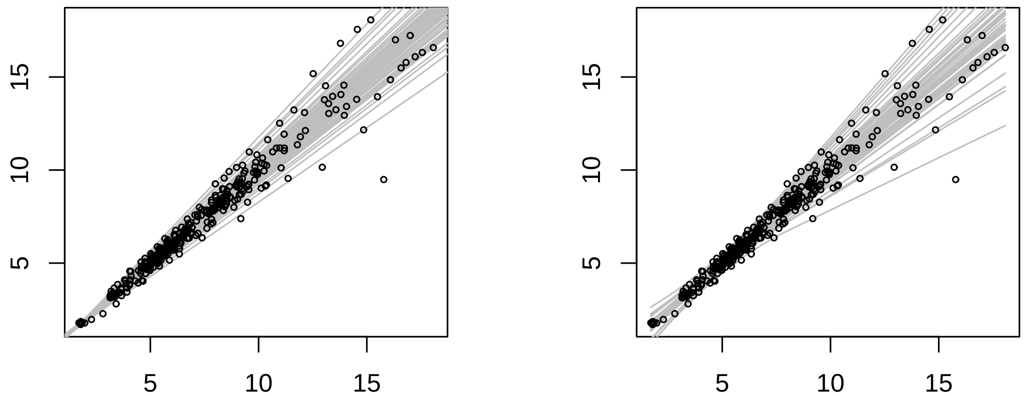

FIGURE 3. QAR(1) Model of U.S. Short Term Interest Rate: The AR(1) scatterplot of the U.S. three month rate is superimposed in the left panel with 49 equally spaced estimates of linear conditional quantile functions. In the right panel the model is augmented with a nonlinear (quadratic) component. The introduction of the quadratic component alleviates some nonmonotonicity in the estimated quantiles at low interest rates.

where $\gamma = A\theta$. And now inside the convex hull of our selected points we have a comonotonic random coefficient representation of the model. In effect, we have simply reparameterized the design so that the $p + 1$ coefficients are the conditional quantile functions of y_t at the selected points. The fact that quantile functions of sums of nonnegative comonotonic random variables are sums of their marginal quantile functions, see e.g. Denneberg(1994) or Bassett, Koenker and Kordas (2004), allows us to interpolate inside the convex hull. Of course, linear extrapolation is also possible but we must be cautious about possible violations of the monotonicity requirement in this region.

The interpretation of linear conditional quantile functions as approximations to the local behavior in central range of the covariate space should always be regarded as provisional; richer data sources can be expected to yield more elaborate nonlinear specifications that would have validity over larger regions. Figure 1 illustrates a

Post-print version of an article published in Journal of the American Statistical Association 101(475): 980-990. doi:10.1198/016214506000000672.

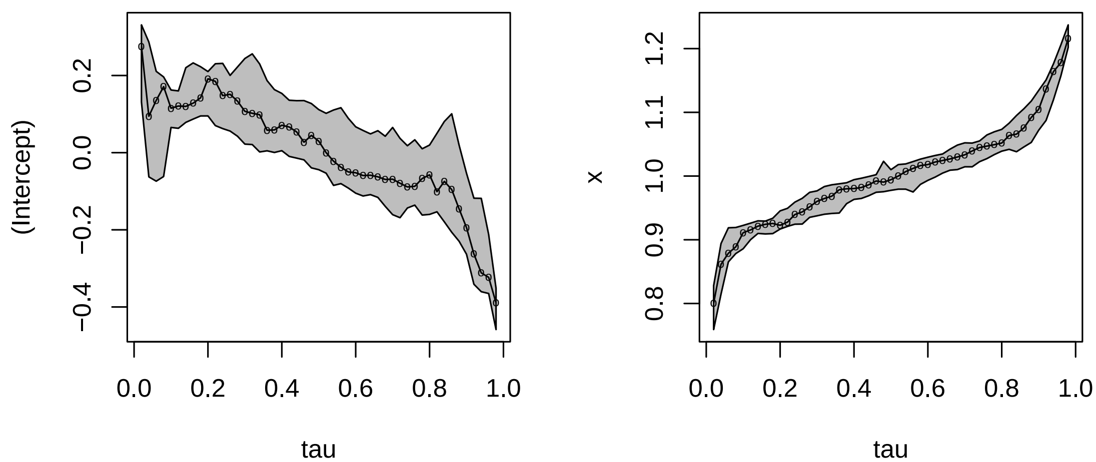

Figure 4. QAR(1) Model of U.S. Short Term Interest Rate: The QAR(1) estimates of the intercept and slope parameters for 19 equally spaced quantile functions are illustrated in the two plots. Note that the slope parameter is, like the prior simulated example, explosive in the upper tail but mean reverting in the lower tail.

realization of the simple QAR(1) model described in Section 2. The black sample path shows 1000 observations generated from the model (4) with AR(1) coefficient $\theta_1(u) = .85 + .25u$ and $\theta_0(u) = \Phi^{-1}(u)$. The grey sample path depicts the a random walk generated from the same innovation sequence, i.e. the same $\theta_0(U_t)$'s but with constant θ_1 equal to one. It is easy to verify that the QAR(1) form of the model satisfies the stationarity conditions of Section 2.2, and despite the explosive character of its upper tail behavior we observe that the series appears quite stationary, at least by comparison to the random walk series. Estimating the QAR(1) model at 19 equally spaced quantiles yields the intercept and slope estimates depicted in Figure 2.

Figure 3 depicts estimated linear conditional quantile functions for short term (three month) US interest rates using the QAR(1) model superimposed on the AR(1) scatter plot. In this example the scatterplot shows clearly that there is more dispersion at higher interest rates, with nearly degenerate behavior at very low rates. The fitted linear quantile regression lines in the left panel show little evidence of

Post-print version of an article published in Journal of the American Statistical Association 101(475): 980-990. doi:10.1198/016214506000000672.

crossing, but at rates below .04 there are some violations of the monotonicity requirement in the fitted quantile functions. Fitting the data using a somewhat more complex nonlinear (in variables) model by introducing a another additive component $\theta_2(\tau)(y_{t-1}-\delta)^2 I(y_{t-1}<\delta)$ with $\delta=8$ in our example we can eliminate the problem of the crossing of the fitted quantile functions. In Figure 4 depicting the fitted coefficients of the QAR(1) model and their confidence region, we see that the estimated slope coefficient of the QAR(1) model has somewhat similar appearance to the simulated example. Even more flexible models may be needed in other settings. A B-spline expansion QAR(1) model for Melbourne daily temperature is described in Koenker(2000) illustrating this approach.

The statistical properties of nonlinear QAR models and associated estimators are much more complicated than the linear QAR model that we study in the present paper. Despite the possible crossing of quantile curves, we believe that the linear QAR model provides a convenient and useful local approximation to nonlinear QAR models. Such simplied QAR models can still deliver important insight about dynamics, e.g. adjustment asymmetries, in economic time series and thus provides a useful tool in empirical diagnostic time series analysis.

5. Inference On The QAR Process

In this section, we turn our attention to inference in QAR models. Although other inference problems can be analyzed, we consider here the following inference problems that are of paramount interest in many applications. The first hypothesis is the quantile regression analog of the classical representation of linear restrictions on θ: (1) $H_{01}: R\theta(\tau)=r$, with known R and r, where R denotes an $q\times p$-dimensional matrix and r is an q-dimensional vector. In addition to the classical inference problem, we are also interested in testing for asymmetric dynamics under the QAR framework. Thus we consider the hypothesis of parameter constancy, which can be formulated in the form of: (2) $H_{02}: R\theta(\tau)=r$, with unknown but estimable r. We consider both the cases at specific quantiles τ (say, median, lower quartile, upper quartile) and the case over a range of quantiles $\tau\in\mathcal{T}$.

5.1. **The Regression Wald Process and Related Tests.** Under the linear hypothesis $H_{01}: R\theta(\tau)=r$ and assumptions A.1-A.3, we have

$$V_n(\tau)=\sqrt{n}\left[R\Omega_1^{-1}\Omega_0\Omega_1^{-1}R^\top\right]^{-1/2}(R\widehat{\theta}(\tau)-r)\Rightarrow B_q(\tau), \tag{11}$$

Post-print version of an article published in Journal of the American Statistical Association 101(475): 980-990. doi:10.1198/016214506000000672.

where $B_q(\tau)$ represents a q-dimensional standard Brownian Bridge. For any fixed τ, $B_q(\tau)$ is $\mathcal{N}(0, \tau(1-\tau)I_q)$. Thus, the regression Wald process can be constructed as

$$W_n(\tau) = n(R\widehat{\theta}(\tau) - r)^\top [\tau(1-\tau)R\widehat{\Omega}_1^{-1}\widehat{\Omega}_0\widehat{\Omega}_1^{-1}R^\top]^{-1}(R\widehat{\theta}(\tau) - r),$$

where $\widehat{\Omega}_1$ and $\widehat{\Omega}_0$ are consistent estimators of Ω_1 and Ω_0. If we are interested in testing $R\theta(\tau) = r$ over $\tau \in \mathcal{T}$, we may consider, say, the following Kolmogorov-Smirnov (KS) type sup-Wald test:

$$KSW_n = \sup_{\tau \in \mathcal{T}} W_n(\tau),$$

If we are interested in testing $R\theta(\tau) = r$ at a particular quantile $\tau = \tau_0$, a Chi-square test can be conducted based on the statistic $W_n(\tau_0)$. The limiting distributions are summarized in the following theorem.

Theorem 5.1. *Under assumptions A.1-A.3 and the linear restriction H_{01},*

$$W_n(\tau_0) \Rightarrow \chi_q^2, \text{ and } KSW_n = \sup_{\tau \in \mathcal{T}} W_n(\tau) \Rightarrow \sup_{\tau \in \mathcal{T}} Q_q^2(\tau),$$

where $Q_q(\tau) = \|B_q(\tau)\| / \sqrt{\tau(1-\tau)}$ is a Bessel process of order q, where $\|\cdot\|$ represents the Euclidean norm. For any fixed τ, $Q_q^2(\tau) \sim \chi_q^2$ is a centered Chi-square random variable with q-degrees of freedom.

5.2. **Testing For Asymmetric Dynamics.** The hypothesis that $\theta_j(\tau)$, $j = 1, \ldots, p$, are constants over τ (i.e. $\theta_j(\tau) = \mu_j$) can be represented in the form of $H_{02} : R\theta(\tau) = r$ by taking $R = [0_{p\times 1} \vdots I_p]$ and $r = [\mu_1, \cdots, \mu_p]^\top$, with unknown parameters $\mu_1, \cdots, \mu_p$. The Wald process and associated limiting theory provide a natural test for the hypothesis $R\theta(\tau) = r$ when r is known. To test the hypothesis with unknown r, appropriate estimator of r is needed. In many econometrics applications, a $\sqrt{n}$-consistent estimator of r is available. If we look at the process

$$\widehat{V}_n(\tau) = \sqrt{n}\left[R\widehat{\Omega}_1^{-1}\widehat{\Omega}_0\widehat{\Omega}_1^{-1}R^\top\right]^{-1/2}(R\widehat{\theta}(\tau) - \widehat{r}),$$

then under H_{02}, we have,

$$\begin{aligned}
\widehat{V}_n(\tau) &= \sqrt{n}\left[R\widehat{\Omega}_1^{-1}\widehat{\Omega}_0\widehat{\Omega}_1^{-1}R^\top\right]^{-1/2}(R\widehat{\theta}(\tau) - r) - \sqrt{n}\left[R\widehat{\Omega}_1^{-1}\widehat{\Omega}_0\widehat{\Omega}_1^{-1}R^\top\right]^{-1/2}(\widehat{r} - r) \\
&\Rightarrow B_q(\tau) - f(F^{-1}(\tau))\left[R\Omega_0^{-1}R^\top\right]^{-1/2} Z
\end{aligned}$$

where $Z = \lim \sqrt{n}(\widehat{r} - r)$. The necessity of estimating r introduces a drift component in addition to the simple Brownian bridge process, invalidating the distribution-free character of the original Kolmogorov-Smirnov (KS) test.

To restore the asymptotically distribution free nature of inference, we employ a martingale transformation proposed by Khmaladze (1981) over the process $\widehat{V}_n(\tau)$. Denote $df(x)/dx$ as $\dot{f}$, and define

$$\dot{g}(r) = (1, (\dot{f}/f)(F^{-1}(r)))^\top, \text{ and } C(s) = \int_s^1 \dot{g}(r)\dot{g}(r)^\top dr,$$

we construct a martingale transformation $\mathcal{K}$ on $\widehat{V}_n(\tau)$ defined as:

$$\widetilde{V}_n(\tau) = \mathcal{K}\widehat{V}_n(\tau) = \widehat{V}_n(\tau) - \int_0^\tau \left[\dot{g}_n(s)^\top C_n^{-1}(s) \int_s^1 \dot{g}_n(r) d\widehat{V}_n(r)\right] ds, \tag{12}$$

where $\dot{g}_n(s)$ and $C_n(s)$ are uniformly consistent estimators of $\dot{g}(r)$ and $C(s)$ over $\tau \in \mathcal{T}$, and propose the following Kolmogorov-Smirnov[2] type test based on the transformed process:

$$KH_n = \sup_{\tau \in \mathcal{T}} \left\| \widetilde{V}_n(\tau) \right\|. \tag{13}$$

Under the null hypothesis, the transformed process $\widetilde{V}_n(\tau)$ converges to a standard Brownian motion. For more discussions of quantile regression inference based on the martingale transformation approach, see, Koenker and Xiao (2002) and references therein. We make the following assumptions on the estimators:

A.4: There exist estimators $\dot{g}_n(\tau)$, $\widehat{\Omega}_0$ and $\widehat{\Omega}_1$ satisfying:
i.: $\sup_{\tau \in} |\dot{g}_n(\tau) - \dot{g}(\tau)| = o_p(1)$,
ii.: $||\widehat{\Omega}_0 - \Omega_0|| = o_p(1)$, $||\widehat{\Omega}_1 - \Omega_1|| = o_p(1)$, $\sqrt{n}(\widehat{r} - r) = O_p(1)$.

Theorem 5.2. *Under the assumptions A.1 - A.4 and the hypothesis H_{02},*

$$\widetilde{V}_n(\tau) \Rightarrow W_q(\tau), \; KH_n = \sup_{\tau \in \mathcal{T}} \left\| \widetilde{V}_n(\tau) \right\| \Rightarrow \sup_{\tau \in \mathcal{T}} \|W_q(\tau)\|,$$

where $W_q(r)$ is a q-dimensional standard Brownian motion.

The martingale transformation is based on function $\dot{g}(s)$ which needs to be estimated. There are several approaches to estimate the score: $\frac{f'}{f}(F^{-1}(s))$. Portnoy and Koenker (1989) studied adaptive estimation and employed kernel-smoothing method

[2] A Cramer-von-Mises type test based on the transformed process can also be constructed and analysed in a similar way.

Post-print version of an article published in Journal of the American Statistical Association 101(475): 980-990. doi:10.1198/016214506000000672.

in estimating the density and score functions, uniform consistency of the estimators is also discussed. Cox (1985) proposed an elegant smoothing spline approach to the estimation of f'/f, and Ng (1994) provided an efficient algorithm for computing this score estimator. Estimation of Ω_0 is straightforward: $\widehat{\Omega}_0 = n^{-1}\sum_t x_t x_t^\top$. For the estimation of $\widehat{\Omega}_1$, see, inter alia, Koenker (1994), Powell (1989), and Koenker and Machado (1999) for related discussions.

6. Monte Carlo

A Monte Carlo experiment is conducted in this section to examine the QAR-based inference procedures. We are particularly interested in time series displaying asymmetric dynamics. Thus, we consider the QAR model with $p = 1$ and test the hypothesis that $\alpha_1(\tau) =$ constant over τ.

The data in our experiments were generated from model (6), where u_t are i.i.d. random variables. We consider the Kolmogorov-Smirnov test KH_n given by (13) for different sample sizes ($n = 100$ and 300) and innovation distributions, and choose $\mathcal{T} = [0.1, 0.9]$. Both normal innovations and student-t innovations are considered. The number of repetitions is 1000.

Representative results of the empirical size and power of the proposed tests are reported in Tables 1-3. We report the empirical size of this test for three choices of α_t : (1) $\alpha_t = 0.95$; (2) $\alpha_t = 0.9$; (3) $\alpha_t = 0.6$. The first two choices (0.95 and 0.9) are large and close to unity so that the corresponding time series display cartain degree of (symmetric) persistence. For models under the alternative, we considered the following four choices of α_t:

$$\begin{aligned}
\alpha_t &= \varphi_1(u_t) = \begin{cases} 1, & u_t \geq 0, \\ 0.8, & u_t < 0, \end{cases} \\
\alpha_t &= \varphi_2(u_t) = \begin{cases} 0.95, & u_t \geq 0, \\ 0.8, & u_t < 0, \end{cases} \\
\alpha_t &= \varphi_3(u_t) = \min\{0.5 + F_u(u_t), 1\}, \\
\alpha_t &= \varphi_4(u_t) = \min\{0.75 + F_u(u_t), 1\}.
\end{aligned} \tag{14}$$

These alternatives deliver processes with different types of asymmetric (or local) persistency. In particular, when $\alpha_t = \varphi_1(u_t)$, $\varphi_3(u_t)$, $\varphi_4(u_t)$, y_t display unit root behavior in the presence of positive or large values of innovations, but have a mean

Post-print version of an article published in Journal of the American Statistical Association 101(475): 980-990. doi:10.1198/016214506000000672.

reversion tendency with negative shocks. The alternative $\alpha_t = \varphi_2(u_t)$ has local to (or weak) unit root behavior in the presence of positive innovations, and behave more stationarily when there are negative shocks.

The construction of tests uses estimators of the density and score. We estimate the density (or sparsity) function using the approach of Siddiqui (1960). The density estimation entails a choice of bandwidth. We consider the bandwidth choices suggested by Hall and Sheather (1988) and Bofinger (1975) and rescaled versions of them. A bandwidth rule that Hall and Sheather (1988) suggested based on Edgeworth expansion for studentized quantiles (and using Gaussian plug-in) is

$$h_{HS} = n^{-1/3} z_\alpha^{2/3} [1.5\phi^2(\Phi^{-1}(t))/(2(\Phi^{-1}(t))^2 + 1)]^{1/3},$$

where z_α satisfies $\Phi(z_\alpha) = 1 - \alpha/2$ for the construction of $1 - \alpha$ confidence intervals. Another bandwidth selection has been proposed by Bofinger (1975) based on minimizing the mean squared error of the density estimator and is of order $n^{-1/5}$. If we plug-in the Gaussian density, we obtain the following bandwidth that has been widely used in practice:

$$h_B = n^{-1/5} [4.5\phi^4(\Phi^{-1}(t))/(2(\Phi^{-1}(t))^2 + 1)^2]^{1/5}.$$

Monte Carlo results indicate that the Hall-Sheather bandwidth provides a good lower bound and the Bofinger bandwidth provides a reasonable upper bound for bandwidth in testing parameter constancy. For this reason, we consider bandwidth choices between h_{HS} and h_B. In particular, we consider rescaled versions of h_B and h_{HS} (θh_B and δh_{HS}, where $0 < \theta < 1$ and $\delta > 1$ are scalars) in our Monte Carlo and representative results are reported. Bandwidth values that are constant over the whole range of quantiles are not recommended. The sampling performance of tests using a constant bandwidth turned out to be poor, and are inferior than bandwidth choices such as the Hall/Sheather or Bofinger bandwidth that varies over the quantiles. For these reason, we focus on bandwidth h_B, h_{HS}, θh_B, and δh_{HS}. The Monte Carlo results indicate that the test using a rescaled version of Bofinger bandwidth ($h = 0.6h_B$) yields good performance in the cases that we study.

The score function was estimated by the method of Portnoy and Koenker (1989) and we choose the Silverman (1986) bandwidth in our Monte Carlo. Our simulation results show that the test is more affected by the estimation of the density than that of the score. Intuitively, the estimator of the density plays the role of a scalar and thus

Post-print version of an article published in Journal of the American Statistical Association 101(475): 980-990. doi:10.1198/016214506000000672.

	Model	$h = 3h_{HS}$	$h = h_{HS}$	$h = h_B$	$h = 0.6h_B$
Size	$\alpha_t = 0.95$	0.073	0.287	0.018	0.056
	$\alpha_t = 0.9$	0.073	0.275	0.01	0.046
	$\alpha_t = 0.6$	0.07	0.287	0.012	0.052
Power	$\alpha_t = \varphi_1(u_t)$	0.474	0.795	0.271	0.391
	$\alpha_t = \varphi_2(u_t)$	0.262	0.620	0.121	0.234
	$\alpha_t = \varphi_3(u_t)$	0.652	0.939	0.322	0.533
	$\alpha_t = \varphi_4(u_t)$	0.159	0.548	0.046	0.114

Table 1. Empirical Size and Power of Tests of Constancy of the Coefficient α with Gaussian Innovations: Models for size employ the indicated constant coefficient; models for power comparisons are those indicated in (14). Sample size is 100, and number of replications is 1000.

has the largest influence. The Monte Carlo results also indicates that the method of Portnoy and Koenker (1989) coupled with the Silverman bandwidth has reasonably good performance. Table 1 reports the empirical size and power for the case with Gaussian innovations and sample size $n = 100$. Table 2 reports results when u_t are student-t innovations (with 3 degrees of freedom) and $n = 100$. Results in Table 2 confirm that, using the quantile regression based approach, power gain can be obtained in the presence of heavy-tailed disturbances. (Such gains obviously depend on choosing quantiles at which there is sufficient conditional density.) Experiments based on larger sample sizes are also conductedand. Table 3 reports the size and power for the case with Gaussian innovations and sample size $n = 300$. These results are qualitatively similar to those of Table 1, but also show that, as the sample sizes increase, the tests do have improved size and power properties, corroborating the asymptotic theory.

7. Empirical Applications

There have been many claims and observations that some economic time series display asymmetric dynamics. For example, it has been observed that increases in the unemployment rate are sharper than declines. If an economic time series displays asymmetric dynamics systematically, then appropriate models are needed to incorporate such behavior. In this section, we apply the QAR model to two economic

	Model	$h = 3h_{HS}$	$h = h_{HS}$	$h = h_B$	$h = 0.6h_B$
Size	$\alpha_t = 0.95$	0.086	0.339	0.011	0.059
	$\alpha_t = 0.9$	0.072	0.301	0.015	0.043
	$\alpha_t = 0.6$	0.072	0.305	0.013	0.038
Power	$\alpha_t = \varphi_1(u_t)$	0.556	0.819	0.319	0.444
	$\alpha_t = \varphi_2(u_t)$	0.348	0.671	0.174	0.279
	$\alpha_t = \varphi_3(u_t)$	0.713	0.933	0.346	0.55
	$\alpha_t = \varphi_4(u_t)$	0.284	0.685	0.061	0.162

TABLE 2. Empirical Size and Power of Tests of Constancy of the Coefficient α with t(3) Innovations: Configurations as in Table 1.

	Model	$h = 3h_{HS}$	$h = h_{HS}$	$h = h_B$	$h = 0.6h_B$
Size	$\alpha_t = 0.95$	0.081	0.191	0.028	0.049
	$\alpha_t = 0.90$	0.098	0.189	0.030	0.056
	$\alpha_t = 0.60$	0.097	0.160	0.020	0.045
Power	$\alpha_t = \varphi_1(u_t)$	0.974	0.992	0.921	0.937
	$\alpha_t = \varphi_2(u_t)$	0.831	0.923	0.685	0.763
	$\alpha_t = \varphi_3(u_t)$	0.998	1.000	0.971	0.989
	$\alpha_t = \varphi_4(u_t)$	0.557	0.897	0.235	0.392

TABLE 3. Empirical Size and Power of Tests of Constancy of the Coefficient α with Gaussian Innovations: Configurations as in Table 1, except sample size is 300.

time series: unemployment rates and retail gasoline prices in the US. Our empirical analysis indicate that both series display asymmetric dynamics.

7.1. **Unemployment Rate.** Many studies on unemployment suggest that the response of unemployment to expansionary or contractionary shocks may be asymmetric. An asymmetric response to different types of shocks has important implications in economic policy. In this section, we examine unemployment dynamics using the proposed procedures.

The data that we consider are quarterly and annual rates of unemployment in the US. In particular, we looked at (seasonally adjusted) quarterly rates, starting from the first quarter of 1948 and ending at the last quarter of 2003, with 224 observations. and the annual rates are from 1890 to 1996. Many empirical studies in the unit root literature have investigated unemployment rate data. Nelson and Plosser (1982)

Post-print version of an article published in Journal of the American Statistical Association 101(475): 980-990. doi:10.1198/016214506000000672.

Frequency	τ	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
Annual	$\delta_0(\tau)$	0.740	0.776	0.929	0.871	0.858	0.793	0.727	0.680	0.599
Quarterly	$\delta_0(\tau)$	0.912	0.908	0.931	0.919	0.951	0.959	0.967	0.962	0.953

Table 4. Estimates of the Largest AR Root at Each Decile of Unemployment

Bandwidth	$0.6h_B$	$3h_{HS}$	5% CV
Annual	4.89	5.12	4.523
Quarterly	4.46	5.36	3.393

Table 5. Kolmogorov Test of Constant AR Coefficient for Unemployment

studied the unit root property of annual US unemployment rates in their seminal work on fourteen macroeconomic time series. Evidence based on the unit root tests suggests that the series is stationary. This series and other type unemployment rates have been often re-examined in later analysis.

We first apply regression (10) on the unemployment rates. We use the BIC criterion of Schwarz (1978) and Rissanen (1978) in selecting the appropriate lag length of the autoregressions. The selected lag length is $p = 3$ for the annual data and $p = 2$ for the quarterly data. The OLS estimation of the largest autoregressive root is 0.718 for the annual series and 0.941 for the quarterly rates. Quantile autoregression was also performed for each deciles. Estimates of the largest autoregressive root at each quantile are reported in Table 4. These estimated values are different over different quantiles, displaying asymmetric dynamics.

We then test asymmetric dynamics using the martingale transformation based Kolmogorov-Smirnov procedure (13) based on quantile autoregression (8). According to the suggestion from the Monte Carlo results, we choose the rescaled Hall and Sheather (1988) bandwidth $3h_{HS}$ and the rescaled Bofinger (1975) bandwidth $0.6h_B$ in estimating the density function. The tests were constructed over $\tau \in T = [0.05, 0.95]$ and results are reported in Table 5. The empirical results indicate that asymmetric behavior exist in these series.

7.2. **Retail Gasoline Price Dynamics.** Our second application investigates the asymmetric price dynamics in the retail gasoline market. We consider weekly data of US regular gasoline retail price from August 20, 1990 to Februry 16, 2004. The sample size is 699. Evidence from OLS-based ADF tests of the null hypothesis of a

τ	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
$\widehat{\delta}_0(\tau)$	0.948	0.958	0.971	0.980	0.996	1.005	1.016	1.024	1.047

TABLE 6. Estimated Largest AR Root at each Decile of Retail Gasoline Price.

unit root is mixed. The unit root null is rejected by the coefficient based test ADF_α, with a test statistic of -17.14 and critical value of -14.1, but can not be rejected by the t ratio based test ADF_t, given the test statistic -2.67 and critical value -2.86. Again we use the BIC criterion to select the lag length to obtain $p = 4$ for these tests.

We next consider quantile regression evidence based on the ADF model (9) on persistency of retail gasoline prices. Table 6 reports the estimates of the largest autoregressive roots $\widehat{\delta}_0(\tau)$ at each decile. These results suggest that the gasoline price series has asymmetric dynamics. The estimate takes quite different values over different quantiles. Estimates, $\widehat{\delta}_0(\tau)$, monotonically increase as we move from lower quantiles to higher quantiles. The autoregressive coefficient values at the lower quantiles are relatively small, indicating that the local behavior of the gasoline price would be stationary. However, at higher quantiles, the largest autoregressive root is close to or even slightly above unity, consequently the time series display unit root or locally explosive behavior at upper quantiles.

A formal test of the null hypothesis that gasoline prices have a constant autoregressive coefficent is conducted using the Kolmogorov-Smirnov procedure (13) based on quantile autoregression (2) and martingale transformation (12). Constancy of coefficients is rejected. The calculated Kolmogorov-Smirnov statistic (using the rescaled Bofinger (1975) bandwidth $0.6h_B$) is 8.347735 (lag length $p = 4$), which is considerably larger than the 5% level critical value of 5.56. However, taking into account the possibility of unit root behavior under the null, we also consider the following (coefficient-based) empirical quantile process $U_n(\tau) = n(\widehat{\delta}_0(\tau) - 1)$, and the Kolmogorov-Smirnov (KS) or Cramer-von-Mises (CvM) type tests:

$$QKS_\alpha = \sup_{\tau \in \mathcal{T}} |U_n(\tau)|, QCM_\alpha = \int_{\tau \in \mathcal{T}} U_n(\tau)^2 d\tau. \tag{15}$$

Using the results of unit root quantile regression asymptotics provided by Koenker and Xiao (2004), we have, under the unit root hypothesis,

$$U_n(\tau) \Rightarrow U(\tau) = \frac{1}{f(F^{-1}(\tau))} \left[\int_0^1 \underline{B}_y^2 \right]^{-1} \int_0^1 \underline{B}_y dB_\psi^\tau. \tag{16}$$

Post-print version of an article published in Journal of the American Statistical Association 101(475): 980-990. doi:10.1198/016214506000000672.

where $\underline{B}_w(r)$ and $B^\tau_\psi(r)$ are limiting processes of $n^{-1/2}\sum_{t=1}^{[nr]}\Delta y_t$ and $n^{-1/2}\sum_{t=1}^{[nr]}\psi_\tau(u_{t\tau})$). We adopt the approach of Koenker and Xiao (2004) and approximate the distributions of the limiting variates by resampling method and construct bootstrap tests for the unit root hypothesis based on (15).

We consider both the QKS_α and QCM_α tests for the null hypothesis of a constant AR coefficient equal to unity. Both tests firmly reject the null with test statistics of 35.79 and 320.41 respectively, and 5% level critical values of 13.22 and 19.72. The critical values were computed based on the resampling procedure described in Koenker and Xiao (2004). These results, together with the point estimates reported in Table 6, indicate that the gasoline price series has asymmetric adjustment dynamics and thus is not well characterized as a constant coefficient unit root process.

8. APPENDIX: PROOFS

8.1. **Proof of Theorem 2.1.** Giving a p-th order autoregression process (5), we denote $E(\alpha_{j,t}) = \mu_j$, and assume that $1-\sum\mu_j \neq 0$. Let $\mu_y = \mu_0/(1-\sum_{j=1}^p \mu_j)$, and denote

$$\underline{y}_t = y_t - \mu_y$$

we have

$$\underline{y}_t = \alpha_{1,t}\underline{y}_{t-1} + \cdots + \alpha_{p,t}\underline{y}_{t-p} + v_t, \tag{17}$$

where

$$v_t = u_t + \mu\sum_{l=1}^{p}(\alpha_{l,t} - \mu_l).$$

It's easy to see that $Ev_t = 0$ and $Ev_tv_s = 0$ for any $t \neq s$ since $E\alpha_{l,t} = \mu_l$ and u_t are independent. In order to derive stationarity conditions for the process $\underline{y}_t$, we first find an $\mathcal{F}_t$-measurable solution for (17). We define the $p\times 1$ random vectors

$$\underline{Y}_t = [\underline{y}_t, \cdots, \underline{y}_{t-p+1}]^\top,\ V_t = [v_t, 0, \cdots, 0]^\top$$

and the $p \times p$ random matrix

$$A_t = \begin{bmatrix} A_{p-1,t} & \alpha_{p,t} \\ I_{p-1} & 0_{p-1} \end{bmatrix},$$

where $A_{p-1,t} = [\alpha_{1,t}, \ldots, \alpha_{p-1,t}]$ and 0_{p-1} is the $(p-1)$-dimensional vector of zeros, then

$$E(V_tV_t^\top) = \begin{bmatrix} \sigma_v^2 & 0_{1\times(p-1)} \\ 0_{(p-1)\times 1} & 0_{(p-1)\times(p-1)} \end{bmatrix} = \Sigma$$

Post-print version of an article published in Journal of the American Statistical Association 101(475): 980-990. doi:10.1198/016214506000000672.

and the original process can be written as

$$\underline{Y}_t = A_t \underline{Y}_{t-1} + V_t$$

By substitution, we have

$$\begin{aligned} \underline{Y}_t &= V_t + A_t V_{t-1} + A_t A_{t-1} V_{t-2} + [A_t \cdots A_{t-m+1}] V_{t-m} + [A_t \cdots A_{t-m}] \underline{Y}_{t-m-1} \\ &= \underline{Y}_{t,m} + R_{t,m} \end{aligned}$$

where

$$\underline{Y}_{t,m} = \sum_{j=0}^{m} B_j V_{t-j}, \;\; R_{t,m} = B_{m+1} \underline{Y}_{t-m-1}, \text{ and } B_j = \left\{ \begin{array}{l} \prod_{l=0}^{j-1} A_{t-l}, \;\; j \geq 1. \\ I, \;\; j = 0. \end{array} \right. .$$

The stationarity of an $\mathcal{F}_t$-measurable solution for y_t involves the convergence of $\{\sum_{j=0}^{m} B_j V_{t-j}\}$ and $\{R_{t,m}\}$ as m increases, for fixed t. Following a similar analysis as Nicholls and Quinn (1982, Chapter 2), We need to verify that $\text{vec}E\left[\underline{Y}_{t,m}\underline{Y}_{t,m}^{\top}\right]$ converges as $m \to \infty$. Notice that B_j is independent with V_{t-j} and $\{u_t, \, t = 0, \pm 1, \pm 2, \cdots\}$ are independent random variables, thus, $\{B_j V_{t-j}\}_{j=0}^{\infty}$ is an orthogonal sequence in the sense that $E[B_j V_{t-j} B_k V_{t-k}] = 0$ for any $j \neq k$. Thus

$$\text{vec}E\left[\underline{Y}_{t,m}\underline{Y}_{t,m}^{\top}\right] = \text{vec}E\left[(\sum_{j=0}^{m} B_j V_{t-j})(\sum_{j=0}^{m} B_j V_{t-j})^{\top}\right] = \text{vec}E\left[\sum_{j=0}^{m} B_j V_{t-j} V_{t-j}^{\top} B_j^{\top}\right]$$

Notice that $\text{vec}(ABC) = (C^{\top} \otimes A)\text{vec}(B)$, and $\left(\prod_{l=0}^{j} A_l\right) \otimes \left(\prod_{k=0}^{j} B_k\right) = \prod_{k=0}^{j}(A_k \otimes B_k)$, we have

$$\begin{aligned} \text{vec}E\left[\sum_{j=0}^{m} B_j V_{t-j} V_{t-j}^{\top} B_j^{\top}\right] &= E\left[\sum_{j=0}^{m} (B_j \otimes B_j)\text{vec}(V_{t-j}V_{t-j}^{\top})\right] \\ &= E\left[\sum_{j=0}^{m} \left(\prod_{l=0}^{j-1} A_{t-l}\right) \otimes \left(\prod_{l=0}^{j-1} A_{t-l}\right) \text{vec}(V_{t-j}V_{t-j}^{\top})\right] \\ &= \sum_{j=0}^{m} \prod_{l=0}^{j-1} E(A_{t-l} \otimes A_{t-l}) \text{vec}E(V_{t-j}V_{t-j}^{\top}) \end{aligned}$$

If we denote

$$A = E[A_t] = \left[\begin{array}{cc} \overline{\mu}_{p-1} & \alpha_p \\ I_{p-1} & 0_{p-1} \end{array} \right],$$

where $\overline{\mu}_{p-1} = [\alpha_1, \ldots, \alpha_{p-1}]$, then $A_t = A + \Xi_t$, where $E(\Xi_t) = 0$, and

$$E(A_{t-l} \otimes A_{t-l}) = E\left[(A + \Xi_t) \otimes (A + \Xi_t)\right] = A \otimes A + \mathrm{E}(\Xi_t \otimes \Xi_t) = \Omega_A$$

then

$$\mathrm{vec}E\left[(\sum_{j=0}^{m} B_j V_{t-j})(\sum_{j=0}^{m} B_j V_{t-j})^\top\right] = \sum_{j=0}^{m} \Omega_A^j \mathrm{vec}(\Sigma).$$

The critical condition for the stationarity of the process $\underline{y}_t$ is that $\sum_{j=0}^{m} \Omega_A^j$ converges as $m \to \infty$.

The matrix Ω_A may be represented in Jordan canonical form as $\Omega_A = P\Lambda P^{-1}$, where Λ has the eigenvalues of Ω_A along its main diagonal. If the eigenvalues of Ω_A have moduli less than unity, Λ^j converges to zero at a geometric rate. Notice that $\Omega_A^j = P\Lambda^j P^{-1}$, following a similar analysis as Nicholls and Quinn (1982, Chapter 2), $\underline{Y}_t$ (and thus y_t) is stationary and can be represented as

$$\underline{Y}_t = \sum_{j=0}^{\infty} B_j V_{t-j}.$$

The central limit theorem then follows from Billingsley (1961) (also see Nicholls and Quinn (1982, Theorem A.1.4)). ■

8.2. **Proof of Theorem 3.1.** If we denote $\widehat{v} = \sqrt{n}(\widehat{\theta}(\tau) - \theta(\tau))$, then $\rho_\tau(y_t - \widehat{\theta}(\tau)^\top x_t) = \rho_\tau(u_{t\tau} - (n^{-1/2}\widehat{v})^\top x_t)$, where $u_{t\tau} = y_t - x_t^\top \theta(\tau)$. Minimization of (8) is equivalent to minimizing:

$$Z_n(v) = \sum_{t=1}^{n} \left[\rho_\tau(u_{t\tau} - (n^{-1/2}v)^\top x_t) - \rho_\tau(u_{t\tau})\right]. \tag{18}$$

If $\widehat{v}$ is a minimizer of $Z_n(v)$, we have $\widehat{v} = \sqrt{n}(\widehat{\theta}(\tau) - \theta(\tau))$. The objective function $Z_n(v)$ is a convex random function. Knight (1989) (also see Pollard (1991) and Knight (1998)) shows that if the finite-dimensional distributions of $Z_n(\cdot)$ converge weakly to those of $Z(\cdot)$ and $Z(\cdot)$ has a unique minimum, the convexity of $Z_n(\cdot)$ implies that $\widehat{v}$ converges in distribution to the minimizer of $Z(\cdot)$.

We use the following identity: if we denote $\psi_\tau(u) = \tau - I(u < 0)$, for $u \neq 0$,

$$\begin{aligned} \rho_\tau(u - v) - \rho_\tau(u) &= -v\psi_\tau(u) + (u - v)\{I(0 > u > v) - I(0 < u < v)\} \\ &= -v\psi_\tau(u) + \int_0^v \{I(u \leq s) - I(u < 0)\} ds. \end{aligned} \tag{19}$$

Post-print version of an article published in Journal of the American Statistical Association 101(475): 980-990. doi:10.1198/016214506000000672.

Thus the objective function of minimization problem can be written as

$$\begin{aligned} & \sum_{t=1}^{n} \left[\rho_\tau(u_{t\tau} - (n^{-1/2}v)^\top x_t) - \rho_\tau(u_{t\tau}) \right] \\ = & -\sum_{t=1}^{n} (n^{-1/2}v)^\top x_t \psi_\tau(u_{t\tau}) + \sum_{t=1}^{n} \int_0^{(n^{-1/2}v)^\top x_t} \{I(u_{t\tau} \le s) - I(u_{t\tau} < 0)\} ds \end{aligned}$$

We first consider the limiting behavior of

$$W_n(v) = \sum_{t=1}^{n} \int_0^{(n^{-1/2}v)^\top x_t} \{I(u_{t\tau} \le s) - I(u_{t\tau} < 0)\} ds.$$

For convenience of asymptotic analysis, we denote

$$W_n(v) = \sum_{t=1}^{n} \xi_t(v), \ \xi_t(v) = \int_0^{(n^{-1/2}v)^\top x_t} \{I(u_{t\tau} \le s) - I(u_{t\tau} < 0)\} ds.$$

We further define $\overline{\xi}_t(v) = E\{\xi_t(v)|\mathcal{F}_{t-1}\}$, and $\overline{W}_n(v) = \sum_{t=1}^{n} \overline{\xi}_t(v)$, then $\{\xi_t(v) - \overline{\xi}_t(v)\}$ is a martingale difference sequence.

Notice that

$$u_{\tau t} = y_t - x_t^\top \alpha(\tau) = y_t - F_{t-1}^{-1}(\tau)$$

$$\begin{aligned} \overline{W}_n(v) &= \sum_{t=1}^{n} E\{\int_0^{(n^{-1/2}v)^\top x_t} [I(u_{t\tau} \le s) - I(u_{t\tau} < 0)] \,|\mathcal{F}_{t-1}\} \\ &= \sum_{t=1}^{n} \int_0^{(n^{-1/2}v)^\top x_t} \left[\int_{F_{t-1}^{-1}(\tau)}^{s+F_{t-1}^{-1}(\tau)} f_{t-1}(r) dr \right] ds \\ &= \sum_{t=1}^{n} \int_0^{(n^{-1/2}v)^\top x_t} \left[\frac{F_{t-1}(s + F_{t-1}^{-1}(\tau)) - F_{t-1}(F_{t-1}^{-1}(\tau))}{s} \right] s ds \end{aligned}$$

Under assumption A.3,

$$\begin{aligned} \overline{W}_n(v) &= \sum_{t=1}^{n} \int_0^{n^{-1/2}v^\top x_t} f_{t-1}(F_{t-1}^{-1}(\tau)) s ds + o_p(1) \\ &= \frac{1}{2n} \sum_{t=1}^{n} f_{t-1}(F_{t-1}^{-1}(\tau)) v^\top x_t x_t^\top v + o_p(1) \end{aligned}$$

By our assumptions and stationarity of y_t, we have

$$\overline{W}_n(v) \Rightarrow \frac{1}{2} v^\top \Omega_1 v$$

Post-print version of an article published in Journal of the American Statistical Association 101(475): 980-990. doi:10.1198/016214506000000672.

Using the same argument as Herce(1996), the limiting distribution of $\sum_t \xi_t(v)$ is the same as that of $\sum_t \overline{\xi}_t(v)$.

For the behavior of the first term, $n^{-1/2}\sum_{t=1}^n x_t\psi_\tau(u_{t\tau})$, in the objective function, notice that $x_t \in \mathcal{F}_{t-1}$ and $E[\psi_\tau(u_{t\tau})|\mathcal{F}_{t-1}] = 0$, $x_t\psi_\tau(u_{t\tau})$ is a martingale difference sequence and thus $n^{-1/2}\sum_{t=1}^n x_t\psi_\tau(u_{t\tau})$ satisfies a central limit theorem. Following the arguments of Portnoy (1984) and Gutenbrunner and Jurečková (1992), the autoregression quantile process is tight and thus the limiting variate viewed as a random function of τ, is a Brownian bridge over $\tau \in \mathcal{T}$,

$$n^{-1/2}\sum_{t=1}^n x_t\psi_\tau(u_{t\tau}) \Rightarrow \Omega_0^{1/2}B_k(\tau).$$

For each fixed τ, $n^{-1/2}\sum_{t=1}^n x_t\psi_\tau(u_{t\tau})$ converges to a q-dimensional vector normal variate with covariance matrix $\tau(1-\tau)\Omega_0$. Thus,

$$\begin{aligned}
& Z_n(v) \\
= \; & \sum_{t=1}^n \Big[\rho_\tau(u_{t\tau} - (n^{-1/2}v)^\top x_t) - \rho_\tau(u_{t\tau})\Big] \\
= \; & -\sum_{t=1}^n (n^{-1/2}v)^\top x_t\psi_\tau(u_{t\tau}) + \sum_{t=1}^n \int_0^{(n^{-1/2}v)^\top x_t} \{I(u_{t\tau} \le s) - I(u_{t\tau} < 0)\}ds. \\
\Rightarrow \; & -v^\top\Omega_0^{1/2}B_k(\tau) + \frac{1}{2}v^\top\Omega_1 v = Z(v)
\end{aligned}$$

By the convexity Lemma of Pollard (1991) and arguments of Knight (1989), notice that $Z_n(v)$ and $Z(v)$ are minimized at $\widehat{v} = \sqrt{n}(\widehat{\alpha}(\tau) - \alpha(\tau))$ and $\Sigma^{1/2}B_k(\tau)$ respectively, by Lemma A of Knight (1989) we have,

$$\Sigma^{-1/2}\sqrt{n}(\widehat{\alpha}(\tau) - \alpha(\tau)) \Rightarrow B_k(\tau).\blacksquare$$

REFERENCES

[1] Balke, N. and T. Fomby, 1997, Threshold Cointegration, *International Economic Review*, 38, 627-645.

[2] Bassett, G., and R. Koenker, 1982, An Empirical Quantile Function for linear models with iid errors, *Journal of the American Statistical Association*, Vol. 77, 407-415.

[3] Bassett, G., and R. Koenker and G. Kordas, 2004, Pessimistic Portfolio Allocation and Choquet Expected Utility, *Journal of Financial Econometrics*, 4, 477-492.

Post-print version of an article published in Journal of the American Statistical Association 101(475): 980-990. doi:10.1198/016214506000000672.

[4] Beaudry, P. and G. Koop, 1993, Do recessions permanently change output?, *Journal of Monetary Economics*, 31, 149-163.

[5] Billingsley, P, 1961, The Lindeberg-Levy Theorem for Martingales, *Proc. Amer. Math. Soc.*, 12, 788-792.

[6] Bradley, M.D. and D.W. Jansen, 1997, Nonlinear business cycle dynamics: Cross-country evidence on the persistence of aggregate shocks, *Economic Inquiry*, 35, 495-509.

[7] Brandt, A., 1986, The stochastic equation $Y_{n+1} = A_n Y_n + B_n$ with stationary coefficients, *Adv. Applied Probability*, 18, 211-220.

[8] Bofinger, E., 1975, Estimation of a density function using order statistics," *Australian Journal of Statistics*, 17, 1-7.

[9] Caner, M. and B. Hansen, 2001, Threshold Autoregression with a unit root, *Econometrica*, 69, 1555-1596.

[10] Cox, D., 1985, A Penalty Method for Nonparametric Estimation of the Logarithmic Derivative of a Density Function," *Annals of Institute of Mathematical Statistics*, 37, 271-288.

[11] Delong, J.B., and Summers, L.H., 1986, Are business cycle symmetrical?, in Gordon, R.J. (ed.), *American Business Cycle*, University of Chicago Press, Chicago.

[12] Denneberg, D. 1994, *Non-additive measure and integral*, Kluwer Academic Publishers.

[13] Enders, W. and C. Granger, 1998, Unit Root tests and asymetric adjustment with an example sing the term structure of interest rates, *Journal of Business and Economic Statistics*, 16, 304-311.

[14] Evans, M. and P. Wachtel, 1993, Inflation regioms and the sources of inflation uncertainty, *Journal of Money, Credit, and Banking*, 25, 475-511.

[15] Gonzalez, M. and J. Gonzalo, 1998, Threshold unit root models, Working paper, U. Carlos III de Madrid.

[16] Gutenbrunner, C., and Jurečková , J. (1992), "Regression Rank Scores and Regression Quantiles," *Annals of Statistics*, 20, 305-330.

[17] Hall, P., and S. Sheather, 1988, On the distribution of a studentized quantile," *J. Royal Statistical Soc. (B)*, 50, 381-391.

[18] Hamilton, J., 1989, A new approach to the economic analysis of nonstationary time series and the business cycle, *Econometrica*, 57, 357-384.

[19] Hansen, B., 2000, Sample splitting and Threshold estimation, *Econometrica*, 68, 575-603.

[20] Hess, G.D. and Iwata, 1997, Asymmetric persistence in GDP? A deeper look at depth, *Journal of Monetary Economics*, 40, 535-554.

[21] Hasan, M.N. and R. Koenker, 1997, Robust rank tests of the unit root hypothesis, *Econometrica* 65, 133-161.

Post-print version of an article published in Journal of the American Statistical Association 101(475): 980-990. doi:10.1198/016214506000000672.

[22] Hercé, M., 1996, Asymptotic Theory of LAD estimation in a unit root process with finite variance errors, *Econometric Theory*, 12, 129-153.

[23] Jurečková and Hallin, 1999, Optimal Tests for Autoregressive Models Based on Autoregression Rank Scores, *The Annals of Statistics*, 27, 1385–1414.

[24] Karlsen, H.A., 1990, Existence of Moments in a stationary stochastic difference equation, *Adv. Applied Probability*, 22, 129-146.

[25] Khmaladze, E., 1981, Martingale Approach to the goodness of fit tests, *Theory Probab. Appl.* 26, 246-265.

[26] Knight, K., 1989, Limit Theory for Autoregressive-parameter estimates in an infinite-variance random walk, The *Canadian Journal of Statistics* 17, 261-278.

[27] Knight, K., 1998, Asymptotics for L1 regression estimates under general conditions, *Annals of Statistics*, 26, 755-770.

[28] Koenker, R., 1994, Confidence Intervals for regression quantiles," in *Proc. of the 5th Prague Symp. on Asym. Stat.*, Springer-Verlag.

[29] Koenker, R. and G. Bassett, 1978, Regression Quantiles, *Econometrica*, 46, 33-49.

[30] Koenker, R. and J. Machado, 1999, Goodness of fit and relatedinference processes for quantile regression, *Journal of the American Statistical Association*, 81, 1296-1310.

[31] Koenker, R. 2000, Galton, Edgeworth, Frisch and prospects for quantile regression in econometrics, *J. of Econometrics*, 95, 347-374.

[32] Koenker, R. and Z. Xiao, 2002, Inference on the Quantile Regression Processes, *Econometrica*, 70, 1583-1612.

[33] Koenker, R. and Z. Xiao, 2004, Unit Root Quantile Regression Inference, *Journal of the American Statistical Association*, 99, 775-787.

[34] Koul, H., and A. K. Saleh, 1995, Autoregression quantiles and related rank-scores processes, *The Annals of Statistics*, 23, 670-689.

[35] Koul, H., and K. Mukherjee, 1994, Regression quantiles and related processes under long range dependent errors. *J. Mult. Analysis*, 51, 318-317.

[36] Kuan, C.M. and Y.L. Huang, 2001, The semi-nonstationary process: Model and empirical evidence, preprint.

[37] Neftci, S., 1984, Are economic time series asymmetric over the business cycle?, *Journal of Political Economy*, 92, 307-328.

[38] Nelson, C.R. and C.I. Plosser, 1982, Trends and random walks in macroeconomic time series: some evidence and implications, *Journal of Monetary Economics*, 10, 139-162.

[39] Ng, P., 1994, Smoothing Spline Score Estimation, *SIAM Journal of Scientific and Statistical Computing*, 15, 1003-1025.

[40] Nicholls, D.F., and B.G. Quinn, 1982, *Random Coefficient Autoregressive Models: An Introduction*, Springer-Verlag.

[41] Pollard, D., 1991, Asymptotics for Least Absolute Deviation Regression Estimators, *Econometric Theory*, 7, 186-199.

[42] Portnoy, S, (1984), Tightness of the Sequence of Empiric cdf Processes Defined from Regression Fractiles, in *Robust and Nonlinear Time Series Analysis*, eds. J. Franke, W. Hardle, and D. Martin, Springer-Verlag: New York.

[43] Portnoy, S., and R. Koenker, 1989, Adaptive L-estimation of linear models, *Annals of Statistics*, 17, 362-381.

[44] Pourahmadi, M., 1986, On stationarity of the solution of a doubly stochastic model, *J. of Time Series Analysis*, 7, 123-131.

[45] Powell, J., 1989, Estimation of monotonic regression models under quantile restrictions, in *Nonparametric and semiparametric methods in econometrics*, J. Powell, G. Tauchen, (ed.), Cambridge University Press.

[46] Rissanen, J., 1978, Modelling by shortest data description, *Automatica*, 14, 465-471.

[47] Schmeidler, D., 1986, Integral representation without additivity, *Proceedings of Amer. Math. Society*, 97, 255-261.

[48] Schwarz, G., 1978, Estimating the dimension of a model, *The Annals of Statistics*, 6, 461-464.

[49] Siddiqui, M., 1960, Distribution of Quantiles from a Bivariate Population, *Journal of Research of the National Bureau of Standards*, 64, 145–150.

[50] Silverman, B., 1986, *Density Estimation for statistics and data analysis*, Chapman and Hall, London.

[51] Tjøstheim, D., 1986, Some doubly stochastic time series models, *J. of Time Series Analysis*, 7, 51-72.

[52] Tong, H., 1990, *Nonlinear Time Series: A Dynamical Approach*, Oxford University Press.

[53] Tsay, R., 1997, Unit root tests with Threshold Innovations, preprint, University of Chicago.

[54] Weiss, A.,1987, Estimating Nonlinear Dynamic Models Using Least Absolute Error Estimation, *Econometric Theory*, 7, 46-68.