UNIT ROOT QUANTILE AUTOREGRESSION INFERENCE

ROGER KOENKER AND ZHIJIE XIAO

ABSTRACT. We study statistical inference in quantile autoregression models when the largest autoregressive coefficient may be unity. The limiting distribution of a quantile autoregression estimator and its t-statistic is derived. The asymptotic distribution is not the conventional Dickey-Fuller distribution, but a linear combination of the Dickey-Fuller distribution and the standard normal, with the weight determined by the correlation coefficient of related time series. Inference methods based on the estimator are investigated asymptotically. Monte Carlo results indicate that the new inference procedures have power gains over the conventional least squares based unit root tests in the presence of non-Gaussian disturbances. An empirical application of the model to US macroeconomic time series data further illustrates the potential of the new approach.

1. INTRODUCTION

An extensive literature in economics and finance suggests that many economic time-series are well characterized as autoregressive processes with a root near unity. Much of the formal inference apparatus used to investigate the so-called unit root hypothesis is, however, designed to provide optimal performance under Gaussian conditions. Under departures from the Gaussian model, particularly for innovation distributions with heavy tails, these methods can exhibit rather poor power performance. Since many applications, particularly in economics and finance, have notoriously heavy-tailed behavior it is important to consider estimation and inference procedures which are robust to departures from Gaussian conditions and are applicable to nonstationary time series.

One way to achieve robustness is the use of M estimation and associated inference apparatus. M-estimation methods for nonstationary time series with non-Gaussian innovations have been studied by Cox and Llatas (1991), Knight (1991), Phillips (1995), Lucas (1995), Rothenberg and Stock (1997), Juhl (1999), and Xiao (2001) among others. In particular, Cox and Llatas (1991) derive the asymptotic distribution of the M estimator for an AR(1) process with a (near) unit root. Knight (1991) studies unit root M estimation in the case with infinite variance errors. Lucas (1995) considers a unit root test based on a nonparametric modification of M estimators, focusing on the Huber and Student-t models.

Quantile regression methods provide an alternative approach for robust inference. Rather than relying exclusively on a single measure of conditional central tendency, the quantile regression approach allows the investigator to explore a range of conditional quantile functions thereby exposing a variety of forms of conditional heterogeneity. There is a considerable literature on quantile autoregression methods in time series including work by Weiss (1987), Knight (1989), Koul and Saleh(1995), Koul and Mukherjee(1994), Hercé (1996), Jurečková and Hallin (1999), and Rogers(2001). Hasan and Koenker (1997) consider rank-type tests based on regression rankscores in an augmented Dickey-Fuller framework.

In this paper we propose new tests of the unit root hypothesis based on the quantile autoregression approach. Both t-ratios based on estimates at selected quantiles and Kolmogorov-Smirnov (or

Cramer-von-Mises) type tests based on estimates over a range of quantiles are considered. Compared to existing procedures in the literature the new tests are targeted toward a somewhat broader class of alternatives including the random coefficient alternatives described in Section 4. The proposed tests have good power under non-Gaussian conditions and sacrifice little efficiency at the Gaussian model. The tests also have good power in the presence of asymmetric dynamics, comparing to the existing tests.

We introduce the model and estimation in Section 2. After proposing the tests and describing their asymptotic behavior in Section 3, we illustrate their performance with a small Monte Carlo experiment in Section 4. An application to U.S. macroeconomic data is given in Section 5.

A few words on notation: we use x_t to denote the vector of regressors which vary with the order of the autoregression, the symbol "$\Rightarrow$" indicates weak convergence of the associated probability measures, $[nr]$ denotes the integer part of nr, and $:=$ is used to signify definitional equivalence. Continuous stochastic processes such as the Brownian motion $B(r)$ on $[0,1]$ are usually written simply as B, and integrals with respect to the Lebesgue measure such as $\int_0^1 B(r)dr$ are simply written as $\int B$.

2. QUANTILE AUTOREGRESSION WITH A UNIT ROOT

2.1. The QAR(1) Model.
We first consider the following autoregression model

$$y_t = \alpha y_{t-1} + u_t, \ t = 1, \cdots, n. \tag{1}$$

focusing on the case that $\alpha = 1$. For simplicity and without essential loss of generality, we focus much of our attention on the first order autoregression in this section, but our analysis is easily extended to the general case – see the discussion in Section 2.3 for extension to the AR(p) model. For results on unit root estimation and testing based on least-squares methods, see, e.g., Dickey and Fuller (1979), Chan and Wei (1987).

If we denote the τ-th quantile of u_t as $Q_u(\tau)$ and let $Q_{y_t}(\tau|y_{t-1})$ denote the τ-th conditional quantile of y_t conditional on y_{t-1}, then

$$Q_{y_t}(\tau|y_{t-1}) = Q_u(\tau) + \alpha y_{t-1}.$$

Let $\alpha_0(\tau) = Q_u(\tau)$, $\alpha_1(\tau) = \alpha$, and define $\alpha(\tau) = (\alpha_0(\tau), \alpha_1(\tau))^\top$, $x_t = (1, y_{t-1})^\top$, we have

$$Q_{y_t}(\tau|y_{t-1}) = x_t^\top \alpha(\tau). \tag{2}$$

In this model the τth conditional quantile function of the response y_t is expressed as a linear function of lagged values of the response. We will explore estimation and inference in the above quantile autoregression (QAR) model in the presence of a unit root.

Estimation of the linear quantile autoregressive model involves solving the problem

$$\min_{\alpha \in \mathbb{R}^2} \sum_{t=1}^n \rho_\tau(y_t - x_t^\top \alpha), \tag{3}$$

where $\rho_\tau(u) = u(\tau - I(u < 0))$ as in Koenker and Bassett (1978). Solutions of (3), $\widehat{\alpha}(\tau)$, will be called τ-th autoregression quantiles; viewed as a function of τ we will refer to $\widehat{\alpha}(\tau)$ as the QAR(1) process. Given $\widehat{\alpha}(\tau)$, the τ-th conditional quantile function of y_t, conditional on the past information, can be estimated by,

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

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

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

for some appropriately chosen sequence of τ's.

2.2. **Limiting Distribution of the QAR(1) Process.** In this section, we describe the limiting behavior of the autoregression quantile process under the unit root hypothesis. Our analysis follows the approach of Knight (1991). See Herce (1996), Hasan and Koenker (1997), and Hasan (2001) for related results.

As will become clear in our later analysis, due to the nonstationarity of y_t, the two components in $\widehat{\alpha}(\tau) = (\widehat{\alpha}_0(\tau), \widehat{\alpha}_1(\tau))$ have different rates of convergence. In particular, $\widehat{\alpha}_1(\tau)$ converges to unity at rate n, while $\widehat{\alpha}_0(\tau)$ converges at rate $\sqrt{n}$. For this reason, we introduce the standardization matrix $D_n = diag(\sqrt{n}, n)$ and denote $\widehat{v} = D_n(\widehat{\alpha}(\tau) - \alpha(\tau))$, and write $\rho_\tau(y_t - \widehat{\alpha}(\tau)^\top x_t)$ as $\rho_\tau(u_{t\tau} - (D_n^{-1}\widehat{v})^\top x_t)$, where $u_{t\tau} = y_t - x_t'\alpha(\tau)$. Minimization of (3) is equivalent to the following problem:

$$\min_v \sum_{t=1}^n \left[\rho_\tau(u_{t\tau} - (D_n^{-1}v)^\top x_t) - \rho_\tau(u_{t\tau}) \right]. \tag{4}$$

If $\widehat{v}$ is a minimizer of $Z_n(v) = \sum_{t=1}^n \left[\rho_\tau(u_{t\tau} - (D_n^{-1}v)^\top x_t) - \rho_\tau(u_{t\tau}) \right]$, we have $\widehat{v} = D_n(\widehat{\alpha}(\tau) - \alpha(\tau))$.

The objective function $Z_n(v)$ is a convex random function. Knight (1989, 1991) and Pollard (1991) show 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 $Z(\cdot)$.

Our asymptotic analysis is based on the following assumptions.

> **Assumption A1:** $\{u_t\}$ *are i.i.d. random variables with mean zero and variance* $\sigma^2 < \infty$.
> **Assumption A2:** *The distribution function of* $\{u_t\}$, F, *has a continuous Lebesgue density,* f, *with* $0 < f(u) < \infty$ *on* $\{u : 0 < F(u) < 1\}$.

Denoting $\psi_\tau(u) = \tau - I(u < 0)$, by definition of $u_{t\tau}$, we have $E[\psi_\tau(u_{t\tau})|\mathcal{F}_{t-1}] = 0$. The asymptotic distribution of the autoregression quantile is closely related to the asymptotic behavior of $n^{-1} \sum_{t=1}^n y_{t-1}\psi_\tau(u_{t\tau})$. Note that both u_t and $\psi_\tau(u_{t\tau})$ have mean 0, and are correlated. Under Assumption A1, the partial sums of the vector process $(u_t, \psi_\tau(u_{t\tau}))$ follow a bivariate invariance principle (see, e.g., Phillips and Durlauf (1986, Theorem 2.1, 474-476, and 486-489); Wooldridge and White (1988, Corollary 4.2); and Hansen 1992):

$$n^{-1/2} \sum_{t=1}^{[nr]} (u_t, \psi_\tau(u_{t\tau}))^\top \Rightarrow (B_u(r), B_\psi^\tau(r))^\top = BM(0, \Sigma(\tau))$$

, where $\Sigma(\tau) = E[(u_t, \psi_\tau(u_{t\tau}))^\top(u_t, \psi_\tau(u_{t\tau}))]$ is the covariance matrix of the bivariate Brownian motion. Consequently, it is easy to verify (e.g. Phillips and Durlauf (1986, Lemma 3.1); and Hansen 1992)) that

$$n^{-1} \sum_{t=1}^n y_{t-1}\psi_\tau(u_{t\tau}) \Rightarrow \int_0^1 B_u dB_\psi^\tau.$$

The random function $n^{-1/2} \sum_{t=1}^{[nr]} \psi_\tau(u_{t\tau})$ converges to a two parameter process $B_\psi^\tau(r) = B_\psi(\tau, r)$. Following the arguments of Portnoy (1984) and Gutenbrunner and Jurečková (1994) it can be shown that the autoregression quantile process is tight and thus the limiting variate $B_\psi^\tau(r)$, viewed as a random function of τ, is a Brownian bridge over $\tau \in [0, 1]$. Thus, the two parameter process $B_\psi^\tau(r)$ is partially Brownian motion and partially Brownian bridge in the sense that for fixed r, $B_\psi^\tau(r) = B_\psi(\tau, r)$ is a rescaled Brownian bridge, while for each τ, $n^{-1/2} \sum_{t=1}^{[nr]} \psi_\tau(u_{t\tau})$ converges weakly to a Brownian motion with variance $\tau(1 - \tau)$. Thus, for each fixed pair (τ, r), $B_\psi^\tau(r) = B_\psi(\tau, r) \sim N(0, \tau(1 - \tau)r)$.

Using the identity (18) given in the Appendix, the objective function of minimization problem (4) can be written as

$$\sum_{t=1}^{n} \left[\rho_\tau(u_{t\tau} - (D_n^{-1}v)^\top x_t) - \rho_\tau(u_{t\tau}) \right]$$

$$= -\sum_{t=1}^{n}(D_n^{-1}v)^\top x_t \psi_\tau(u_{t\tau}) + \sum_{t=1}^{n} \int_0^{(D_n^{-1}v)^\top x_t} \{I(u_{t\tau} \leq s) - I(u_{t\tau} < 0)\} ds.$$

The following Lemma gives asymptotic results that are useful in deriving the limiting distribution of $D_n(\widehat{\alpha}(\tau) - \alpha(\tau))$.

Lemma 2.1 *Let y_t be determined by (1) with $\alpha = 1$, under Assumptions A1-A2,*

$$D_n^{-1} \sum_{t=1}^{n} x_t \psi_\tau(u_{t\tau}) \Rightarrow \int_0^1 \overline{B}_u dB_\psi^\tau, \tag{5}$$

$$\sum_{t=1}^{n} \int_0^{(D_n^{-1}v)^\top x_t} \{I(u_{t\tau} \leq s) - I(u_{t\tau} < 0)\} ds \Rightarrow \frac{1}{2} f(F^{-1}(\tau)) v^\top \left[\int_0^1 \overline{B}_u \overline{B}_u^\top \right] v,$$

where $\overline{B}_u(r) = [1, B_u(r)]^\top$.

The limiting distribution (5) can be written as $(\int_0^1 dB_\psi^\tau, \int_0^1 B_u dB_\psi^\tau)$ where the first component, $\int_0^1 dB_\psi^\tau$, is simply $N(0, \tau(1-\tau))$. The limiting distribution of the QAR estimator for the unit root model is summarized in the following Theorem.

Theorem 2.1 *If y_t is determined by (1) with $\alpha = 1$, under Assumptions A1-A2,*

$$D_n(\widehat{\alpha}(\tau) - \alpha(\tau)) \Rightarrow \frac{1}{f(F^{-1}(\tau))} \left[\int_0^1 \overline{B}_u \overline{B}_u^\top \right]^{-1} \int_0^1 \overline{B}_u dB_\psi^\tau.$$

As an immediate consequence of the above Theorem, we have the following corollary which is useful for construction of tests of the unit root hypothesis.

Corollary 2.1 *Under the assumptions of Theorem 2.1,*

$$n(\widehat{\alpha}_1(\tau) - 1) \Rightarrow \frac{1}{f(F^{-1}(\tau))} \left[\int_0^1 \underline{B}_u^2 \right]^{-1} \int_0^1 \underline{B}_u dB_\psi^\tau,$$

where $\underline{B}_u(r) = B_u(r) - \int_0^1 B_u$ is the demeaned version of Brownian motion B_u.

2.3. **Higher Order QAR Models.** One of the most important extensions of the first order autoregression formulation of the unit root model is the augmented Dickey-Fuller (1979) (ADF) regression model

$$y_t = \alpha_1 y_{t-1} + \sum_{j=1}^{q} \alpha_{j+1} \Delta y_{t-j} + u_t. \tag{6}$$

In this model, the autoregressive coefficient α_1 plays an important role in measuring persistency in economic and financial time series. Under regularity conditions, if $\alpha_1 = 1$, y_t contains a unit root and is persistent; and if $|\alpha_1| < 1$, y_t is stationary. Denoting the σ-field generated by $\{u_s, s \leq t\}$ by $\mathcal{F}_t$, the τ-th conditional quantile of y_t, conditional on $\mathcal{F}_{t-1}$, is given by

$$Q_{y_t}(\tau|\mathcal{F}_{t-1}) = Q_u(\tau) + \alpha_1 y_{t-1} + \sum_{j=1}^{q} \alpha_{j+1} \Delta y_{t-j}.$$

Let $\alpha_0(\tau) = Q_u(\tau))$, and $\alpha_j(\tau) = \alpha_j$, $j = 1, ..., q + 1$, and define

$$\alpha(\tau) = (\alpha_0(\tau), \alpha_1, \cdots, \alpha_{q+1}), \ x_t = (1, y_{t-1}, \Delta y_{t-1}, \cdots, \Delta y_{t-q})',$$

we have

$$Q_{y_t}(\tau|\mathcal{F}_{t-1}) = x_t'\alpha(\tau). \tag{7}$$

Again the τth conditional quantile function of the response y_t is expressed as a linear function of lagged values of the response.

Let $\widehat{\alpha}(\tau) = (\widehat{\alpha}_0(\tau), \widehat{\alpha}_1, \cdots, \widehat{\alpha}_p)$, $p = q + 1$, and $D_n = diag(\sqrt{n}, n, \sqrt{n}, \cdots, \sqrt{n})$, then the analysis of $\widehat{\alpha}(\tau)$ follows a similar procedure to that of first order quantile autoregression with a unit root. We replace Assumption A1 by the following modification.

Assumption A1': *The roots of $A(L) = 1 - \sum_{j=1}^q \alpha_{j+1}L^j$ all lie outside the unit circle, and $\{u_t\}$ are i.i.d. random variables with mean zero and variance $\sigma^2 < \infty$.*

Denote $w_t = \Delta y_t$, then, under the unit root hypothesis and Assumption A1'

$$n^{-1/2}\sum_{t=1}^{[nr]}(w_t, \psi_\tau(u_{t\tau}))^\top \Rightarrow (B_w(r), B_\psi^\tau(r))^\top = BM(0, \underline{\Sigma}(\tau)),$$

where

$$\underline{\Sigma}(\tau) = \left[\begin{array}{cc} \sigma_w^2 & \sigma_{w\psi}(\tau) \\ \sigma_{w\psi}(\tau) & \sigma_\psi^2(\tau) \end{array} \right]$$

is the long run covariance matrix of the bivariate Brownian motion and can be written as $\Sigma_0(\tau) + \Sigma_1(\tau) + \Sigma_1^\top(\tau)$, where $\Sigma_0(\tau) = E[(w_t, \psi_\tau(u_{t\tau}))^\top(w_t, \psi_\tau(u_{t\tau}))]$ and

$$\Sigma_1(\tau) = \sum_{s=2}^\infty E[(w_1, \psi_\tau(u_{1\tau}))^\top(w_s, \psi_\tau(u_{s\tau}))].$$

We summarize the limiting distribution of $\widehat{\alpha}(\tau)$ in the following Theorem.

Theorem 2.2 *Let y_t be determined by (6), under Assumptions A1', A2, and the unit root assumption $\alpha_1 = 1$,*

$$D_n(\widehat{\alpha}(\tau) - \alpha(\tau)) \Rightarrow \frac{1}{f(F^{-1}(\tau))} \left[\begin{array}{cc} \int_0^1 \overline{B}_w\overline{B}_w^\top & 0_{2\times q} \\ 0_{q\times 2} & \Omega_\Phi \end{array} \right]^{-1} \left[\begin{array}{c} \int_0^1 \overline{B}_w dB_\psi^\tau \\ \Phi \end{array} \right],$$

where $\overline{B}_w(r) = [1, B_w(r)]^\top$, $\Phi = [\Phi_1, \cdots, \Phi_q]^\top$ is a q-dimensional normal variate with covariance matrix $\tau(1 - \tau)\Omega_\Phi$ where

$$\Omega_\Phi = \left[\begin{array}{ccc} \nu_0 & \cdots & \nu_{q-1} \\ \vdots & \ddots & \vdots \\ \nu_{q-1} & \cdots & \nu_0 \end{array} \right], \ \nu_j = E[w_t w_{t-j}],$$

and Φ is independent with $\int_0^1 \overline{B}_w dB_\psi^\tau$.

Remark 2.1 As an immediate by-product of Theorem 2.2, the limiting distribution of $n(\widehat{\alpha}_1(\tau) - 1)$ is invariant to the estimation of $\widehat{\alpha}_j(\tau)(j = 2, ...p)$ and the lag length p, which is a result similar to the conventional ADF regression.

Corollary 2.2 *Under the assumptions of Theorem 2.2,*

$$\left[\begin{array}{c} \sqrt{n}(\widehat{\alpha}_0(\tau) - \alpha_0(\tau)) \\ n(\widehat{\alpha}_1(\tau) - 1) \end{array} \right] \Rightarrow \frac{1}{f(F^{-1}(\tau))} \left[\int_0^1 \overline{B}_w \overline{B}_w^\top \right]^{-1} \int_0^1 \overline{B}_w dB_\psi^\tau.$$

In particular,

$$n(\widehat{\alpha}_1(\tau) - 1) \Rightarrow \frac{1}{f(F^{-1}(\tau))} \left[\int_0^1 \underline{B}_w^2 \right]^{-1} \int_0^1 \underline{B}_w dB_\psi^\tau, \tag{8}$$

where $\underline{B}_w(r) = B_w(r) - \int_0^1 B_w$ is the corresponding demeaned Brownian motion.

3. INFERENCE ON THE QAR PROCESS

Inference based on the autoregression quantile process provides a more robust approach to testing the unit root hypothesis. Like the conventional augmented Dickey-Fuller (ADF) t-ratio test, we consider the t-ratio statistic

$$t_n(\tau) = \frac{f(\widehat{F^{-1}(\tau)})}{\sqrt{\tau(1-\tau)}} \left(Y_{-1}^\top P_X Y_{-1} \right)^{1/2} (\widehat{\alpha}_1(\tau) - 1),$$

where $f(\widehat{F^{-1}(\tau)})$ is a consistent estimator of $f(F^{-1}(\tau))$, Y_{-1} is the vector of lagged dependent variables (y_{t-1}) and P_X is the projection matrix onto the space orthogonal to $X = (1, \Delta y_{t-1}, \cdots, \Delta y_{t-q})$. Under the unit root hypothesis, we have using the results in the previous section,

$$t_n(\tau) \Rightarrow t(\tau) = \frac{1}{\sqrt{\tau(1-\tau)}} \left[\int_0^1 \underline{B}_w^2 \right]^{-1/2} \int_0^1 \underline{B}_w dB_\psi^\tau. \tag{9}$$

At any fixed τ the test statistic $t_n(\tau)$ is simply the quantile regression counterpart of the well-known ADF t-ratio test for a unit root. The limiting distribution of $t_n(\tau)$ is nonstandard and depends on nuisance parameters $(\sigma_w^2, \sigma_{w\psi}(\tau))$ since B_w and B_ψ^τ are correlated Brownian motions. The above limiting distribution is similar to distributions appearing in various unit root tests using other methods. In particular, similar limiting distributions arise in Lucas (1995) for unit root tests based on his nonparametric modified M-estimators, in Hasan and Koenker (1997) for their unmodified statistic S_T based on rank scores, and, as we show below, in Hansen (1995) for his least-sqares based covariate augmented Dickey-Fuller test. In this section, we consider two options to facilitate inference based on the QAR processes.

The limiting distribution of $t_n(\tau)$ can be decomposed as a linear combination of two (independent) distributions, with weights determined by a long-run (zero frequency) correlation coefficient that can be consistently estimated. Consequently, the limiting distribution can be easily approximated using simulation methods. In fact, required critical values are already tabulated in the literature and thus are available for use in applications. This decomposition facilitates our first approach of unit root test. In the second approach, we abandon the asymptotically distribution free nature of tests and use critical values generated by resampling methods. We explore both approaches[1] in the following analysis.

Unit root tests may be constructed based on a quantile autoregression at some selected representative quantiles (say, median, lower quartile, upper quartile, or deciles). Alternatively, we could examine the unit root property over a range of quantiles $\tau \in \mathcal{T}$. We first consider testing procedures based on quantile regression at a selected representative quantile.

---

[1]A third approach is to construct a transformation of the original statistic $t_n(\tau)$ that annihilates the nuisance parameter, and thereby provides a distributional-free form of inference In the presence of Gaussian innovations, performance of this "fully-modified" test is not as good as the tests based on the unmodified statistic. For this reason, we focus our attention on the two procedures proposed in this section.

x

3.1. **Decomposing the Limiting Distribution of $t_n(\tau)$.** Following Phillips and Hansen (1990) and Phillips (1995)) we have the decomposition,

$$\int_0^1 \underline{B}_w \, dB_\psi^\tau = \int \underline{B}_w \, dB_{\psi.w}^\tau + \lambda_{\omega\psi}(\tau) \int \underline{B}_w \, dB_w,$$

where $\lambda_{\omega\psi}(\tau) = \sigma_{w\psi}(\tau)/\sigma_w^2$ and $B_{\psi.w}^\tau$ is a Brownian motion with variance

$$\sigma_{\psi.w}^2(\tau) = \sigma_\psi^2(\tau) - \sigma_{w\psi}^2(\tau)/\sigma_w^2$$

and is independent of $\underline{B}_w$. The limiting distribution of $t_n(\tau)$ can therefore be decomposed as

$$\frac{1}{\sqrt{\tau(1-\tau)}} \frac{\int \underline{B}_w \, dB_{\psi.w}^\tau}{\left(\int_0^1 \underline{B}_w^2\right)^{1/2}} + \frac{\lambda_{w\psi}(\tau)}{\sqrt{\tau(1-\tau)}} \frac{\int \underline{B}_w \, dB_w}{\left(\int_0^1 \underline{B}_w^2\right)^{1/2}}.$$

For convenience of exposition, we may re-write the Brownian motions $B_w(r)$ and $B_{\psi.w}^\tau(r)$ as

$$B_w(r) = \sigma_w W_1(r), \ B_{\psi.w}^\tau(r) = \sigma_{\psi.w}(\tau) W_2(r),$$

$$\underline{B}_w(r) = \sigma_w \underline{W}_1(r), \ \underline{W}_1(r) = W_1(r) - \int_0^1 W_1(s) \, ds,$$

where $W_1(r)$ and $W_2(r)$ are standard Brownian motions and are independent of one another. Note that $\sigma_\psi^2(\tau) = \tau(1-\tau)$, so it is easy to show that the limiting distribution of $t_n(\tau)$ can be written as,

$$\delta \left(\int_0^1 \underline{W}_1^2\right)^{-1/2} \int_0^1 \underline{W}_1 \, dW_1 + \sqrt{1-\delta^2} N(0,1), \tag{10}$$

where

$$\delta = \delta(\tau) = \frac{\sigma_{w\psi}(\tau)}{\sigma_w \sigma_\psi(\tau)} = \frac{\sigma_{w\psi}(\tau)}{\sigma_w \sqrt{\tau(1-\tau)}}.$$

Thus we conclude that the limiting distribution of the $t_n(\tau)$ is a mixture of a Dickey-Fuller component,

$$\left(\int_0^1 \underline{W}_1^2\right)^{-1/2} \int_0^1 \underline{W}_1 \, dW_1,$$

and a standard normal component (which is independent of the Dickey-Fuller component), with the weights determined by the parameter δ. Notice that σ_w^2 is the long-run (zero frequency) variance of $\{w_t\}$, $\sigma_\psi^2(\tau)$ is the long-run variance of $\{\psi_\tau(u_{t\tau})\}$, and $\sigma_{w\psi}(\tau)$ is the long-run covariance of $\{w_t\}$ and $\{\psi_\tau(u_{t\tau})\}$, thus $\delta = \delta(\tau)$ is simply the long-run correlation coefficient between $\{w_t\}$ and $\{\psi_\tau(u_{t\tau})\}$.

Given a consistent estimate of δ, the limiting distribution of $t_n(\tau)$ can be approximated by a direct simulation. The limiting distribution (10) is the same as that of the covariate-augmented Dickey-Fuller (CADF) test of Hansen (1995). Tables of critical values for the statistic $t_n(\tau)$ are provided by Hansen (1995, page 1155) for values of δ^2 in steps of 0.1. For intermediate values of δ^2, Hansen suggest using critical values obtained by interpolation[2]. For convenience, we give the table of critical values from Hansen in the Appendix as Table A. In practice, to use the correct critical values from Table A, we estimate δ^2 by $\widehat{\delta^2} = \widehat{\sigma}_{w\psi}^2(\tau)/[\tau(1-\tau)\widehat{\sigma}_w^2]$ and then use the estimated $\widehat{\delta^2}$ to select the appropriate row from the Table. See Section 3.4 for further details.

---

[2]An alternative approach would be to fit a polynomial in δ^2 to certain order and approximate the critical values for any δ^2 by the fitted regression.

3.2. **Calculating Critical Values Using Resampling.** The second approach is to generate critical values for the unmodified statistics using resampling methods. We use the usual notation $*$ to signify the bootstrap samples and use P^* to denote the probability conditional on the original sample. We may consider the following resampling procedure.

(1) First, let $w_t = \Delta y_t$, $(t = 2, ..., n)$, we fit the following q-th order autoregression[3] by OLS:

$$w_t = \sum_{j=1}^q \widehat{\beta}_j w_{t-j} + \widehat{u}_t,\ t = q+1, ..., n,$$

and obtain estimates $\widehat{\beta}_1, \cdots \cdots, \widehat{\beta}_q$, and the residuals $\widehat{u}_t$.

(2) Draw i.i.d. variables $\{u_t^*\}_{t=q+1}^n$ from the centered residuals $\widehat{u}_t - \frac{1}{n-q}\sum_{j=q+1}^n \widehat{u}_j$ and generate w_t^* from u_t^* using the fitted autoregression:

$$w_t^* = \sum_{j=1}^q \widehat{\beta}_j w_{t-j}^* + u_t^*,\ t = q+1, ..., n,$$

with $w_j^* = \Delta y_j$ for $j = 1, ..., q$.

(3) Then we can generate y_t^* under the null restriction of a unit root: $y_t^* = y_{t-1}^* + w_t^*$, with $y_1^* = y_1$.

(4) Finally, we estimate the following p-th order autoregressive quantile regression

$$y_t^* = \alpha_0 + \alpha_1 y_{t-1}^* + \sum_{j=1}^q \alpha_{j+1}\Delta y_{t-j}^* + u_t. \tag{11}$$

and denote the estimator of $\alpha_1(\tau)$ by $\widehat{\alpha}_1^*(\tau)$. Corresponding to $t_n(\tau)$, we construct

$$t_n^*(\tau) = \frac{f(\widehat{F^{-1}(\tau)})}{\sqrt{\tau(1-\tau)}} \left(Y_{-1}^{*\top} P_X^* Y_{-1}^*\right)^{1/2} \left(\widehat{\alpha}_1^*(\tau) - 1\right).$$

In the above procedure, we generate y_t^* under the null hypothesis of unit root to ensure the nonstationarity of the generated sample $\{y_t^*\}$ and thus make the subsequent bootstrap test valid. The asymptotic validity of the bootstrap procedure relies on a bootstrap invariance principle, i.e., the weak convergence of the bootstrap partial sum process to Brownian motion that holds almost surely for all sample realizations. Park (2002) establishes a bootstrap invariance principle for sieve bootstrap that allows the AR lag length q to go to infinity. In a subsequent paper, Chang, Park and Song (2002) develop a multivariate bootstrap invariance principle. The theory they derive in the vector time series model can be used here to derive a bivariate bootstrap invariance principle that validates the above resampling procedure. (For other versions of bootstrap invariance principles, see, e.g., Kinateder (1992), Ferretti and Romo (1996), and Giersbergen (1996).) The limiting null distribution of the test statistics can then be approximated by repeating steps 2-4 many times. Let $C_t^*(\tau, \theta)$ be the (100θ)-th quantiles, i.e.,

$$\mathrm{P}^*\left[t_n^*(\tau) \le C_t^*(\tau, \theta)\right] = \theta,$$

then the unit root hypothesis will be rejected at the $(1 - \theta)$ level if $t_n(\tau) \le C_t^*(\tau, \theta)$.

Alternatively, instead of using resampling methods, we may directly simulate the Brownian motions. In particular, we may replace step 4 of quantile regression (11) by directly approximating

---

[3]We may also use the Yule-Walker method, which is asymptotically equivalent to the OLS method, to estimate the autoregression.

$\int_0^1 \underline{B}_w^2$ and $\int_0^1 \underline{B}_w dB_\psi^\tau$ using

$$\frac{1}{n^2} \sum_t (y_t^* - \overline{y}^*)^2 \text{ and } \frac{1}{n} \sum_t (y_t^* - \overline{y}^*) \psi_\tau(u_{t\tau}^*) ,$$

where $\overline{y}^* = n^{-1} \sum y_t^*$, and $u_{t\tau}^* = u_t^* - \widetilde{F}_u^{-1}(\tau)$, where $\widetilde{F}_u^{-1}(\tau)$ is the quantile function of u_t^*. Thus, the limiting null distribution of $t_n(\tau)$ can be approximated based on the following quantities

$$\frac{1}{\sqrt{\tau(1-\tau)}} \left[\sum_t (y_t^* - \overline{y}^*)^2 \right]^{-1/2} \left[\sum_t (y_t^* - \overline{y}^*) \psi_\tau(u_{t\tau}^*) \right].$$

Since we simply calculate sample moment and avoid solving the linear programming in each repetition in this alternative procedure, computationally this is faster.

3.3. **Other Tests.** In addition to the t-ratio statistic $t_n(\tau)$, just as the ADF coefficient test and the Phillips-Perron Z_α test, we may also use the coefficient-based statistic in the QAR model for unit root testing. We may define the following coefficient-based statistic

$$U_n(\tau) = n(\widehat{\alpha}_1(\tau) - 1).$$

Under the unit root hypothesis and our assumptions,

$$U_n(\tau) \Rightarrow U(\tau) = \frac{1}{f(F^{-1}(\tau))} \left[\int_0^1 \underline{B}_w^2 \right]^{-1} \int_0^1 \underline{B}_w dB_\psi^\tau. \tag{12}$$

At fixed τ the test statistic $U_n(\tau)$ is the quantile regression counterpart of the coefficient based ADF test. Like the t-ratio statistic, the limiting distribution of $U_n(\tau)$ is not standard and depend on nuisance parameters. We can consider similar options as those we used for the t-statistic. For example, notice that $\underline{B}_w$ and B_ψ^τ are Brownian motions and can be approximated by sums of Gaussian random variables, thus the distributions of the limiting variates $\left[\int_0^1 \underline{B}_w^2 \right]^{-1} \int_0^1 \underline{B}_w dB_\psi^\tau$ may be approximated by a direct simulation or resampling.

Another approach to test the unit root property is to examine the unit root property over a range of quantiles $\tau \in \mathcal{T}$, instead of focusing only on a selected quantile. For example, we may construct Kolmogorov-Smirnov (KS) or Cramer-von-Mises (CM) type tests based on the regression quantile process for $\tau \in \mathcal{T}$.

Consider $\tau \in \mathcal{T} = [\tau_0, 1 - \tau_0]$ for some small $\tau_0 > 0$, we propose the following quantile regression-based statistics for testing the null hypothesis of a unit root:

$$QKS_\alpha = \sup_{\tau \in \mathcal{T}} |U_n(\tau)| , \ QKS_t = \sup_{\tau \in \mathcal{T}} |t_n(\tau)| , \tag{13}$$

and

$$QCM_\alpha = \int_{\tau \in \mathcal{T}} U_n(\tau)^2 d\tau, \ QCM_t = \int_{\tau \in \mathcal{T}} t_n(\tau)^2 d\tau. \tag{14}$$

In practice, we may calculate $U_n(\tau)$ and $t_n(\tau)$ at, say, $\{\tau_i = i/n\}_{i=1}^n$, and thus the statistics QKS_α and QKS_t can be constructed by taking maximum over $\tau_i \in \mathcal{T}$ and QCM_α and QCM_t are obtained using numerical integration.

The limiting distributions of these tests are given by $\sup_{\tau \in \mathcal{T}} |U(\tau)| \sup_{\tau \in \mathcal{T}} |t(\tau)|$, $\int_{\tau \in \mathcal{T}} U(\tau)^2 d\tau$, and $\int_{\tau \in \mathcal{T}} t(\tau)^2 d\tau$ respectively. Again, we may approximate these limiting distributions by direct simulation or resampling methods. To resample the limiting distributions we follow steps (1), (2), and (3) in the above procedure, and replace step (4) by

(4′) We estimate the p-th order autoregressive quantile regression (11) at, say, $\{\tau_i \in I\}_{i=1}^n$, where $I = \{i : \tau_i = i/n, \text{ and } \tau_i \in \mathcal{T}\}$, denote the τ_i-th quantile autoregression estimator of α_1 by $\widehat{\alpha}_1^*(\tau_i)$, and calculate QKS_α^*, QKS_t^*, QCM_α^*, and QCM_t^* based on $U_n^*(\tau_i) = n(\widehat{\alpha}_1^*(\tau_i) - 1)$ and $t_n^*(\tau_i)$:

$$QKS_t^* = \max_{i \in I} |t_n^*(\tau_i)|, QCM_t^* = \sum_{i \in I} t_n^*(\tau_i)^2 \left(\tau_i - \tau_{i-1}\right),$$

$$QKS_\alpha^* = \max_{i \in I} |U_n^*(\tau_i)|, QCM_\alpha^* = \sum_{i \in I} U_n^*(\tau_i)^2 \left(\tau_i - \tau_{i-1}\right).$$

The limiting null distribution of the test statistics can again be approximated by repeating steps 2-4′. Let $C_{KS_\alpha}(\theta)$, $C_{KS_t}(\theta)$, $C_{CM_\alpha}(\theta)$ and $C_{CM_t}(\theta)$ be the (100θ)-th quantiles, i.e.,

$$\mathrm{P}^* \left[QKS_\alpha^* \le C_{KS_\alpha}(\theta)\right] = \mathrm{P}^* \left[QKS_t^* \le C_{KS_t}(\theta)\right] = \theta,$$

$$\mathrm{P}^* \left[QCM_\alpha^* \le C_{CM_\alpha}(\theta)\right] = \mathrm{P}^* \left[QCM_t^* \le C_{CM_t}(\theta)\right] = \theta,$$

then the unit root hypothesis will be rejected at the $(1 - \theta)$ level if, say, $QKS_\alpha > C_{KS_\alpha}(\theta)$.

3.4. **Estimation of Nuisance Parameters.** Our proposed tests require estimates of the quantile density function $f(F^{-1}(\tau))$ and the variance and covariance parameters σ_w^2, $\sigma_{w\psi}(\tau)$ and λ_w. There is a large related literature on estimating $f(F^{-1}(\tau))$, including, e.g. Siddiqui (1960), Bofinger (1975). Following Siddiqui(1960), and noting that, $dF^{-1}(t)/dt = (f(F^{-1}(t)))^{-1}$, it is natural to use the estimator,

$$f_n(F_n^{-1}(t)) = \frac{2h_n}{F_n^{-1}(t + h_n) - F_n^{-1}(t - h_n)}, \tag{15}$$

where $F_n^{-1}(s)$ is an estimate of $F^{-1}(s)$ and h_n is a bandwidth which tends to zero as $n \to \infty$.

One way of estimating $F^{-1}(s)$ is to use a variant of the empirical quantile function for the linear model proposed in Bassett and Koenker (1982),

$$\widehat{Q}(\tau|\overline{x}) = \overline{x}^\top \widehat{\alpha}(\tau). \tag{16}$$

If we use (16) in the formula (15), the density $f(F^{-1}(t))$ can be estimated by

$$f_n(F_n^{-1}(t)) = \frac{2h_n}{\overline{x}^\top \left(\widehat{\alpha}(t + h_n) - \widehat{\alpha}(t - h_n)\right)}.$$

For the long-run variance and covariance parameters, we may use the kernel estimators

$$\sigma_w^2 = \sum_{h=-M}^{M} k\left(\frac{h}{M}\right) C_{ww}(h), \sigma_{w\psi}(\tau) = \sum_{h=-M}^{M} k\left(\frac{h}{M}\right) C_{w\psi}(h), \lambda_w = \sum_{h=1}^{M} k\left(\frac{h}{M}\right) C_{ww}(h),$$

where $k(\cdot)$ is the lag window defined on $[-1, 1]$ with $k(0) = 1$, and M is the bandwidth (truncation) parameter satisfying the property that $M \to \infty$ and $M/n \to 0$ (say $M = O(n^{1/3})$ for many commonly used kernels, as in Andrews, 1991) as the sample size $n \to \infty$. The quantities $C_{ww}(h)$, and $C_{w\psi}(h)$ are sample covariances defined by $C_{ww}(h) = n^{-1} \sum' w_t w_{t+h}$, $C_{w\psi}(h) = n^{-1} \sum' w_t \psi_\tau(\widehat{u}_{(t+h)\tau})$, where $\sum'$ signifies summation over $1 \le t, t + h \le n$. Candidate kernel functions can be found in standard texts (e.g., Hannan, 1970).

4. Monte Carlo Results

In this section, we report on a Monte Carlo experiment designed to examine the finite sample performance of the inference procedures that we proposed in Section 3. There is a large literature of Monte Carlo studies on the size and power properties of the traditional unit root tests (see, e.g., Dickey and Fuller (1979), Said and Dickey (1985), Schwert (1989), Stock (1995)). One general conclusion to emerge is that, although difference exists across tests, the discriminatory power in the traditional tests between models with a root at unity and a root close to unity is generally low. In this section, we examine, for various error distributions, the finite sample properties for the quantile autoregression-based procedures, and compare them to the conventional unit root tests based on least squares regression. The experiment is designed particularly to explore the comparison for the heavy-tailed data, but includes a comparision for Gaussian time series as a benchmark.

The Monte Carlo considers the following design which is the leading case studied in the literature:

$$y_t = \alpha y_{t-1} + u_t, \tag{17}$$

where u_t are i.i.d. random variables. Different values of α and different types of error distribution are examined in the experiment. In particular, four values of α were considered: 1.0, 0.95, 0.9, 0.85. In addition, we also consider an alternative with random coefficient $\alpha = \alpha_t = \min\{0.5 + \Phi(u_t), 1\}$, where u_t is standard normal and Φ is the CDF of standard normal, this alternative is denoted as $\alpha = \alpha_t$ in Table 1. When $\alpha = 1$, the rejection rate gives the empirical size of the tests. Other cases deliver the empirical power.

Both normal and nonnormal disturbances are considered in our experiment. We are interested in the sampling performance of these tests with the presence of different type of error distributions, especially heavy-tailed innovations, which is an important feature in many financial and economic data. We report results for the following cases of error distribution: (1) u_t are i.i.d.N(0,1) variates; (2) u_t are student-t distributed variables with 2 degrees of freedom; (3) u_t are student-t distributed variables with 3 degrees of freedom; (4) u_t are student-t distributed variables with 4 degrees of freedom. The last three disturbances have heavy-tail distributions, and Assumption A1 does not hold in the second case where u_t are t variates with 2 degrees of freedom.

We report the Monte Carlo results for the following tests[4]: (1) The t-ratio test $t_n(\tau)$ based on quantile autoregression at $\tau = 0.5$, using the critical values in Table A; (2) The t-ratio test $t_n^*(\tau)$ based on quantile autoregression at $\tau = 0.5$, using the bootstrapped critical values; (3) The Kolmogorov-Smirnov (KS) type test (QKS_α) based on quantile autoregression with $\mathcal{T} = [0.1, 0.9]$, using the bootstrapped critical values; (4) The Cramer-von-Mises (CM) type test (QCM_α) with $\mathcal{T} = [0.1, 0.9]$ and the bootstrapped critical values; (5) The unmodified rank test of Hasan and Koenker (1997) using the Wilcoxon score function;[5] (6) The unmodified rank test of Hasan and Koenker (1997) using the normal (van der Waerden) score function; (7) The unmodified rank test of Hasan and Koenker (1997) using the sign score function; (8) The classical ADF t-ratio test (ADF_t); (9) The classical ADF coefficient-based test (ADF_α); (10) The Phillips-Perron semiparametric Z_t test; (11) The Phillips-Perron semiparametric Z_α test. The first four tests are based on the QAR model proposed in this paper, the 5-7th tests are rank tests using different scores, and the last four tests are OLS-based procedures that are widely used in applications. The order of ADF regressions

---

[4]The results here are drawn from a larger experiment with additional tests, including various procedures constructed based on $U_n(\tau)$ and $t_n(\tau)$, and modified versions of the rank tests. Qualitatively similar results are found from these tests. To conserve space, we report only the results from the 11 representative procedures.

[5]The unmodified versions of the Hasan and Koenker tests are based on their S_T statistic, defined in their equation (3.4) with the score functions normalized to have $\mathcal{L}_2$ norm one. Critical values for these tests are based on the procedure discussed at the end of Section 3.1 using an estimated value of δ^2 and the Hansen (1995) table. Performance of this unmodified version of the test is somewhat superior to the modified version particularly near the Gaussian model. This is consistent with the findings reported in Thompson (2001).

Table 1: Size and Power, Case with Gaussian Innovations										
$n = 100$										
	ADF_t	ADF_α	Z_t	Z_α	HK_W	HK_N	HK_S	QKS	QCM	$t_n(\tau)$	$t_n^*(\tau)$
$\alpha = 1$	0.061	0.056	0.070	0.060	0.043	0.036	0.051	0.05	0.06	0.122	0.06
$\alpha = 0.95$	0.121	0.134	0.130	0.139	0.101	0.096	0.084	0.110	0.120	0.194	0.11
$\alpha = 0.90$	0.246	0.261	0.238	0.252	0.241	0.252	0.173	0.211	0.230	0.281	0.196
$\alpha = 0.85$	0.581	0.680	0.611	0.672	0.482	0.468	0.318	0.486	0.514	0.578	0.430
$\alpha = \alpha_t$	0.52	0.665	0.535	0.656	0.093	0.163	0.024	0.850	0.780	0.150	0.250
$n = 200$										
$\alpha = 1$	0.06	0.05	0.07	0.06	0.054	0.053	0.044	0.06	0.05	0.106	0.06
$\alpha = 0.95$	0.359	0.391	0.348	0.372	0.277	0.288	0.218	0.289	0.350	0.391	0.26
$\alpha = 0.90$	0.82	0.922	0.84	0.906	0.765	0.783	0.493	0.73	0.78	0.78	0.64
$\alpha = 0.85$	1	1	0.99	1	0.968	0.977	0.747	0.90	0.95	0.94	0.88
$\alpha = \alpha_t$	0.920	0.935	0.918	0.936	0.221	0.431	0.018	0.975	0.990	0.051	0.064

and the QAR regressions are set at 2. We use the Bofinger (1975) bandwidth in estimation of the sparsity function in constructing the t statistic. The Cramer-von-Mises statistic is calculated using numerical integration with step size equals 0.01, and the Kolmogorov-Smirnov (KS) type test is calculated using the same step size. The semiparametric tests Z_t and Z_α are calculated using the procedure PPZAZT of COINT 2.0 (Ouliaris and Phillips 1994). The number of repetitions in the bootstrapping process is 2000. For each test, the number of repetition is 1000. Two sample sizes are studied: $n = 100$ and $n = 200$.

Table 1 reports the empirical size and power for the case with Gaussian innovations. In the presence of Gaussian errors, the OLS based tests have better performance than procedures based on quantile regression. In this case, the t-ratio test $t_n(\tau)$ using the critical values in Table A has the largest size distortions than other tests. Although the performance is improved as the sample size increase from 100 to 200, the results are qualitatively very similar. Table 1 also gives the empirical power of these tests against the random coefficient alternative $\alpha = \alpha_t$. In this case, the Kolmogorov-Smirnov (KS) or Cramer-von-Mises (CM) type tests based on the regression quantile process for $\tau \in \mathcal{T}$ have the highest power over all procedures.

Table 2, 3, and 4 report results for cases where the errors have student-t innovations with degrees of freedom 2, 3 and 4 respectively. Results in these tables indicate that the QAR based procedures are in general superior in the presence of heavy-tail disturbances. The OLS based tests have lower power. From results in Tables 2 and 3, we can see that the power gain by using quantile based method can be quite substantial over certain range of parameter values. Although (from Table 1) there is a loss in power by using the quantile autoregression based tests under normality, the power loss is small relative to the gain in power in the presence of heavy tailed distribution.

A comparison can also be made among the four procedures based on the QAR model. The tests using bootstrapped critical values have better size properties than the t-test $t_n(\tau)$ using the critical values in Table A. However, $t_n(\tau)$ using the critical values in Table A has the best power in the presence of heavy-tailed innovations. For the comparison between the two QAR-based tests using information over $\mathcal{T} = [0.1, 0.9]$, the Cramer-von-Mises type test (QCM) has relatively better finite sample performance than the Kolmogorov-Smirnov type test (QKS).

A final comparison can be made between the proposed quantile autoregression based tests and the rank type tests. The QAR-based tests proposed in this paper are constructed from the unrestricted

Table 2: Size and Power, Case with $t(2)$ Innovations										
$n = 100$										
	ADF_t	ADF_α	Z_t	Z_α	HK_W	HK_N	HK_S	QKS	QCM	$t_n(\tau)$	$t_n^*(\tau)$
$\alpha = 1$	0.062	0.048	0.066	0.058	0.045	0.040	0.050	0.054	0.051	0.058	0.052
$\alpha = 0.95$	0.114	0.173	0.127	0.193	0.505	0.392	0.504	0.365	0.444	0.526	0.38
$\alpha = 0.9$	0.334	0.530	0.356	0.561	0.808	0.723	0.743	0.698	0.782	0.834	0.730
$\alpha = 0.85$	0.616	0.779	0.622	0.781	0.930	0.886	0.880	0.815	0.909	0.961	0.872
$n = 200$										
$\alpha = 1$	0.08	0.06	0.08	0.06	0.056	0.062	0.060	0.061	0.053	0.057	0.052
$\alpha = 0.95$	0.29	0.36	0.30	0.36	0.897	0.835	0.843	0.813	0.880	0.922	0.870
$\alpha = 0.9$	0.89	0.92	0.90	0.912	1.000	0.998	0.991	0.982	0.996	0.999	0.991
$\alpha = 0.85$	0.97	1	0.97	0.99	1	1	0.998	0.995	1	1	1

Table 3: Size and Power, Case with $t(3)$ Innovations										
$n = 100$										
	ADF_α	ADF_α	Z_t	Z_α	HK_W	HK_N	HK_S	QKS	QCM	$t_n(\tau)$	$t_n^*(\tau)$
$\alpha = 1$	0.056	0.048	0.057	0.057	0.048	0.055	0.044	0.053	0.052	0.078	0.054
$\alpha = 0.95$	0.126	0.203	0.147	0.211	0.244	0.191	0.258	0.180	0.240	0.330	0.232
$\alpha = 0.9$	0.37	0.45	0.38	0.46	0.552	0.443	0.495	0.420	0.600	0.62	0.47
$\alpha = 0.85$	0.54	0.68	0.55	0.68	0.804	0.723	0.702	0.610	0.840	0.85	0.70
$n = 200$										
$\alpha = 1$	0.06	0.07	0.07	0.06	0.059	0.048	0.050	0.06	0.05	0.08	0.05
$\alpha = 0.95$	0.30	0.43	0.31	0.44	0.641	0.535	0.528	0.61	0.70	0.79	0.67
$\alpha = 0.9$	0.85	0.94	0.87	0.95	0.977	0.960	0.892	0.95	0.97	0.98	0.95
$\alpha = 0.85$	0.99	1	0.99	1	1	1	0.986	0.99	1	1	1

Table 4: Size and Power, Case with $t(4)$ Innovations										
$n = 100$										
	ADF_t	ADF_α	Z_t	Z_α	HK_W	HK_N	HK_S	QKS	QCM	$t_n(\tau)$	$t_n^*(\tau)$
$\alpha = 1$	0.066	0.070	0.068	0.071	0.058	0.043	0.060	0.047	0.056	0.074	0.057
$\alpha = 0.95$	0.137	0.205	0.146	0.224	0.173	0.134	0.160	0.190	0.250	0.270	0.172
$\alpha = 0.9$	0.352	0.501	0.355	0.512	0.475	0.407	0.393	0.510	0.550	0.575	0.474
$\alpha = 0.85$	0.68	0.77	0.69	0.78	0.714	0.643	0.564	0.76	0.79	0.83	0.70
$n = 200$										
$\alpha = 1$	0.055	0.057	0.057	0.062	0.051	0.051	0.047	0.051	0.052	0.078	0.061
$\alpha = 0.95$	0.31	0.47	0.33	0.46	0.486	0.397	0.407	0.51	0.60	0.65	0.55
$\alpha = 0.9$	0.87	0.91	0.88	0.90	0.938	0.899	0.805	0.92	0.93	0.95	0.92
$\alpha = 0.85$	0.97	0.99	0.99	1	0.995	0.994	0.952	0.98	1	0.99	0.97

Table 5: Descriptive Statistics			
	1 month rate	3 month rate	annual rate
Skewness	-1.8435	-1.543	-0.1410
Kurtosis	28.062	24.41	3.88
Jarque-Bera	9759	7117	2.80

quantile autoregression, while the rank tests of Hasan and Koenker are based on the restricted (under the unit root null) quantile regression. The Monte Carlo results indicate that these tests have similar performance when the alternatives are constant coefficient process. In the presence of random coefficient alternative $\alpha = \alpha_t$, the KS or CM type tests proposed in this paper have better performance than other tests, although the rank tests are also based on quantile regression estimates over a range of quantiles. Such a difference might be attributed to bias in estimating α_1 in constructing the rank tests based on the restricted (null) model under the alternative.

5. US Interest Rate Dynamics

We now apply the QAR-based unit root tests to several US interest rate series. The behavior of short-term interest rates is central to much of the theoretical and empirical work in macroeconomics and finance. However, there is still no consensus on the dynamics of short term interest rates. In this section, we examine the unit root property of several interest rate series using our proposed procedures. We focus on the interest rate itself and do not consider multifactor (term structure) models.

Many empirical studies in the unit root literature have investigated U.S. interest rate data. Nelson and Plosser (1982) studied the unit root property of US annual interest rates in their seminal work on fourteen macroeconomic time series. This series and other type interest rates have been often re-examined. Among various empirical findings, two important features have been well-documented: (1) First, evidence based on the traditional unit root tests has accumulated suggesting that there is a unit root in interest rates. See, e.g., Nelson and Plosser (1982), Schotman and Van Dijk (1991), El-Jahel et al. (1997), Ball and Torous, (1996); (2) Another well-documented characteristic of the interest rate time series is its non-Gaussianity, the leptokurtosis and heavy-tailed features in these time series are usually accentuated when the data are sampled more frequently. In this section, we revisit the interest rate series using the proposed QAR methods.

The time series that we consider are one month, three month, and annual interest rates in the US. In particular, we looked at (seasonally adjusted) one month and three month commercial paper rates, and the annual bond yield from the extended Nelson-Plosser data. Both the one month rate and the three month rate start from April, 1971 and end at June, 2002, with 378 observations. The annual data are from 1900 to 1988. We first apply the augmented Dickey-Fuller (ADF) unit root tests to these series. In the ADF regressions, the BIC criterion of Schwarz (1978) and Rissanen (1978) is used in selecting the appropriate lag length of the autoregressions. The OLS based ADF regression estimates of the largest autoregressive roots of the three interest series are all very close to unity (see the estimates of the largest AR coefficients in Tables 6A, 7A, 8A). Tables 6A, 7A, and 8A report the ADF test statistics for the 1 month, 3 month and annual series respectively. The unit root hypothesis can not be rejected by the traditional ADF test at 5% level of significance, leading to the conclusion that the interest rate series exhibit unit roots.

Table 5 presents some descriptive statistics of the ADF regression residuals of three interest rate time series. All series exhibit negative skewness. The kurtosis of all these series exceed 3. Tests based on the Jarque-Bera procedure suggest departures from Gaussianity in the 1 month and 3 month series.

Table 6A				
	ADF_α	ADF_t	QKS_α	QCM_α
Test Statistics	-11.54	-2.22	41.39**	326.21**
5% Critical Values	-14.1	-2.86	20.04	48.73
OLS Estimator		$\widehat{\alpha}_1 = 0.976$		

Table 6B							
Quantiles	$\widehat{\alpha}_1(\tau)$	$t_n(\tau)$	$\widehat{\delta}^2$	$U_n(\tau)$	Critical Values for $U_n(\tau)$		
					2.5% c.v	5% c.v.	95% c.v.	97.5% c.v.
0.1	0.886	-4.65	0.142	-41.4**	-22.75	-18.06	2.88	4.19
0.2	0.929	-4.74	0.184	-26.4**	-12.31	-10.11	1.62	2.40
0.3	0.961	-4.20	0.201	-14.3**	-7.48	-5.94	1.05	1.55
0.4	0.981	-2.44	0.248	-7.06*	-3.69	-3.10	0.49	0.75
0.5	0.994	-0.75	0.177	-2.08	-3.61	-2.85	0.49	0.74
0.6	1.014	1.84	0.160	5.39#	-5.65	-4.51	0.76	1.17
0.7	1.029	3.18	0.163	11.13##	-7.75	-6.30	1.12	1.67
0.8	1.055	3.82	0.137	20.47##	-11.17	-9.14	1.67	2.36
0.9	1.111	5.55	0.116	41.39##	-18.83	-15.11	2.49	3.73

The values of $U_n(\tau)$ denoted by an (*) are significant the 5% level when the alternative is $H_{1A} : \alpha_1 < 1$. Those with an (**) are significant at the 1% level with alternative H_{1A}. Similarly, the values of $\widehat{\alpha}_1(\tau)$ denoted by an (#) are significant the 5% level when the alternative is $H_{1B} : \alpha_1 > 1$. Those with an (##) are significant at the 1% level when the alternative is H_{1B}.

Table 7A				
	ADF_α	ADF_t	QKS_α	QCM_α
Test Statistics	-11.14	-2.17	43.03**	341.33**
5% Critical Values	-14.1	-2.86	19.74	40.67
OLS Estimator		$\widehat{\alpha}_1 = 0.977$		

Table 7B							
Quantiles	$\widehat{\alpha}_1(\tau)$	$t_n(\tau)$	$\widehat{\delta}^2$	$U_n(\tau)$	Critical Values for $U_n(\tau)$		
					2.5%	5%	95%	97.5%
0.1	0.884	-4.74	0.189	-43.03	-19.99	-16.51	2.54	3.88
0.2	0.926	-5.52	0.200	-27.55	-9.45	-7.62	1.24	1.84
0.3	0.959	-4.30	0.165	-15.22**	-6.59	-5.50	1.03	1.47
0.4	0.984	-1.71	0.188	-5.83	-4.39	-3.50	0.62	0.86
0.5	0.991	-0.76	0.161	-3.51	-4.03	-3.26	0.51	0.76
0.6	1.012	1.65	0.179	4.82	-5.53	-4.48	0.76	1.19
0.7	1.034	3.23	0.136	12.62	-7.79	-6.44	1.06	1.52
0.8	1.065	4.76	0.185	23.91	-10.95	-8.78	1.60	2.21
0.9	1.107	6.35	0.118	39.73	-20.41	-15.99	2.93	4.13

Table 8A				
	ADF_α	ADF_t	QKS_α	QCM_α
Test Statistics	-3.15	-1.02	14.65	67.39*
5% Critical Values	-14.1	-2.86	23.49	65.42
OLS Estimator		$\widehat{\alpha}_1 = 0.974$		

Table 8B								
Quantiles	$\widehat{\alpha}_1(\tau)$	$t_n(\tau)$	$\widehat{\delta^2}$	$U_n(\tau)$	Critical Values for $U_n(\tau)$			
					2.5%	5%	95%	97.5%
0.1	0.829	-5.47	0.142	-14.64	-26.89	-22.13	3.33	4.95
0.2	0.865	-4.33	0.101	-11.61	-12.12	-9.15	1.24	2.02
0.3	0.965	-2.42	0.200	-2.99	-6.00	-4.50	0.89	1.25
0.4	0.981	-1.01	0.120	-1.64	-5.24	-4.12	0.92	1.22
0.5	1.004	0.06	0.168	0.39	-5.52	-4.40	0.90	1.35
0.6	1.052	1.52	0.268	4.44	-6.88	-5.48	1.17	1.74
0.7	1.126	4.02	0.163	10.84	-9.53	-7.69	1.61	2.45
0.8	1.165	4.79	0.194	14.25	-13.32	-11.11	2.22	3.36
0.9	1.126	3.76	0.257	10.83	-17.88	-14.65	3.06	4.75

We reconsider these series using the proposed QAR methods and report the results in Tables 6, 7, 8 respectively. In particular, we applied the following four tests to the interest rate series: (1) The Kolmogorov-Smirnov (KS) type test QKS_α based on quantile autoregression with $\mathcal{T} = [0.1, 0.9]$, using the bootstrapped critical values; (2) The Cramer-von-Mises (CM) type test QCM_α with $\mathcal{T} = [0.1, 0.9]$ and the bootstrapped critical values; (3) The t-ratio test $t_n(\tau)$ based on quantile autoregression at each decile using the critical values in Table A; (4) The coefficient based tests $U_n(\tau)$ based on quantile autoregression at each decile using the bootstrapped critical values. The first two tests provide a general analysis of the unit root behavior based on a range of quantiles. The third and forth tests try to provide a more detailed examination on the unit root property of these series at each decile.

Tables 6A, 7A and 8A report the QKS_α and QCM_α tests for the three time series respectively. The 5% level critical values calculated based on the resampling procedure given in Section 3 are also reported in these tables. For both the 1 month and 3 month data, the unit root hypothesis is rejected at 1% level by both tests. For the annual data, the unit root hypothesis is only marginally rejected by the Cramer-von-Mises test at 5% level, but not rejected by the Kolmogorov-Smirnov test QKS_α. In summary, there is a strong evidence that the short term interest rate series (1 month and 3 month rates) are not constant unit root process.

Tables 6B, 7B and 8B take a closer look on the interest rate dynamics by examining the unit root behavior at various quantiles. The second columns in these tables report the estimates of the largest autoregressive roots at each decile. The evidence based on these point estimates of the largest autoregressive root at each quantile suggests that the interest rate series are not constant unit root processes. From all these three tables we can see that there is asymmetry in the persistency. The largest autoregressive coefficient estimate $\widehat{\alpha}_1(\tau)$ has different values over different quantiles, displaying asymmetric dynamics over the business cycle. In particular, $\widehat{\alpha}_1(\tau)$ increases when we move from lower quantiles to higher quantiles. The autoregressive coefficient values at the lower quantiles are smaller than those at higher quantiles, indicating that the local behavior of the interest rate

during a recession would be much more stationary than its behavior during an expansion. Interest rates appear to exhibit asymmetric adjustment dynamics. In the presence of positive shocks to the economy, the interest rate is more persistent. This finding of asymmetric dynamics is consistent with the interest rate smoothing by the Federal Reserve Board. It may be more acceptable for the Fed to lower rates more quickly and by a larger amount than to raise rates in the same way. Instead, the Fed tends to gradually raise rates in small amounts for a longer period of time. Consequently, interest rates are more persistent in the presence of positive shocks than to negative ones.

The third columns in Tables 6B, 7B, 8B report the calculated coefficient statistic $U_n(\tau)$ for the three time series. Given the possibility of both locally stationary and locally explosive behavior at different quantiles, we consider both the one-sided and the two-sided alternative hypotheses. Columns 6 to 7 of these tables reports 2.5%, 5%, 95% and 97.5% quantiles (and thus the generated critical values) of the null distribution of $U_n(\tau)$ calculated under the unit root null using the resampling procedure described in Section 3.

We also consider tests for the unit root hypothesis based the autoregression estimates $\widehat{\alpha}_1(\tau)$ at selected quantiles. The third and fifth columns in Tables 6B, 7B and 8B report the calculated t-statistic $t_n(\tau)$ and coefficient statistic $U_n(\tau)$ for the three time series. The estimated δ^2 are reported in the forth columns of these tables. The majority of results reject the unit root null hypothesis. In order to test the unit root hypothesis against different alternatives: $H_{1A} : \alpha_1 < 1$; $H_{1B} : \alpha_1 > 1$; and $H_{1C} : \alpha_1 \neq 1$, using the coefficient-based statistic $U_n(\tau)$ at each specified quantiles, we report, from columns 6 to 9 in Tables 6B, 7B and 8B, quantiles (and thus the critical values) of the null distribution of $U_n(\tau)$ calculated under the unit root null and based on the resampling procedure in Section 3.

If we test the unit root hypothesis at these specified quantiles, we can see that only at quantiles that are around median can the unit root hypothesis not be rejected. At both low quantiles and high quantiles the unit root hypothesis is rejected. At low quantiles, the autoregressive roots are usually smaller than unity. At high quantiles, the estimate become larger than one, displaying mildly explosive behavior. Combining this evidence with the results of tables 6A, 7A, and 8A, we find significant support for asymmetry in the business cycle dynamics of short term interest rates.

We believe that the quantile regression based inference procedures have some advantages over the least squared based tests in analyzing dynamics and persistency in time series with heavy-tailed distributions. Quantile regression methods offer a mechanism for estimating models for the conditional median function and the full range of other conditional quantile functions. By supplementing the estimation of conditional mean functions with techniques for estimating an entire family of conditional quantile functions, quantile regression provides a relatively complete statistical analysis of the stochastic relationships among random variables. In addition, it also provide a more robust and efficient approach than the least square method when the data is non-Gaussian or is contaminated by outliers.

6. Appendix: Tables and Proofs

6.1. Table 1: Asymptotic Critical Values of The t-Statistic $t_n(\tau)$ Given By (10).

δ^2	1%	5%	10%
0.9	-3.39	-2.81	-2.50
0.8	-3.36	-2.75	-2.46
0.7	-3.30	-2.72	-2.41
0.6	-3.24	-2.64	-2.32
0.5	-3.19	-2.58	-2.25
0.4	-3.14	-2.51	-2.17
0.3	-3.06	-2.40	-2.06
0.2	-2.91	-2.28	-1.92
0.1	-2.78	-2.12	-1.75

6.2. Proof of Theorem 2.1 and Theorem 2.2.

We follow the approach of Knight (1989) (also see Pollard (1991)) which is based on a convexity lemma that the quantile regression objective function satisfies. We use the following identity: if we denote $\psi_\tau(u) = \tau - I(u < 0)$, for $u \neq 0$,

$$
\begin{aligned}
\rho_\tau(u - v) - \rho_\tau(u) &= -v\psi_\tau(u) + (u - v)\{I(0 > u > v) - I(0 < u < v)\} \\
&= -v\psi_\tau(u) + \int_0^v \{I(u \leq s) - I(u < 0)\}ds.
\end{aligned}
\tag{18}
$$

Let

$$
u_{t\tau} = y_t - x_t'\alpha(\tau) = u_t - F_u^{-1}(\tau),
\tag{19}
$$

then $u_{t\tau}$ satisfies the quantile restriction that

$$
Q_{u_{t\tau}}(\tau|\mathcal{F}_{t-1}) = 0.
\tag{20}
$$

If we denote $v = D_n(\alpha - \alpha(\tau))$, where $D_n = diag(\sqrt{n}, n, \sqrt{n}\cdots, \sqrt{n})$, the minimization (3) is equivalent to

$$
\min_v \sum_{t=1}^n \left[\rho_\tau(u_{t\tau} - v'D_n^{-1}x_t) - \rho_\tau(u_{t\tau})\right].
$$

Using identity (18), we have

$$
\sum_{t=1}^n \left[\rho_\tau(u_{t\tau} - v'D_n^{-1}x_t) - \rho_\tau(u_{t\tau})\right]
$$

$$
= -\sum_{t=1}^n v'D_n^{-1}x_t\psi_\tau(u_{t\tau}) + \sum_{t=1}^n (u_{t\tau} - v'D_n^{-1}x_t)\{I(0 > u_{t\tau} > v'D_n^{-1}x_t) - I(0 < u_{t\tau} < v'D_n^{-1}x_t)\}.
$$

Notice again that u_t are uncorrelated with y_{t-1}, under Assumption A1 (or A1′), we have

$$
n^{-1}\sum_{t=1}^n y_{t-1}\psi_\tau(u_{t\tau}) \Rightarrow \int_0^1 B_w dB_\psi^\tau.
$$

We also need to consider the limiting distribution of

$$
\begin{bmatrix}
\frac{1}{\sqrt{n}} \sum_t \Delta y_{t-1} \psi_\tau(u_{t\tau}) \\
\vdots \\
\frac{1}{\sqrt{n}} \sum_t \Delta y_{t-q} \psi_\tau(u_{t\tau})
\end{bmatrix}.
\tag{21}
$$

If we denote that $E[w_t w_{t-j}] = \nu_j$, it can be shown that (21) converges to a q-dimensional normal variate $\Phi = [\Phi_1, \cdots, \Phi_q]^\top$ with covariance matrix $\tau(1-\tau)\Omega_\Phi$ where

$$
\Omega_\Phi = \begin{bmatrix}
\nu_0 & \cdots & \nu_{q-1} \\
\vdots & \ddots & \vdots \\
\nu_{q-1} & \cdots & \nu_0
\end{bmatrix},
$$

and Φ is independent with $\int_0^1 \overline{B}_w dB_\psi^\tau$. Thus,

$$
D_n^{-1} \sum_{t=1}^n x_t \psi_\tau(u_{t\tau}) =
\begin{bmatrix}
\frac{1}{\sqrt{n}} \sum_t \psi_\tau(u_{t\tau}) \\
\frac{1}{n} \sum_t y_{t-1} \psi_\tau(u_{t\tau}) \\
\frac{1}{\sqrt{n}} \sum_t \Delta y_{t-1} \psi_\tau(u_{t\tau}) \\
\vdots \\
\frac{1}{\sqrt{n}} \sum_t \Delta y_{t-q} \psi_\tau(u_{t\tau})
\end{bmatrix}
\Rightarrow
\begin{bmatrix}
\int_0^1 dB_\psi^\tau \\
\int_0^1 B_w dB_\psi^\tau \\
\Phi_1 \\
\vdots \\
\Phi_q
\end{bmatrix}
=
\begin{bmatrix}
\int_0^1 \overline{B}_w dB_\psi^\tau \\
\Phi
\end{bmatrix}
:= \Phi^*
$$

where Φ is a q-dimensional normal variate with covariance matrix $\tau(1-\tau)\Omega_\Phi$, and is independent with $\int_0^1 \overline{B}_w dB_\psi^\tau$.

We now consider the limit of

$$
\sum_{t=1}^n (u_{t\tau} - v' D_n^{-1} x_t) I(0 < u_{t\tau} < v' D_n^{-1} x_t).
$$

For convenience of asymptotic analysis, we denote

$$
U_n(v) = \sum_{t=1}^n z_t(v), \text{ where } z_t(v) = (v' D_n^{-1} x_t - u_{t\tau}) I(0 < u_{t\tau} < v' D_n^{-1} x_t).
$$

To avoid technical problems in taking conditional expectations, following Knight (1989), we consider truncation of $v_2 n^{-1/2} y_{t-1}$ at some finite number $m > 0$ and denote

$$
\begin{aligned}
U_{nm}(v) &= \sum_{t=1}^n z_{tm}(v), \\
z_{tm}(v) &= (v' D_n^{-1} x_t - u_{t\tau}) I(0 < u_{t\tau} < v' D_n^{-1} x_t) M_t \\
M_t &= I(0 \le v_2 n^{-1/2} y_{t-1} \le m).
\end{aligned}
$$

We further define

$$
\overline{z}_{tm}(v) = \mathrm{E}\{(v' D_n^{-1} x_t - u_{t\tau}) I(0 < u_{t\tau} < v' D_n^{-1} x_t) M_t | \mathcal{F}_{t-1}\},
$$

and

$$
\overline{U}_{nm}(v) = \sum_{t=1}^n \overline{z}_{tm}(v),
$$

then $\{z_{tm}(v) - \overline{z}_{tm}(v)\}$ is a martingale difference sequence. Notice that

$$\overline{U}_{nm}(v)$$

$$= \sum_{t=1}^{n} E\{(v'D_n^{-1}x_t - u_{t\tau})I(0 < u_{t\tau} < v'D_n^{-1}x_t)M_t|\mathcal{F}_{t-1}\}$$

$$= \sum_{t=1}^{n} \int_{F_u^{-1}(\tau)}^{[v'D_n^{-1}x_t+F_u^{-1}(\tau)]M_t} [(v'D_n^{-1}x_t + F_u^{-1}(\tau))M_t - r]f_u(r)dr$$

$$= \sum_{t=1}^{n} \int_{F_u^{-1}(\tau)}^{[v'D_n^{-1}x_t+F_u^{-1}(\tau)]M_t} [\int_{r}^{[v'D_n^{-1}x_t+F_u^{-1}(\tau)]M_t} ds]f_u(r)dr$$

$$= \sum_{t=1}^{n} \int_{F_u^{-1}(\tau)\leq r\leq[v'D_n^{-1}x_t+F_u^{-1}(\tau)]M_t)} \int_{r\leq s\leq[v'D_n^{-1}x_t+F_u^{-1}(\tau)]M_t} f_u(r)dsdr$$

$$= \sum_{t=1}^{n} \int_{F_u^{-1}(\tau)\leq s\leq[v'D_n^{-1}x_t+F_u^{-1}(\tau)]M_t} \int_{F_u^{-1}(\tau)\leq r\leq s} f_u(r)drds$$

$$= \sum_{t=1}^{n} \int_{F_u^{-1}(\tau)}^{[v'D_n^{-1}x_t+F_u^{-1}(\tau)]M_t} \left[\int_{F_u^{-1}(\tau)}^{s} f_u(r)dr\right] ds$$

$$= \sum_{t=1}^{n} \int_{F_u^{-1}(\tau)}^{[v'D_n^{-1}x_t+F_u^{-1}(\tau)]M_t} [s - F_u^{-1}(\tau)] \left[\frac{F_u(s) - F_u(F_u^{-1}(\tau))}{s - F_u^{-1}(\tau)}\right] ds.$$

Under Assumption A2,

$$\overline{U}_{nm}(v) = \sum_{t=1}^{n} \int_{F_u^{-1}(\tau)}^{[v'D_n^{-1}x_t+F_u^{-1}(\tau)]M_t} [s - F_u^{-1}(\tau)]f_u[F_u^{-1}(\tau)]ds + o_p(1)$$

$$= \sum_{t=1}^{n} f_u[F_u^{-1}(\tau)] \left\{\frac{[s - F_u^{-1}(\tau)]^2}{2}\Big|_{F^{-1}(\tau)}^{[v'D_n^{-1}x_t+F^{-1}(\tau)]M_t}\right\} + o_p(1)$$

$$= \frac{1}{2}\sum_{t=1}^{n} f_u[F^{-1}(\tau)]v'[D_n^{-1}x_tx_t'D_n^{-1}]vM_t + o_p(1).$$

Thus

$$\overline{U}_{nm}(v) \Rightarrow \eta_m = \frac{1}{2}f_u[F_u^{-1}(\tau)]v'\Psi_{1m}v,$$

where

$$\Psi_{1m} = \left[\begin{array}{cc} \int_0^1 \overline{B}_w\overline{B}_w' I(0 \leq v_2'B_w(s) \leq m) & 0_q' \\ 0_q' & \Omega_\Phi \end{array}\right].$$

We now follow the arguments of Pollard (1984, p171), notice that $(v'D_n^{-1}x_t)I(0 \leq v_2n^{-1/2}y_{t-1} \leq m) \xrightarrow{P} 0$ uniformly in t,

$$\sum_{t=1}^{n} E[z_{tm}(v)^2|\mathcal{F}_{t-1}] \leq \max\{(v'D_n^{-1}x_t)I(0 \leq v_2n^{-1/2}y_{t-1} \leq m)\} \sum \overline{z}_{tm}(v) \xrightarrow{P} 0.$$

Thus the following summation of martingale difference sequence

$$\sum_{t}\{z_{tm}(v) - \overline{z}_{tm}(v)\}$$

converges to zero in probability. Thus the limiting distribution of $\sum_t z_{tm}(v)$ is the same as that of $\sum_t \overline{z}_{tm}(v)$, i.e.,

$$U_{nm}(v) \Rightarrow \eta_m.$$

Let $m \to \infty$, we have

$$\eta_m \Rightarrow \eta = \frac{1}{2} f(F^{-1}(\tau)) v' \Psi_1 v I(v_2 B_w(s) > 0),$$

and

$$\Psi_1 = \left[\begin{array}{cc} \int_0^1 \overline{B}_w \overline{B}_w' I(0 \leq v_2' B_w(s)) & 0_q' \\ 0_q' & \Omega_\Phi \end{array} \right].$$

Now, by a similar argument as Herce(1996), we show that

$$\lim_{m \to \infty} \lim_{n \to \infty} \sup \Pr[|U_n(v) - U_{nm}(v)| \geq \varepsilon] = 0.$$

Similarly, we can show that $\sum_{t=1}^n (u_{t\tau} - (D_n^{-1}v)'x_t)\{I(0 > u_{t\tau} > (D_n^{-1}v)'x_t)$ converges to

$$\frac{1}{2} f(F^{-1}(\tau)) v' \Psi_2 v I(v_2 B_w(s) \leq 0),$$

with

$$\Psi_2 = \left[\begin{array}{cc} \int_0^1 \overline{B}_w \overline{B}_w' I(v_2' B_w(s) \leq 0) & 0_q' \\ 0_q' & \Omega_\Phi \end{array} \right].$$

Thus,

$$\sum_{t=1}^n (u_{t\tau} - (D_n^{-1}v)'x_t)\{I(0 > u_{t\tau} > (D_n^{-1}v)'x_t) - I(0 < u_{t\tau} < (D_n^{-1}v)'x_t)\} \Rightarrow f(F^{-1}(\tau)) v' \Psi v,$$

where

$$\Psi = \left[\begin{array}{cc} \int_0^1 \overline{B}_w \overline{B}_w' & 0_q' \\ 0_q' & \Omega_\Phi \end{array} \right].$$

As a result,

$$
\begin{aligned}
Z_n(v) &= \sum_{t=1}^n \left[\rho_\tau(u_{t\tau} - (D_n^{-1}v)'z_t) - \rho_\tau(u_{t\tau}) \right] \\
&= -\sum_{t=1}^n (D_n^{-1}v)'z_t \psi_\tau(u_{t\tau}) \\
&\quad + \sum_{t=1}^n (u_{t\tau} - (D_n^{-1}v)'z_t)\{I(0 > u_{t\tau} > (D_n^{-1}v)'z_t) - I(0 < u_{t\tau} < (D_n^{-1}v)'z_t)\} \\
&\Rightarrow -v'\Phi^* + f(F^{-1}(\tau)) v' \Psi v \\
&:= Z(v)
\end{aligned}
$$

By the convexity Lemma of Pollard (1991) and arguments of Knight (1989), notice that $Z_n(v)$ and $Z(v)$ are minimized at $\widehat{v} = D_n(\widehat{\alpha}(\tau) - \alpha(\tau))$ and

$$\frac{1}{f(F^{-1}(\tau))} \left[\begin{array}{cc} \int_0^1 \overline{B}_w \overline{B}_w^\top & 0_{2 \times q} \\ 0_{q \times 2} & \Omega_\Phi \end{array} \right]^{-1} \left[\begin{array}{c} \int_0^1 \overline{B}_w dB_\psi^\tau \\ \Phi \end{array} \right],$$

respectively, by Lemma A of Knight (1989) we have,

$$D_n(\widehat{\alpha}(\tau) - \alpha(\tau)) \Rightarrow \frac{1}{f(F^{-1}(\tau))} \left[\begin{array}{cc} \int_0^1 \overline{B}_w \overline{B}_w^\top & 0_{2 \times q} \\ 0_{q \times 2} & \Omega_\Phi \end{array} \right]^{-1} \left[\begin{array}{c} \int_0^1 \overline{B}_w dB_\psi^\tau \\ \Phi \end{array} \right].$$

Portnoy (1984) shows that the quantile regression process is tight. Using the argument of Portnoy (1984), we obtain that the limiting variate $B_\psi^\tau(r)$, as a random function of τ, is a Brownian bridge over $\tau \in \mathcal{T}$.   ■

REFERENCES

Ball, C.A., and Torous, W.N. (1996), "Unit Roots and the Estimation of Interest Rate Dynamics," *Journal of Empirical Finance*, **3**, 215-238.

Bassett, G., and Koenker, R. (1982), "An Empirical Quantile Function for Linear Models with iid Errors," *Journal of the American Statistical Association*, **77**, 407-415.

Billingsley, P. (1961), "The Lindeberg-Levy Theorem for Martingales," *Proc. Amer. Math. Soc.*, **12**, 788-792.

Bofinger, E. (1975), "Estimation of a Density Function Using Order Statistics," *Australian J. of Statistics*, **17**, 1-7.

Chan, N.H., and Wei,c.z. (1987), "Asymptotic Inference for Nearly Nonstationary AR(1) Processes," *Annals of Statistics*, **15**, 1050-1063.

Chang, Y., Park, J., and Song, K. (2001), "Bootstrapping Cointegrating Regressions," Working Paper, Rice University.

Cox, D., and Llatas, I. (1991), "Maximum Likelihood Type Estimation for Nearly Non-Stationary Autoregressive Time Series," *Annals of Statistics*, **19**, 1109-1128.

Davis, R., Knight, K., and Liu, J. (1992), "M-estimation for Autoregressions with Infinite Variance," *Stochastic Processes and Their Applications*, **40**, 145-180.

Dickey, D.A., and Fuller, W.A. (1979), "Distribution of Estimators for Autoregressive Time Series with a Unit Root," *Journal of the American Statistical Association*, **74**, 427-431.

El-Jahel, L., Lindberg, H., and Perraudin, W. (1997), "Interest Rate Distributions, Yield Curve Modelling and Monetary Policy," in *Mathematics of Derivative Securities*, eds. Dempster and Pliska, Cambridge University Press, pp.1-35.

Ferretti, N., and Romo, J. (1995), "Unit Root Bootstrap Tests for AR(1) Models," *Biometrica*, **83**, 849-860.

van Giersbergen, N.P.A. (1996), "Bootstrapping Unit Root Tests in AR(1) Model with Drift," working paper, University of Amsterdam.

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

Hannan, E.J. (1973), "Central Limit Theorems for Time Series Regression," *Z. Wahrsch. Verw. Gebiete*, **26**, 157-170.

Hansen, B. (1992), "Convergence to Stochastic Integrals for Dependent Heterogeneous Processes," *Econometric Theory*, **8**, 489-500.

———(1995), "Rethinking the Univariate Approach to Unit Root Tests: How to Use Covariates to Increase Power," *Econometric Theory*, **11**, 1148-1171.

Hasan, M.N. (2001), "Rank Tests of Unit Root Hypothesis with Infinite Variance Errors," *Journal of Econometrics*, **104**, 49-65.

——— and Koenker, R. (1997), "Robust Rank Tests of the Unit Root Hypothesis," *Econometrica*, **65**, 133-161.

Herce, M. (1996), "Asymptotic Theory of LAD Estimation in a Unit Root Process with Finite Variance Errors," *Econometric Theory*, **12**, 129-153.

Juhl, T. (1999), "Testing for Cointegration Using M Estimators," Working Paper, University of Illinois.

Jureckova, J., and Hallin, M. (1999), "Optimal Tests for Autoregressive Models Based on Autoregression Rank Scores," *The Annals of Statistics*, **27**, 1385–1414.

Knight, K. (1989), "Limit Theory for Autoregressive Parameter Estimates in an Infinite-Variance Random Walk," *The Canadian Journal of Statistics*, **17**, 261-278.

Knight, K. (1991), "Limiting Theory for M Estimators in an Integrated Infinite Variance Process," *Econometric Theory*, **7**, 200-212.

Koenker, R., and Bassett, G. (1978), "Regression Quantiles", *Econometrica*, **46**, 33-49.

Koul, H., and Saleh, A.K. (1995), "Autoregression Quantiles and Related Rank-Scores Processes," *The Annals of Statistics*, **23**, 670-689.

Lucas, A. (1994), "An Outlier Robust Unit Root Test with an Application to the Extended Nelson-Plosser Data," *Journal of Econometrics* **66**, 153-174.

———(1995), "Unit Root Tests Based on M Estimators," *Econometric Theory*, **11**, 331-346.

Nelson, C.R., and Plosser, C.I. (1982), "Trends and Random Walks in Macroeconomic Time Series: Some Evidence and Implications," *Journal of Monetary Economics*, **10**, 139-162.

Ouliaris, S., and Phillips, P.C. B. (1994), *COINT2.0*, Maple Valley, WA: Aptech Systems.

Park, J. (2002), "An Invariance Principle for Sieve Bootstrap in Time Series," *Econometric Theory*, **18**, 469-490.

Phillips, P.C.B., and Durlauf, S. (1986), "Multiple Time Series Regression with Integrated Processes," *Review of Economic Studies*, **53**, 473-495.

——— and Hansen, B. (1990), "Statistical Inference in Instrumental Variables Regression with I(1) Processes," *Review of Economic Studies*, **57**, 99-125.

———(1987), "Time Series Regression with a Unit Root," *Econometrica*, **55**, 277-301.

———(1995), "Fully Modified Least Squares and Vector Autoregression," *Econometrica*, **63**, 1023-1078.

Pollard, D. (1984), *Convergence of Stochastic Processes*, Springer-Verlag, New York.

———(1991), "Asymptotics for Least Absolute Deviation Regression Estimators," *Econometric Theory*, **7**, 186-199.

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.

Rissanen, J. (1978), "Modelling by Shortest Data Description," *Automatica,* **14**, 465-471.

Rogers, A. (2001), "Least Absolute Deviations Regression under Nonstandard Conditions," *Econometric Theory*, **17**, 820-852.

Rothenberg, T., and Stock, J. (1997), "Inference in a Nearly Integrated Autoregressive Model with Nonnormal Innovations," *Jornal of Econometrics*, **80**, 269-286.

Said, S., and Dickey, D. (1984), "Testing for Unit Roots in Autoregressive-Moving Average Models of Unknown Order," *Biometrika*, **71**, 599-608.

Schotman, P., and van Dijk, H.K. (1991), "On the Bayesian Routes to Unit Root," *Journal of Applied Econometrics*, **6**, 387-402.

Schwarz, G. (1978), "Estimating the Dimension of a Model," *The Annals of Statistics*, **6**, 461-464.

Schwert, G.W. (1989), "Tests for Unit Roots: A Monte Carlo Investigation," *Journal of Business and Economic Statistics*, **7**, 147-159.

Siddiqui, M. (1960), "Distribution of Quantiles from a Bivariate Population," *Journal of Research of National Bureau of Standards*, **64B**, 145-150

Stock, J. (1995), "Unit Roots, Structural Breaks and Trends," in eds. R. Engle and McFadden, *Handbook of Econometrics*, pp. 2739-2841.

Thompson, S. (2001), "Robust Unit Root Testing with Correct Size and Good Power," working paper Harvard University.

Xiao, Z. (2001), "Likelihood Based Inference in Trending Time Series with a Root Near Unity," *Econometric Theory*, **17**, 1082-1112.

Weiss, A. (1987), "Estimating Nonlinear Dynamic Models Using Least Absolute Error Estimation,"
        Econometric Theory, **7**, 46-68.
Wooldridge, J., and White, H. (1988), "Some Invariance Principles and Central Limit Theorems for
        Dependent Heterogeneous Process," *Econometric Theory*, **4**, 210-230.