Unit Root Quantile Autoregression Inference

Roger KOENKER and Zhijie XIAO

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 rather 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 U.S. macroeconomic time series data further illustrates the potential of the new approach.

KEY WORDS: Brownian bridge; Kolmogorov–Smirnov tests; Quantile regression process.

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. Because 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 to use 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) derived the asymptotic distribution of the M estimator for an AR(1) process with a (near) unit root. Knight (1991) studied unit root M estimation in the case with infinite variance errors. Lucas (1995) considered 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 various forms of conditional heterogeneity. There is a considerable literature on quantile autoregression (QAR) 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) considered rank-type tests based on regression rank scores in an augmented Dickey–Fuller (ADF) framework.

In this article we propose new tests of the unit root hypothesis based on the quantile autoregression approach. We consider 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. Compared with 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 compared with 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. Finally, we give an application to U.S. macroeconomic data in Section 5.

A few words on notation. We use x_t to denote the vector of regressors that 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 written simply as $\int B$.

2. QUANTILE AUTOREGRESSION WITH A UNIT ROOT

2.1 The QAR(1) Model

We first consider the autoregression model

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

focusing on the case where $\alpha = 1$. For simplicity and without essential loss of generality, in this section we focus much of our attention on the first-order autoregression, 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 let $Q_u(\tau)$ denote the τth quantile of u_t and let $Q_{y_t}(\tau|y_{t-1})$ denote the τth conditional quantile of y_t conditional on y_{t-1}, then

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

Roger Koenker is Professor of Economics and Statistics (E-mail: *rkoenker@uiuc.edu*) and Zhijie Xiao is Associate Professor of Economics (E-mail: *zxiao@uiuc.edu*), University of Illinois, Urbana-Champaign, IL 61820. This research was supported in part by National Science Foundation grant SES-02-40781. The authors thank the editor, an associate editor, and three referees for their detailed and constructive comments.

Journal of the American Statistical Association
September 2004, Vol. 99, No. 467, Theory and Methods
DOI 10.1198/016214504000001114

Letting $\alpha_0(\tau) = Q_u(\tau)$, $\alpha_1(\tau) = \alpha$, 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 foregoing QAR model in the presence of a unit root.

Estimation of the linear QAR model involves solving the problem

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

where $\rho_\tau(u) = u(\tau - I(u < 0))$ as given by Koenker and Bassett (1978). We call solutions of (3), $\widehat{\alpha}(\tau)$, τth autoregression quantiles; viewed as a function of τ, we refer to $\widehat{\alpha}(\tau)$ as the QAR(1) process. Given $\widehat{\alpha}(\tau)$, the τth 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,

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

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 QAR process under the unit root hypothesis. Our analysis follows the approach of Knight (1991). (See Hercé 1996; Hasan and Koenker 1997; Hasan 2001, for related results.)

As will become clear in our later analysis, because of 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, whereas $\widehat{\alpha}_0(\tau)$ converges at rate $\sqrt{n}$. For this reason, we introduce the standardization matrix $D_n = \mathrm{diag}(\sqrt{n}, n)$ 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:

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

If $\widehat{v}$ is a minimizer of $Z_n(v) = \sum_{t=1}^{n}[\rho_\tau(u_{t\tau} - (D_n^{-1}v)^\top x_t) - \rho_\tau(u_{t\tau})]$, then 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) showed that if the finite-dimensional distributions of $Z_n(\cdot)$ converge weakly to those of $Z(\cdot)$ and $Z(\cdot)$ has a unique minimum, then 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 iid random variables with mean 0 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 QAR 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, thm. 2.1, pp. 474–476 and 486–489; Wooldridge and White 1988, cor. 4.2; Hansen 1992),

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

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, lem. 3.1; 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 (1992), it can be shown that the QAR 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, whereas 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), $B_\psi^\tau(r) = B_\psi(\tau, r) \sim \mathrm{N}(0, \tau(1-\tau)r)$.

Using the identity (A.1) given in the Appendix, the objective function of minimization problem (4) can be written as

$$\begin{aligned}
&\sum_{t=1}^{n} \big[\rho_\tau\big(u_{t\tau} - (D_n^{-1}v)^\top x_t\big) - \rho_\tau(u_{t\tau})\big] \\
&\quad = -\sum_{t=1}^{n} (D_n^{-1}v)^\top x_t \psi_\tau(u_{t\tau}) \\
&\qquad + \sum_{t=1}^{n} \int_0^{(D_n^{-1}v)^\top x_t} \{I(u_{t\tau} \le s) - I(u_{t\tau} < 0)\}\, ds.
\end{aligned}$$

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

Lemma 1. Let y_t be determined by (1) with $\alpha = 1$. Then, under Assumptions A1 and A2,

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

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

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 $\mathrm{N}(0,\tau(1-\tau))$. The limiting distribution of the QAR estimator for the unit root model is summarized in the following theorem.

Theorem 1. If y_t is determined by (1) with $\alpha=1$, then under Assumptions A1 and A2,

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

As an immediate consequence of this theorem, we have the following corollary, which is useful for constructing tests of the unit root hypothesis.

Corollary 1. Under the assumptions of Theorem 1,

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

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 ADF regression model (Dickey and Fuller 1979),

$$y_t=\alpha_1y_{t-1}+\sum_{j=1}^{q}\alpha_{j+1}\Delta y_{t-j}+u_t. \tag{6}$$

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 $\alpha_1=1$, then y_t contains a unit root and is persistent, and if $|\alpha_1|<1$, then y_t is stationary. Denoting the σ-field generated by $\{u_s, s\le t\}$ by $\mathcal{F}_t$, the τth conditional quantile of y_t, conditional on $\mathcal{F}_{t-1}$, is given by

$$Q_{y_t}(\tau|\mathcal{F}_{t-1})=Q_u(\tau)+\alpha_1y_{t-1}+\sum_{j=1}^{q}\alpha_{j+1}\Delta y_{t-j}.$$

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

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

Then 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 $\widehat{\alpha}(\tau)=(\widehat{\alpha}_0(\tau),\widehat{\alpha}_1,\dots,\widehat{\alpha}_p)$, $p=q+1$, and $D_n=\mathrm{diag}(\sqrt{n},n,\sqrt{n},\dots,\sqrt{n})$. Then the analysis of $\widehat{\alpha}(\tau)$ follows a similar procedure to that of first-order QAR 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 iid random variables with mean 0 and variance $\sigma^2<\infty$.

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

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

where

$$\underline{\Sigma}(\tau)=\begin{bmatrix}\sigma_w^2 & \sigma_{w\psi}(\tau)\\ \sigma_{w\psi}(\tau) & \sigma_\psi^2(\tau)\end{bmatrix}$$

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

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

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

Theorem 2. Let y_t be determined by (6), under Assumptions A1′ and A2 and the unit root assumption $\alpha_1=1$,

$$D_n\big(\widehat{\alpha}(\tau)-\alpha(\tau)\big)$$
$$\Rightarrow\frac{1}{f(F^{-1}(\tau))}\begin{bmatrix}\int_0^1\overline{B}_w\overline{B}_w^\top & 0_{2\times q}\\ 0_{q\times 2} & \Omega_\Phi\end{bmatrix}^{-1}\begin{bmatrix}\int_0^1\overline{B}_w\,dB_\psi^\tau\\ \Phi\end{bmatrix},$$

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

$$\Omega_\Phi=\begin{bmatrix}v_0 & \cdots & v_{q-1}\\ \vdots & \ddots & \vdots\\ v_{q-1} & \cdots & v_0\end{bmatrix},\qquad v_j=E[w_tw_{t-j}],$$

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

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

Corollary 2. Under the assumptions of Theorem 2,

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

In particular,

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

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

3. INFERENCE ON THE QUANTILE AUTOREGRESSION PROCESS

Inference based on the QAR process provides a more robust approach to testing the unit root hypothesis. Like the conventional ADF t-ratio test, we consider the t-ratio statistic

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

where $\widehat{f(F^{-1}(\tau))}$ 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 $X = (1, \Delta y_{t-1}, \ldots, \Delta y_{t-q})$. Under the unit root hypothesis, we have, using the results in the previous section,

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

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 $(\sigma_w^2, \sigma_{w\psi}(\tau))$, because B_w and B_ψ^τ are correlated Brownian motions. The foregoing limiting distribution is similar to distributions appearing in various unit root tests using other methods. In particular, similar limiting distributions were used by Lucas (1995) for unit root tests based on his nonparametric modified M estimators, Hasan and Koenker (1997) for their unmodified statistic S_T based on rank scores, and, as we show later, Hansen (1995) for his least squares-based covariate ADF 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 in the following analysis. A third approach would be 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.

Unit root tests may be constructed based on a QAR 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.

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

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

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

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

and

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

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

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

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-ADF (CADF) test of Hansen (1995). Tables of critical values for the statistic $t_n(\tau)$ have been provided by Hansen (1995, p. 1155) for values of δ^2 in steps of .1. For intermediate values of δ^2, Hansen suggest using critical values obtained by interpolation. 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. For convenience, we give the table of critical values from Hansen in the Appendix A.1. In practice, to use the correct critical values from this table, we estimate δ^2

by $\widehat{\delta}^2 = \widehat{\sigma}^2_{w\psi}(\tau)/[\tau(1-\tau)\widehat{\sigma}^2_w]$ and then use the estimated $\widehat{\delta}^2$ to select the appropriate row from the table; see Section 3.4 for further details.

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. Let $w_t = \Delta y_t$, $t = 2, \ldots, n$, then fit the following qth order autoregression by ordinary least squares (OLS):

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

and obtain estimates $\widehat{\beta}_1, \ldots, \widehat{\beta}_q$, and the residuals $\widehat{u}_t$. We may also use the Yule–Walker method, which is asymptotically equivalent to the OLS method, to estimate the autoregression.

2. Draw iid 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

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

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

3. Generate y_t^* under the null restriction of a unit root, $y_t^* = y_{t-1}^* + w_t^*$, with $y_1^* = y_1$.
4. Estimate the following pth-order QAR regression:

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

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

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

In the foregoing 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; that is, the weak convergence of the bootstrap partial sum process to Brownian motion that holds almost surely for all sample realizations. Park (2002) established a bootstrap-invariance principle for sieve bootstrap that allows the autoregressive lag length q to go to infinity. In later work, Chang, Park, and Song (2001) developed a multivariate bootstrap invariance principle. The theory that they derived in the vector time series model can be used here to derive a bivariate bootstrap-invariance principle that validates the foregoing resampling procedure. (For other versions of bootstrap-invariance principles, see, e.g., Kinateder 1992; Ferretti and Romo 1995; van 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, that is,

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

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

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

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 quantities

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

Because 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 coefficient-based statistic

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

Under the unit root hypothesis and our assumptions,

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

At fixed τ, the test statistic $U_n(\tau)$ is the quantile regression counterpart of the CADF test. Like the t-ratio statistic, the limiting distribution of $U_n(\tau)$ is not standard and depends on nuisance parameters. We can consider similar options as those that 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 $[\int_0^1 \underline{B}_w^2]^{-1}\int_0^1 \underline{B}_w\, dB_\psi^\tau$ 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 on only 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:

$$\mathrm{QKS}_\alpha = \sup_{\tau\in\mathcal{T}} |U_n(\tau)|, \qquad \mathrm{QKS}_t = \sup_{\tau\in\mathcal{T}} |t_n(\tau)| \qquad (13)$$

and

$$\mathrm{QCM}_\alpha = \int_{\tau\in\mathcal{T}} U_n(\tau)^2\, d\tau, \qquad \mathrm{QCM}_t = \int_{\tau\in\mathcal{T}} t_n(\tau)^2\, d\tau. \quad (14)$$

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

$$\sup_{\tau \in \mathcal{T}} |U(\tau)|, \qquad \sup_{\tau \in \mathcal{T}} |t(\tau)|,$$

and

$$\int_{\tau \in \mathcal{T}} U(\tau)^2 \, d\tau, \qquad \int_{\tau \in \mathcal{T}} t(\tau)^2 \, d\tau.$$

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 foregoing procedure, and replace step 4 by the foloowing:

4′. We estimate the pth-order QAR 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_ith QAR 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 $t_n^*(\tau_i)$:

$$\mathrm{QKS}_t^* = \max_{i \in I} |t_n^*(\tau_i)|,$$

$$\mathrm{QCM}_t^* = \sum_{i \in I} t_n^*(\tau_i)^2 (\tau_i - \tau_{i-1})$$

and

$$\mathrm{QKS}_\alpha^* = \max_{i \in I} |U_n^*(\tau_i)|,$$

$$\mathrm{QCM}_\alpha^* = \sum_{i \in I} U_n^*(\tau_i)^2 (\tau_i - \tau_{i-1}).$$

The limiting null distribution of the test statistics can again be approximated by repeating steps 2–4′. Let $C_{\mathrm{KS}_\alpha}(\theta)$, $C_{\mathrm{KS}_t}(\theta)$, $C_{\mathrm{CM}_\alpha}(\theta)$, and $C_{\mathrm{CM}_t}(\theta)$ be the (100θ)th quantiles, that is,

$$\mathrm{P}^*[\mathrm{QKS}_\alpha^* \le C_{\mathrm{KS}_\alpha}(\theta)] = \mathrm{P}^*[\mathrm{QKS}_t^* \le C_{\mathrm{KS}_t}(\theta)] = \theta$$

and

$$\mathrm{P}^*[\mathrm{QCM}_\alpha^* \le C_{\mathrm{CM}_\alpha}(\theta)] = \mathrm{P}^*[\mathrm{QCM}_t^* \le C_{\mathrm{CM}_t}(\theta)] = \theta.$$

Then the unit root hypothesis will be rejected at the $(1-\theta)$ level if, say, $\mathrm{QKS}_\alpha > C_{\mathrm{KS}_\alpha}(\theta)$.

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 work by Siddiqui (1960) and 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 that tends toward 0 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 by Bassett and Koenker (1982),

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

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

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

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 the sample size $n \to \infty$. The quantities $C_{ww}(h)$ and $C_{w\psi}(h)$ are sample covariances defined by $C_{ww}(h) = n^{-1}\sum' w_t w_{t+h}$ and $C_{w\psi}(h) = n^{-1}\sum' w_t \times \psi_\tau(\widehat{u}_{(t+h)\tau})$, where $\sum'$ signifies summation over $1 \le t$, $t+h \le n$. Candidate kernel functions can be found in standard texts (e.g., Hannan 1973).

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 1984; 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 QAR-based procedures and compare them with 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 iid random variables. We examined different values of α and different types of error distributions in the experiment. In particular, we consider four values of α: 1.0, .95, .9, .85. In addition, we also considered an alternative with random coefficient $\alpha = \alpha_t = \min\{.5 + \Phi(u_t), 1\}$, 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.

We consider both normal and nonnormal disturbances 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 iid

Table 1. Size and Power, Case With Gaussian Innovations

α	ADF_t	ADF_α	Z_t	Z_α	HK_W	HK_N	HK_S	QKS	QCM	$t_n(\tau)$	$t_n^*(\tau)$
$n = 100$											
1	.061	.056	.070	.060	.043	.036	.051	.05	.06	.122	.06
.95	.121	.134	.130	.139	.101	.096	.084	.110	.120	.194	.11
.90	.246	.261	.238	.252	.241	.252	.173	.211	.230	.281	.196
.85	.581	.680	.611	.672	.482	.468	.318	.486	.514	.578	.430
α_t	.52	.665	.535	.656	.093	.163	.024	.850	.780	.150	.250
$n = 200$											
1	.06	.05	.07	.06	.054	.053	.044	.06	.05	.106	.06
.95	.359	.391	.348	.372	.277	.288	.218	.289	.350	.391	.26
.90	.82	.922	.84	.906	.765	.783	.493	.73	.78	.78	.64
.85	1	1	.99	1	.968	.977	.747	.90	.95	.94	.88
α_t	.920	.935	.918	.936	.221	.431	.018	.975	.990	.051	.064

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; and (4) u_t are Student-t-distributed variables with 4 degrees of freedom. The last three disturbances have heavy-tailed 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: (1) the t-ratio test $t_n(\tau)$ based on QAR at $\tau = .5$, using the critical values in Appendix A.1; (2) the t-ratio test $t_n^*(\tau)$ based on QAR at $\tau = .5$, using the bootstrapped critical values; (3) the KS-type test (QKS_α) based on QAR with $\mathcal{T} = [.1, .9]$, using the bootstrapped critical values; (4) the CM-type test (QCM_α) with $\mathcal{T} = [.1, .9]$ and the bootstrapped critical values; (5) the unmodified rank test of Hasan and Koenker (1997) using the Wilcoxon score function; (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; and (11) the Phillips–Perron semiparametric Z_α test. 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.

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 1. 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 by Thompson (2001). The first four tests are based on the QAR model proposed in this article, the tests 5–7 are rank tests using different scores, and the last four tests are OLS-based procedures used widely in applications. The order of ADF regressions 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 CM statistic is calculated using numerical integration with step size equals .01, and the 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 2,000. For each test, the number of repetition is 1,000. Two sample sizes are studied, $n = 100$ and 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 Appendix A.1 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 KS- or CM-type tests based on the quantile regression process for $\tau \in \mathcal{T}$ have the highest power over all procedures.

Tables 2, 3, and 4 report results for cases where the errors have Student-t innovations with degrees of freedom 2, 3, and 4.

Table 2. Size and Power, Case With t(2) Innovations

α	ADF_t	ADF_α	Z_t	Z_α	HK_W	HK_N	HK_S	QKS	QCM	$t_n(\tau)$	$t_n^*(\tau)$
$n = 100$											
1	.062	.048	.066	.058	.045	.040	.050	.054	.051	.058	.052
.95	.114	.173	.127	.193	.505	.392	.504	.365	.444	.526	.38
.9	.334	.530	.356	.561	.808	.723	.743	.698	.782	.834	.730
.85	.616	.779	.622	.781	.930	.886	.880	.815	.909	.961	.872
$n = 200$											
1	.08	.06	.08	.06	.056	.062	.060	.061	.053	.057	.052
.95	.29	.36	.30	.36	.897	.835	.843	.813	.880	.922	.870
.9	.89	.92	.90	.912	1.000	.998	.991	.982	.996	.999	.991
.85	.97	1	.97	.99	1	1	.998	.995	1	1	1

Table 3. Size and Power, Case With t(3) Innovations

α	ADF_t	ADF_α	Z_t	Z_α	HK_W	HK_N	HK_S	QKS	QCM	$t_n(\tau)$	$t_n^*(\tau)$
$n = 100$											
1	.056	.048	.057	.057	.048	.055	.044	.053	.052	.078	.054
.95	.126	.203	.147	.211	.244	.191	.258	.180	.240	.330	.232
.9	.37	.45	.38	.46	.552	.443	.495	.420	.600	.62	.47
.85	.54	.68	.55	.68	.804	.723	.702	.610	.840	.85	.70
$n = 200$											
1	.06	.07	.07	.06	.059	.048	.050	.06	.05	.08	.05
.95	.30	.43	.31	.44	.641	.535	.528	.61	.70	.79	.67
.9	.85	.94	.87	.95	.977	.960	.892	.95	.97	.98	.95
.85	.99	1	.99	1	1	1	.986	.99	1	1	1

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 the 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 QAR-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 Appendix A.1. However, $t_n(\tau)$ using the critical values in Appendix A.1 has the best power in the presence of heavy-tailed innovations. For the comparison between the two QAR-based tests using information over $\mathcal{T} = [.1, .9]$, the CM-type test (QCM) has relatively better finite sample performance than the KS-type test (QKS).

A final comparison can be made between the proposed QAR-based tests and the rank-type tests. The QAR-based tests proposed in this article are constructed from the unrestricted QAR, whereas 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 processes. In the presence of random coefficient alternative $\alpha = \alpha_t$, the KS- or CM-type test proposed in this article 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. U.S. INTEREST RATE DYNAMICS

We now apply the QAR-based unit root tests to several U.S. 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 U.S. annual interest rates in their seminal work on 14 macroeconomic time series. This series and other types of interest rates have often been reexamined. Among various empirical findings, two important features have been well documented:

1. 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, Lindberg, and Perraudin 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 1 month, 3 month, and annual interest rates in the U.S. In particular, we looked at (seasonally adjusted) 1-month and 3-month commercial paper rates, and the annual bond yield from the extended Nelson–

Table 4. Size and Power, Case With t(4) Innovations

α	ADF_t	ADF_α	Z_t	Z_α	HK_W	HK_N	HK_S	QKS	QCM	$t_n(\tau)$	$t_n^*(\tau)$
$n = 100$											
1	.066	.070	.068	.071	.058	.043	.060	.047	.056	.074	.057
.95	.137	.205	.146	.224	.173	.134	.160	.190	.250	.270	.172
.9	.352	.501	.355	.512	.475	.407	.393	.510	.550	.575	.474
.85	.68	.77	.69	.78	.714	.643	.564	.76	.79	.83	.70
$n = 200$											
1	.055	.057	.057	.062	.051	.051	.047	.051	.052	.078	.061
.95	.31	.47	.33	.46	.486	.397	.407	.51	.60	.65	.55
.9	.87	.91	.88	.90	.938	.899	.805	.92	.93	.95	.92
.85	.97	.99	.99	1	.995	.994	.952	.98	1	.99	.97

Table 5. Descriptive Statistics

	1-month rate	3-month rate	Annual rate
Skewness	−1.8435	−1.543	−.1410
Kurtosis	28.062	24.41	3.88
Jarque–Bera	9759	7117	2.80

Plosser data. Both the 1-month rate and the 3-month rate start on April 1971 and end in June 2002, with 378 observations. The annual data are from 1900 to 1988. We first apply the ADF unit root tests to these series. In the ADF regressions, the Bayesian information 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 autoregressive coefficients in Tables 6, 8, and 10). Tables 6, 8, and 10 report the ADF test statistics for the 1-month, 3-month, and annual series. The unit root hypothesis can not be rejected by the traditional ADF test at the 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.

We reconsider these series using the proposed QAR methods and report the results in Tables 6–11. In particular, we applied the following four tests to the interest rate series: (1) the KS-type test QKS_α based on quantile autoregression with $\mathcal{T} = [.1, .9]$, using the bootstrapped critical values; (2) the CM-type test QCM_α with $\mathcal{T} = [.1, .9]$ 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 Appendix A.1; (4) the coefficient-based tests $U_n(\tau)$ based on QAR 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 fourth tests try to provide a more detailed examination on the unit root property of these series at each decile.

Tables 6, 8, and 10 report the QKS_α and QCM_α tests for the three time series. 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 CM-type test at 5% level, but not rejected by the KS-type 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.

Table 7. QAR Results for 1-Month Rate

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%
.1	.886	−4.65	.142	−41.4**	−22.75	−18.06	2.88	4.19
.2	.929	−4.74	.184	−26.4**	−12.31	−10.11	1.62	2.40
.3	.961	−4.20	.201	−14.3**	−7.48	−5.94	1.05	1.55
.4	.981	−2.44	.248	−7.06*	−3.69	−3.10	.49	.75
.5	.994	−.75	.177	−2.08	−3.61	−2.85	.49	.74
.6	1.014	1.84	.160	5.39#	−5.65	−4.51	.76	1.17
.7	1.029	3.18	.163	11.13##	−7.75	−6.30	1.12	1.67
.8	1.055	3.82	.137	20.47##	−11.17	−9.14	1.67	2.36
.9	1.111	5.55	.116	41.39##	−18.83	−15.11	2.49	3.73

NOTE: 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 $H_{1B}: \alpha_1 > 1$, and those with an (##) are significant at the 1% level when the alternative is H_{1B}.

Tables 7, 9, and 11 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 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 the 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 over a longer period. Consequently, interest rates are more persistent in the presence of positive shocks than to negative ones.

The third columns in Tables 7, 9, and 11 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. The sixth and seventh columns of these tables report 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.

Table 6. Unit Root Tests for 1-Month Rate

	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	$\widehat{\alpha}_1 = .976$			

*Significant at the 1% level with alternative H_{1A}.

Table 8. Unit Root Tests for 3-Month Rate

	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	$\widehat{\alpha}_1 = .977$			

*Significant at the 1% level with alternative H_{1A}.

Table 9. QAR Results for 3-Month Rate

Quan-tiles	$\widehat{\alpha}_1(\tau)$	$t_n(\tau)$	$\widehat{\delta}^2$	$U_n(\tau)$	Critical values for $U_n(\tau)$ 2.5%	5%	95%	97.5%
.1	.884	−4.74	.189	−43.03	−19.99	−16.51	2.54	3.88
.2	.926	−5.52	.200	−27.55	−9.45	−7.62	1.24	1.84
.3	.959	−4.30	.165	−15.22**	−6.59	−5.50	1.03	1.47
.4	.984	−1.71	.188	−5.83	−4.39	−3.50	.62	.86
.5	.991	−.76	.161	−3.51	−4.03	−3.26	.51	.76
.6	1.012	1.65	.179	4.82	−5.53	−4.48	.76	1.19
.7	1.034	3.23	.136	12.62	−7.79	−6.44	1.06	1.52
.8	1.065	4.76	.185	23.91	−10.95	−8.78	1.60	2.21
.9	1.107	6.35	.118	39.73	−20.41	−15.99	2.93	4.13

NOTE: The values of $U_n(\tau)$ denoted by an (**) are significant at the 1% level with alternative $H_{1A}:\alpha_1 < 1$.

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 7, 9, and 11 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 fourth columns of these tables. Most of the results reject the unit root null hypothesis. 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 7, 9, and 11, 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, then 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 1, displaying mildly explosive behavior. Combining this evidence with the results of Tables 6, 8, and 10, 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 squares-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 squares method when the data is non-Gaussian or is contaminated by outliers.

Table 10. Unit Root Tests for Annual Rate

	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		$\widehat{\alpha}_1 = .974$		

*Significant at the 1% level with alternative H_{1A}.

Table 11. QAR Results for Annual Rate

Quan-tiles	$\widehat{\alpha}_1(\tau)$	$t_n(\tau)$	$\widehat{\delta}^2$	$U_n(\tau)$	Critical values for $U_n(\tau)$ 2.5%	5%	95%	97.5%
.1	.829	−5.47	.142	−14.64	−26.89	−22.13	3.33	4.95
.2	.865	−4.33	.101	−11.61	−12.12	−9.15	1.24	2.02
.3	.965	−2.42	.200	−2.99	−6.00	−4.50	.89	1.25
.4	.981	−1.01	.120	−1.64	−5.24	−4.12	.92	1.22
.5	1.004	.06	.168	.39	−5.52	−4.40	.90	1.35
.6	1.052	1.52	.268	4.44	−6.88	−5.48	1.17	1.74
.7	1.126	4.02	.163	10.84	−9.53	−7.69	1.61	2.45
.8	1.165	4.79	.194	14.25	−13.32	−11.11	2.22	3.36
.9	1.126	3.76	.257	10.83	−17.88	−14.65	3.06	4.75

APPENDIX: TABLES AND PROOFS

A.1 Asymptotic Critical Values of the t-Statistic $t_n(\tau)$ Given by (10)

δ^2	1%	5%	10%
.9	−3.39	−2.81	−2.50
.8	−3.36	−2.75	−2.46
.7	−3.30	−2.72	−2.41
.6	−3.24	−2.64	−2.32
.5	−3.19	−2.58	−2.25
.4	−3.14	−2.51	−2.17
.3	−3.06	−2.40	−2.06
.2	−2.91	−2.28	−1.92
.1	−2.78	−2.12	−1.75

A.2 Proof of Theorems 1 and 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 $\psi_\tau(u) = \tau - I(u < 0)$, then, for $u \neq 0$,

$$\begin{aligned}
&\rho_\tau(u-v) - \rho_\tau(u) \\
&\quad = -v\psi_\tau(u) + (u-v)\{I(0 > u > v) - I(0 < u < v)\} \\
&\quad = -v\psi_\tau(u) + \int_0^v \{I(u \le s) - I(u < 0)\}\,ds. \qquad \text{(A.1)}
\end{aligned}$$

Let

$$u_{t\tau} = y_t - x_t'\alpha(\tau) = u_t - F_u^{-1}(\tau). \qquad \text{(A.2)}$$

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

$$Q_{u_{t\tau}}(\tau|\mathcal{F}_{t-1}) = 0. \qquad \text{(A.3)}$$

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

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

Using identity (A.1), we have

$$\begin{aligned}
&\sum_{t=1}^n \big[\rho_\tau(u_{t\tau} - v'D_n^{-1}x_t) - \rho_\tau(u_{t\tau})\big] \\
&\quad = -\sum_{t=1}^n v'D_n^{-1}x_t\psi_\tau(u_{t\tau})
\end{aligned}$$

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

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

$$\begin{bmatrix}\frac{1}{\sqrt{n}}\sum_t\Delta y_{t-1}\psi_\tau(u_{t\tau})\\ \vdots\\ \frac{1}{\sqrt{n}}\sum_t\Delta y_{t-q}\psi_\tau(u_{t\tau})\end{bmatrix}. \tag{A.4}$$

If we denote $E[w_t w_{t-j}]=\nu_j$, we can show that (A.4) converges to a q-dimensional normal variate $\Phi=[\Phi_1,\ldots,\Phi_q]^\top$ 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

$$U_n(v)=\sum_{t=1}^{n}z_t(v),$$

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 $v_2n^{-1/2}y_{t-1}$ at some finite number $m>0$ and denote

$$U_{nm}(v)=\sum_{t=1}^{n}z_{tm}(v),$$
$$z_{tm}(v)=(v'D_n^{-1}x_t-u_{t\tau})I(0<u_{t\tau}<v'D_n^{-1}x_t)M_t,$$

and

$$M_t=I(0\le v_2n^{-1/2}y_{t-1}\le m).$$

We further define

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

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

$$\begin{aligned}
&\overline{U}_{nm}(v)\\
&=\sum_{t=1}^{n}E\left\{(v'D_n^{-1}x_t-u_{t\tau})I(0<u_{t\tau}<v'D_n^{-1}x_t)M_t|\mathcal{F}_{t-1}\right\}\\
&=\sum_{t=1}^{n}\int_{F_u^{-1}(\tau)}^{[v'D_n^{-1}x_t+F_u^{-1}(\tau)]M_t}\left[\left(v'D_n^{-1}x_t+F_u^{-1}(\tau)\right)M_t-r\right]\\
&\qquad\times f_u(r)\,dr\\
&=\sum_{t=1}^{n}\int_{F_u^{-1}(\tau)}^{[v'D_n^{-1}x_t+F_u^{-1}(\tau)]M_t}\left[\int_r^{[v'D_n^{-1}x_t+F_u^{-1}(\tau)]M_t}ds\right]f_u(r)\,dr\\
&=\sum_{t=1}^{n}\int_{F_u^{-1}(\tau)}^{[v'D_n^{-1}x_t+F_u^{-1}(\tau)]M_t}\left[\int_r^{[v'D_n^{-1}x_t+F_u^{-1}(\tau)]M_t}f_u(r)\,ds\right]dr\\
&=\sum_{t=1}^{n}\int_{F_u^{-1}(\tau)}^{[v'D_n^{-1}x_t+F_u^{-1}(\tau)]M_t}\left[\int_{F_u^{-1}(\tau)}^{s}f_u(r)\,dr\right]ds\\
&=\sum_{t=1}^{n}\int_{F_u^{-1}(\tau)}^{[v'D_n^{-1}x_t+F_u^{-1}(\tau)]M_t}\left[\int_{F_u^{-1}(\tau)}^{s}f_u(r)\,dr\right]ds\\
&=\sum_{t=1}^{n}\int_{F_u^{-1}(\tau)}^{[v'D_n^{-1}x_t+F_u^{-1}(\tau)]M_t}[s-F_u^{-1}(\tau)]\\
&\qquad\times\left[\frac{F_u(s)-F_u(F_u^{-1}(\tau))}{s-F_u^{-1}(\tau)}\right]ds.
\end{aligned}$$

Under Assumption A2,

$$\begin{aligned}
\overline{U}_{nm}(v)&=\sum_{t=1}^{n}\int_{F_u^{-1}(\tau)}^{[v'D_n^{-1}x_t+F_u^{-1}(\tau)]M_t}[s-F_u^{-1}(\tau)]f_u[F_u^{-1}(\tau)]\,ds\\
&\quad+o_p(1)\\
&=\sum_{t=1}^{n}f_u[F_u^{-1}(\tau)]\left\{\frac{[s-F_u^{-1}(\tau)]^2}{2}\Bigg|_{F^{-1}(\tau)}^{[v'D_n^{-1}x_t+F^{-1}(\tau)]M_t}\right\}\\
&\quad+o_p(1)\\
&=\frac{1}{2}\sum_{t=1}^{n}f_u[F^{-1}(\tau)]v'[D_n^{-1}x_tx_t'D_n^{-1}]vM_t+o_p(1).
\end{aligned}$$

Thus

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

where

$$\Psi_{1m}=\begin{bmatrix}\int_0^1\overline{B}_w\overline{B}_w'I(0\le v_2'B_w(s)\le m) & 0_q'\\ 0_q' & \Omega_\Phi\end{bmatrix}.$$

We now follow the arguments of Pollard (1984, p. 171). Noting that

$(v'D_n^{-1}x_t)I(0 \le v_2 n^{-1/2} y_{t-1} \le m) \xrightarrow{P} 0$ uniformly in t, we have

$$\sum_{t=1}^{n} E\big[z_{tm}(v)^2 | \mathcal{F}_{t-1}\big] \le \max\big\{(v'D_n^{-1}x_t)I(0 \le v_2 n^{-1/2} y_{t-1} \le m)\big\} \sum \bar{z}_{tm}(v) \xrightarrow{P} 0.$$

Thus the summation of martingale difference sequence

$$\sum_t \{z_{tm}(v) - \bar{z}_{tm}(v)\}$$

converges to 0 in probability. Therefore, the limiting distribution of $\sum_t z_{tm}(v)$ is the same as that of $\sum_t \bar{z}_{tm}(v)$, that is,

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

Letting $m \to \infty$, we have

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

and

$$\Psi_1 = \begin{bmatrix} \int_0^1 \overline{B}_w \overline{B}_w' I(0 \le v_2' B_w(s)) & 0_q' \\ 0_q' & \Omega_\Phi \end{bmatrix}.$$

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

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

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

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

with

$$\Psi_2 = \begin{bmatrix} \int_0^1 \overline{B}_w \overline{B}_w' I(v_2' B_w(s) \le 0) & 0_q' \\ 0_q' & \Omega_\Phi \end{bmatrix}.$$

Thus

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

where

$$\Psi = \begin{bmatrix} \int_0^1 \overline{B}_w \overline{B}_w' & 0_q' \\ 0_q' & \Omega_\Phi \end{bmatrix}.$$

As a result,

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

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

$$\frac{1}{f(F^{-1}(\tau))} \begin{bmatrix} \int_0^1 \overline{B}_w \overline{B}_w^\top & 0_{2\times q} \\ 0_{q\times 2} & \Omega_\Phi \end{bmatrix}^{-1} \begin{bmatrix} \int_0^1 \overline{B}_w \, dB_\psi^\tau \\ \Phi \end{bmatrix}.$$

By lemma A of Knight (1989), we have

$$D_n\big(\widehat{\alpha}(\tau) - \alpha(\tau)\big) \Rightarrow \frac{1}{f(F^{-1}(\tau))} \begin{bmatrix} \int_0^1 \overline{B}_w \overline{B}_w^\top & 0_{2\times q} \\ 0_{q\times 2} & \Omega_\Phi \end{bmatrix}^{-1} \begin{bmatrix} \int_0^1 \overline{B}_w \, dB_\psi^\tau \\ \Phi \end{bmatrix}.$$

Portnoy (1984) showed 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 $\tau \in \mathcal{T}$.

[Received December 2002. Revised January 2004.]

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," *Proceeding of American Mathematical Society*, 12, 788–792.

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

Chan, N. H., and Wei, C. Z. (1987), "Asymptotic Inference for Nearly Nonstationary AR(1) Processes," *The 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," *The Annals of Statistics*, 19, 1109–1128.

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. M. Dempster and S. Pliska, Cambridge, U.K.: Cambridge University Press, pp. 1–35.

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

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

Hannan, E. J. (1973), "Central Limit Theorems for Time Series Regression," *Zeitschrift für Wahrsheinlichkeitstheorie und verwandte 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.

Hasan, M. N., and Koenker, R. (1997), "Robust Rank Tests of the Unit Root Hypothesis," *Econometrica*, 65, 133–161.

Hercé, 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.

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

Kinateder, J. (1992), "An Invariance Principle Applicable to the Bootstrap," in *Exploring the Limits of Bootstrap*, eds. R. LePage and L. Billard, New York: Wiley.

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

——— (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 Mukherjee, R. (1994), "Regression Quantiles and Related Processes Under Long Range Dependent Errors," *Journal of Multivariate Analysis*, 51, 318–337.

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. (1987), "Time Series Regression With a Unit Root," *Econometrica*, 55, 277–301.

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

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

Phillips, P. C. B., and Hansen, B. (1990), "Statistical Inference in Instrumental Variables Regression With I(1) Processes," *Review of Economic Studies*, 57, 99–125.

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

——— (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, New York: Springer-Verlag, pp. 231–246.

Rissanen, J. (1978), "Modeling 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," *Journal 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 & 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 *Handbook of Econometrics*, Vol. 4, eds. R. F. Engle and D. L. McFadden, New York: Elsevier, pp. 2739–2841.

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

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

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.

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