

QUANTILE AUTOREGRESSION, UNIT ROOTS AND ASYMMETRIC INTEREST RATE DYNAMICS

ROGER KOENKER AND ZHIJIE XIAO

ABSTRACT. We consider quantile autoregression models in which the autoregressive coefficients may vary over the unit interval. The models can capture systematic influences of conditioning variables on the location, scale and shape of the conditional distribution of the response, and therefore constitute a significant extension of classical constant coefficient linear time series models in which the effect of conditioning is confined to a location shift. The models may be interpreted as a special case of the general random coefficient autoregression model with strongly dependent coefficients. The statistical properties of the model and the associated estimators are studied. The limiting distributions of the autoregression quantile process are derived for both stationary versions of the model and for a unit root form of the model. Inference methods for both variants of the model are also investigated. An empirical application of the model to US short-term interest rate data displays asymmetric interest rate dynamics.

1. INTRODUCTION

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

In this paper we seek to relax this restriction and consider linear quantile autoregression models whose autoregressive (slope) parameters may vary with $\tau \in [0, 1]$. We were initially motivated to explore these models in the hope that they might expand the modeling options for economic time series that displayed “unit-root behavior”. We will show that some forms of the model can exhibit unit-root like tendencies with occasional episodes of mean reversion sufficient to insure stationarity. The models

Version: December 13, 2002. Corresponding Author: Roger Koenker, Department of Economics, University of Illinois at Urbana-Champaign, 1206 S. Sixth St., Champaign, IL 61820, USA. The authors wish to thank the NSF for financial support.

lead to interesting new hypotheses and inference apparatus for both stationary and non-stationary time series.

2. THE MODEL

There is a substantial theoretical literature, including Weiss (1987), Knight (1989), Koul and Saleh(1995), Koul and Mukherjee(1994), Hercé (1996), Hasan and Koenker (1997), Jurečková and Hallin (1999), dealing with the linear quantile autoregression model. In this model the τ th conditional quantile function of the response y_t is expressed as a linear function of lagged values of the response. But a striking feature of this literature is that it has focused exclusively on the case of iid innovations in which the conditioning variables play their classical role of shifting the *location* of the conditional density of y_t , but they have no effect on conditional scale or shape. In this paper we wish to study estimation and inference in a more general class of quantile autoregressive (QAR) models in which all of the autoregressive coefficients are allowed to be τ -dependent, and therefore capable of altering the location, scale and shape of the conditional densities. We will write the general form of the model as

$$Q_{y_t}(\tau|y_{t-1}, \dots, y_{t-p}) = \alpha_0(\tau) + \alpha_1(\tau)y_{t-1} + \dots + \alpha_p(\tau)y_{t-p}.$$

or somewhat more compactly as,

$$(2.1) \quad Q_{y_t}(\tau|\mathcal{F}_{t-1}) = x_t^\top \alpha(\tau).$$

where \mathcal{F}_t denotes the σ -field generated by $\{y_s, s \leq t\}$, and $x_t = (1, y_{t-1}, \dots, y_{t-p})^\top$.

To motivate the model we will emphasize the simplest, first-order version of it,

$$(2.2) \quad Q_{y_t}(\tau|\mathcal{F}_{t-1}) = \alpha_0(\tau) + \alpha_1(\tau)y_{t-1}.$$

To fix ideas, it may be useful to observe that since $Q_{y_t}(\tau|\mathcal{F}_{t-1})$ is the conditional quantile function, the model may also be expressed as

$$y_t = \alpha_0(U_t) + \alpha_1(U_t)y_{t-1}.$$

where the U_t are taken to be uniformly distributed on the unit interval and iid. The classical Gaussian AR(1) model is obtained by setting $\alpha_0(u) = \sigma\Phi^{-1}(u)$ and $\alpha_1(u) = \alpha_1$, a constant. The formulation in (2) reveals that the model may be interpreted as rather unusual form of random coefficient autoregressive (RCAR) model. Such models arise naturally in many time series applications. Discussions of the role of RCAR models can be found in, inter alia, Nicholls and Quinn (1982), Tjøstheim(1986), Pourahmadi (1986), Brandt (1986), Karlsen(1990), and Tong (1990). In contrast to most of the literature on RCAR models, in which the coefficients are assumed to be stochastically independent of one another, the QAR model has coefficients that are functionally dependent. Since monotonicity is required of the quantile functions we will see that this imposes some discipline on the forms taken by the α functions. This discipline essentially requires that the vector $\alpha(\tau)$, or some affine transformation of

it, be monotonic in each coordinate. This condition insures that the random vector $\alpha_t = \alpha(U_t)$ is comonotonic, as will be elaborated in Section 3.

We are particularly interested in exploring the ability of the QAR model to account for unit-root like behavior in economic time series. And we will argue that QAR models can play a useful role in expanding the modeling space between classical stationary linear time series models and their unit root alternatives. To illustrate this with the QAR(1) model, consider the model (2) with $\alpha_0(\tau) = \sigma\Phi^{-1}(\tau)$ and $\alpha_1(\tau) = \max\{\gamma_0 + \gamma_1\tau, 1\}$ for $\gamma_0 \in (0, 1)$ and $\gamma_1 > 0$. In this model if $u_t > (1 - \gamma_0)/\gamma_1$ the model generates the y_t according to the unit root model, but for sufficiently small realizations of u_t we have a mean reversion tendency. Thus the model exhibits a form of asymmetric persistence in the sense that sequences of strongly positive innovations tend to reinforce its unit root like behavior, while occasional negative realizations induce mean reversion and thus undermine the persistence of the process. We will have more to say about these phenomena in Section 6, when we discuss an application of the QAR model to interest rate dynamics.

Estimation of the linear quantile autoregressive model involves solving the problem

$$(2.3) \quad \min_{\alpha \in \mathbb{R}^{p+1}} \sum_{t=1}^n \rho_{\tau}(y_t - x_t^{\top} \alpha),$$

where $\rho_{\tau}(u) = u(\tau - I(u < 0))$ as in Koenker and Bassett (1978). Solutions, $\hat{\alpha}(\tau)$ will be call autoregression quantiles. Given $\hat{\alpha}(\tau)$, the τ -th conditional quantile function of y_t , conditional on past information, can be estimated by,

$$\hat{Q}_{y_t}(\tau|x_{t-1}) = x_t^{\top} \hat{\alpha}(\tau),$$

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

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

for some appropriately chosen sequence of τ 's.

3. QAR(1) MODEL

In this section we briefly describe some essential features of the QAR(1) model and some associated estimation methods, first for the stationary version of the model, and then for the unit root case.

3.1. Stationary QAR(1). We begin by reviewing some basic properties of the random coefficient AR(1) model. Let $\{u_t\}$ be a sequence of *iid* random variables with mean 0 and variance $\sigma^2 < \infty$. We consider the following process:

$$(3.1) \quad y_t = \alpha_t y_{t-1} + u_t,$$

where α_t is a function of u_t . We are interested in cases that the random autoregressive coefficient α_t satisfies $\alpha_t \leq 1$, but, as will become clear later in this paper, under appropriate regularity assumptions, our analysis and some results can be naturally

extended to cases with explosive roots over a range of quantiles (see Remark 3.2 below for some related discussions).

Note that if $|\alpha_t| \leq 1$, with $|\alpha_t| < 1$ with probability > 0 , and if $\|u_t\|_r = \{Eu_t^r\}^{1/r} < \infty$ exists, then the r -th moment of α_t is always strictly less than 1. Denoting $\|\alpha_t\|_r < 1$ as a_r , we have

$$\|y_t\|_r \leq a_r \|y_{t-1}\|_r + \mu_r.$$

Consequently

$$\|y_t\|_r \leq (1 - a_r)^{-1} \mu_r < \infty.$$

The next result summarizes some important properties of this process.

Theorem 3.1 *If y_t is determined by (3.1), $Eu_t^2 = \sigma^2 < \infty$, $|\alpha_t| \leq 1$, and $\Pr(|\alpha_t| < 1) > 0$, then y_t is covariance stationary and*

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

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

Remark 3.1 Under the assumptions in Theorem 3.1, by recursively substituting in (3.1), we can see that

$$(3.2) \quad y_t = \sum_{j=0}^{\infty} \beta_{t,j} v_{t-j}, \text{ where } \beta_{t,0} = 1, \text{ and } \beta_{t,j} = \prod_{i=0}^{j-1} \alpha_{t-i}, \text{ for } j \geq 1.$$

is a stationary \mathcal{F}_t -measurable solution to (3.1). In addition, if $\sum_{j=0}^{\infty} \beta_{t,j} v_{t-j}$ converges in L^p , then y_t has a finite p -th order moment.

Remark 3.2 From the proof we can see that even with $\alpha_t > 1$ over some range of quantiles, as long as $\omega_\alpha^2 = E(\alpha_t)^2 < 1$, y_t is still covariance stationary. Thus, a quantile autoregressive process allows for some (transient) forms of explosive behavior while maintaining stationarity in the long run.

Remark 3.3 The \mathcal{F}_t -measurable solution of (3.1) gives a doubly stochastic $MA(\infty)$ representation of y_t . In particular, the impulse response of y_t to a shock u_{t-j} is stochastic and is given by $\beta_{t,j}$. On the other hand, although the impulse response of the quantile autoregressive process is stochastic, it does converge (to zero) in mean square (and thus in probability) as $j \rightarrow \infty$, corroborating the stationarity of y_t .

Remark 3.4 If we consider a conventional AR(1) process with autoregressive coefficient μ_α and denote the corresponding process by \underline{y}_t , the long-run variance of y_t (given by ω_y^2) is (as expected) larger than that of \underline{y}_t . The additional variance the QAR process y_t comes from the variation of α_t . In fact, ω_y^2 can be decomposed into the

summation of the long-run variance of \underline{y}_t and an additional term that is determined by the variance of α_t :

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

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

Remark 3.5 If we denote the autocovariance function of y_t by $\gamma_y(h)$, it is easy to verify that

$$\gamma_y(h) = \mu_\alpha^{|h|} \sigma_y^2, \text{ where } \sigma_y^2 = \frac{\sigma^2}{1 - \omega_\alpha^2}.$$

Given the model (3.1), if we denote the τ -th quantile of α_t as $Q_\alpha(\tau)$, $x_t = (1, y_{t-1})^\top$ and $\alpha(\tau) = (Q_u(\tau), Q_\alpha(\tau))^\top$, we have

$$(3.3) \quad Q_{y_t}(\tau | \mathcal{F}_{t-1}) = x_t^\top \alpha(\tau),$$

and we estimate $\alpha(\tau)$ by (2.3).

If we denote $\hat{v} = \sqrt{n}(\hat{\alpha}(\tau) - \alpha(\tau))$, then $\rho_\tau(y_t - \hat{\alpha}(\tau)^\top x_t) = \rho_\tau(u_{t\tau} - (n^{-1/2}\hat{v})^\top x_t)$, where $u_{t\tau} = y_t - x_t^\top \alpha(\tau)$. Minimization of (2.3) is equivalent to minimizing:

$$(3.4) \quad Z_n(v) = \sum_{t=1}^n [\rho_\tau(u_{t\tau} - (n^{-1/2}v)^\top x_t) - \rho_\tau(u_{t\tau})].$$

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

Denoting $\psi_\tau(u) = \tau - I(u < 0)$, we have $E[\psi_\tau(u_{t\tau}) | \mathcal{F}_{t-1}] = 0$. Using the identity (A.2) given in the Appendix, the objective function of minimization problem (3.4) can be written as

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

We denote the conditional distribution function

$$F_{t-1}(\bullet) = \Pr[y_t < \bullet | \mathcal{F}_{t-1}],$$

and its derivative as $f_{t-1}(\bullet)$. The following Lemmas give asymptotic results that are useful in deriving the limiting distribution of $\sqrt{n}(\hat{\alpha}(\tau) - \alpha(\tau))$.

Lemma 3.1 *Under assumptions given in Theorem 3.1,*

$$n^{-1/2} \sum_{t=1}^n x_t \psi_\tau(u_{t\tau}) \Rightarrow N(0, \tau(1-\tau)\Omega_0).$$

where $\Omega_0 = E(x_t x_t^\top) = \text{diag}[1, \gamma_0]$, $\gamma_0 = E[y_t^2]$.

Lemma 3.2 *Under assumptions given in Theorem 3.1,*

$$\sum_{t=1}^n \int_0^{(n^{-1/2}v)^\top x_t} \{I(u_{t\tau} \leq s) - I(u_{t\tau} < 0)\} ds \Rightarrow \frac{1}{2} v^\top \Omega_1 v$$

where $\Omega_1 = E[f_{t-1}[F_{t-1}^{-1}(\tau)]x_t x_t^\top]$.

The matrices Ω_0 and Ω_1 are limiting forms of the matrices of $n^{-1} \sum_t x_t x_t^\top$ and $n^{-1} \sum_t f_{t-1}[F_{t-1}^{-1}(\tau)]x_t x_t^\top$ respectively. Given the results of Lemmas 3.1 and 3.2, and by application of the convexity lemma, we can derive the asymptotic distribution of $\hat{\alpha}(\tau)$.

Theorem 3.2 *Under the assumptions given in Theorem 3.1,*

$$\sqrt{n}(\hat{\alpha}(\tau) - \alpha(\tau)) \Rightarrow N(0, \tau(1-\tau)\Omega_1^{-1}\Omega_0\Omega_1^{-1}).$$

Corollary 3.1 *Under the assumptions given in Theorem 3.1, in the special case that $\alpha_t = \alpha = \text{constant}$,*

$$\sqrt{n}(\hat{\alpha}(\tau) - \alpha(\tau)) \Rightarrow \frac{1}{f[F^{-1}(\tau)]} N(0, \tau(1-\tau)\Omega_0^{-1}),$$

where $f(\cdot)$ and $F(\cdot)$ are the density and distribution functions of u_t , respectively.

Remark 3.6 As in other linear quantile regression models, the linear quantile autoregressive model usually must be cautiously interpreted as a local approximation to a more complex nonlinear global model. If we interpret the linear form of the model literally then obviously at some point, or points, there will be crossings of the conditional quantile functions – unless these functions are precisely parallel, in which case we have the pure location shift model for the covariate effects. This crossing problem is actually more acute in the autoregressive case since the support of the design space, i.e. the set of x_t that occur with positive probability, is determined within the model. Nevertheless, we may still regard the linear models specified here as valid approximations over a region of interest. Such approximations should always be regarded as provisional; richer data sources can be expected to yield more elaborate

nonlinear specifications that would have validity over larger regions. The B-spline expansion QAR(1) model for Melbourne daily temperature described in Koenker(2000) illustrates this approach.

It should be stressed that the estimated conditional quantile functions,

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

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

$$y_t = x_t^\top \alpha(U_t)$$

we express the observable random variable y_t as a linear function conditioning covariates. But rather than assuming that the coordinates of the vector $\alpha_t = \alpha(U_t)$ are independent random variables we adopt a diametrically opposite viewpoint – that they are perfectly functionally dependent. If the functions $(\alpha_0, \dots, \alpha_p)$ are monotonically increasing then the random vector α_t is comonotonic in the sense of Schmeidler (1986). This is often the case, as our empirical examples will illustrate, but there are important cases for which this monotonicity fails. What then?

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

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

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

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

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

3.2. Unit Root Quantile Autoregression. Our analysis in the previous section focused on the case that $\alpha_t \leq 1$ and the strict inequality ($\alpha_t < 1$) holds with positive probability. We have shown that in this case the time series y_t is stationary and the autoregression quantiles converge at rate root- n . This analysis includes the special

case that $|\alpha_t| = \alpha = \text{constant} < 1$, in which case the model reduces to the conventional stationary autoregressive process.

Arguably the most important alternative in many macroeconomic time series models is the unit root model:

$$y_t = y_{t-1} + u_t,$$

which may be treated as a special case of (3.1) but with $\alpha_t = 1$. In this case, y_t is nonstationary and, of course, the previous autoregressive quantile regression asymptotic analysis no longer holds. In this section, we consider quantile autoregression with a unit root.

In the presence of a unit root, $\alpha(\tau) = (\alpha_0(\tau), \alpha_1(\tau))^\top = (Q_u(\tau), 1)^\top$. Herce (1996) studied asymptotic properties of least absolute deviation estimator of a unit root model. In this section, we consider quantile unit root regressions. Some major steps of the asymptotic analysis for the unit root quantile regression estimator are similar to those in the stationary case, but an important difference also exists due to the nonstationarity. Since u_t are iid $(0, \sigma^2)$, the partial sum process constructed from u_t converges to a Brownian motion $B_u(\bullet)$ with variance σ^2 . Because of the nonstationarity of y_t , the two components in $\hat{\alpha}(\tau) = (\hat{\alpha}_0(\tau), \hat{\alpha}_1(\tau))$ will now have different rates of convergence. In particular, $\hat{\alpha}_1(\tau)$ will converge to unity at rate n . We denote $\hat{v} = D_n(\hat{\alpha}(\tau) - \alpha(\tau))$, where $D_n = \text{diag}(\sqrt{n}, n)$, and write $\rho_\tau(y_t - \hat{\alpha}(\tau)^\top x_t)$ as $\rho_\tau(u_{t\tau} - (D_n^{-1}\hat{v})^\top x_t)$. Minimization of (2.3) is now equivalent to:

$$(3.5) \quad \min_v \sum_{t=1}^n [\rho_\tau(u_{t\tau} - (D_n^{-1}v)^\top x_t) - \rho_\tau(u_{t\tau})].$$

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

In the presence of a unit root, $\sum_{t=1}^n y_{t-1}\psi_\tau(u_{t\tau})$ has a different limit than the stationary case. Notice that $E[\psi_\tau(u_{t\tau})] = 0$, u_t and $\psi_\tau(u_{t\tau})$ are correlated with each other and 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)):

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

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. Notice that $n^{-1/2} \sum_{t=1}^{\lfloor nr \rfloor} \psi_\tau(u_{t\tau})$ converges to a two parameter process $B_\psi^\tau(r) = B_\psi(\tau, r)$ on $(\tau, r) \in [0, 1]^2$ which is a ‘‘mixture of Brownian motion and Brownian bridge’’ in the sense: For fixed r , $B_\psi^\tau(r) = B_\psi(\tau, r)$ is a rescaled Brownian bridge; For each τ , $n^{-1/2} \sum_{t=1}^{\lfloor nr \rfloor} \psi_\tau(u_{t\tau})$ converge weakly to a Brownian motion with variance $\tau(1 - \tau)$. Thus, for each pair (τ, r) , $B_\psi^\tau(r) = B_\psi(\tau, r) \sim N(0, \tau(1 - \tau)r)$.

Notice that u_t are i.i.d, it is easy to verify that

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

Again, using identity (A.2) in the Appendix, the objective function of minimization problem (3.5) can be written as the sum of

$$- \sum_{t=1}^n (D_n^{-1}v)^\top x_t \psi_\tau(u_{t\tau})$$

and

$$\sum_{t=1}^n (u_{t\tau} - (D_n^{-1}v)^\top x_t) \{I(0 > u_{t\tau} > (D_n^{-1}v)^\top x_t) - I(0 < u_{t\tau} < (D_n^{-1}v)^\top x_t)\},$$

and the asymptotics of each component is summarized in the following Lemma.

Lemma 3.3 *If y_t is determined by (3.1), $Eu_t^2 = \sigma^2 < \infty$, under the unit root assumption,*

$$(3.6) \quad D_n^{-1} \sum_{t=1}^n x_t \psi_\tau(u_{t\tau}) \Rightarrow \int_0^1 \bar{B}_u dB_\psi^\tau,$$

$$\begin{aligned} & \sum_{t=1}^n (u_{t\tau} - (D_n^{-1}v)^\top x_t) \{I(0 > u_{t\tau} > (D_n^{-1}v)^\top x_t) - I(0 < u_{t\tau} < (D_n^{-1}v)^\top x_t)\} \\ & \Rightarrow \frac{1}{2} f(F^{-1}(\tau)) v^\top \left[\int_0^1 \bar{B}_u \bar{B}_u^\top \right] v. \end{aligned}$$

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

The limiting distribution of the first component in (3.6), $\int_0^1 dB_\psi^\tau$, is simply $N(0, \tau(1-\tau))$ and is the same as that in Lemma 3.1. It is the second component, $\int_0^1 B_u dB_\psi^\tau$, that differs from the stationary result. The limiting distribution of the QAR estimator for the unit root model is summarized in the following result.

Theorem 3.3 *If y_t is determined by (3.1), $Eu_t^2 = \sigma^2 < \infty$, then, under the unit root assumption,*

$$\begin{bmatrix} \sqrt{n}(\hat{\alpha}_0(\tau) - \alpha_1(\tau)) \\ n(\hat{\alpha}_1(\tau) - 1) \end{bmatrix} \Rightarrow \frac{1}{f(F^{-1}(\tau))} \left[\int_0^1 \bar{B}_u \bar{B}_u^\top \right]^{-1} \int_0^1 \bar{B}_u dB_\psi^\tau.$$

As an immediate by-product of Theorem 3.3, we have

Corollary 3.2 *If y_t is determined by (3.1), $Eu_t^2 = \sigma^2 < \infty$, under the unit root assumption,*

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

where \underline{B}_u is the demeaned version of Brownian motion B_u .

4. HIGHER ORDER QAR

One of the most important extensions of the first order autoregression formulation of the unit root model is probably the augmented Dickey-Fuller (1979) (ADF) type regression model

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

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.

In this section, we consider a quantile regression extension of the ADF type regression by allowing α_1 to take different values over different quantiles of u_t . Similar to the AR(1) case in the previous sections, we assume that $\alpha_{1t} = g_1(u_t)$ is a function of u_t and consider the following p -th order quantile autoregression (QAR(p))

$$y_t = \alpha_{1t} y_{t-1} + \sum_{j=1}^q \alpha_{j+1} \Delta y_{t-j} + u_t,$$

where $q = p - 1$. More generally, we may allow for randomness in other coefficients and assume $\alpha_{j,t} = g_j(u_t)$, ($j = 2, \dots, p$), to be functions of u_t . In this case we have

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

If we denote $A_t(L) = 1 - \alpha_{1t}L - \sum_{j=1}^q \alpha_{j+1,t}L^j(1-L)$, y_t can be alternatively expressed as

$$A_t(L)y_t = u_t.$$

For convenience of our later analysis, we denote $E(\alpha_{j,t})$ as α_j , ($j = 1, \dots, p$), and $E[A_t(L)] = 1 - \alpha_1 L - \sum_{j=1}^q \alpha_{j+1} L^j(1-L)$ as $A(L)$.

We may also introduce an intercept term μ in the above model and consider time series $y_t = \mu + \underline{y}_t$, where $A_t(L)\underline{y}_t = u_t$. In this case, we obtain a QAR(p) with an intercept

$$(4.1) \quad y_t = \alpha_{0,t} + \alpha_{1t} y_{t-1} + \sum_{j=1}^q \alpha_{j+1,t} \Delta y_{t-j} + u_t,$$

where $\alpha_{0,t} = \mu(1 - \alpha_{1t})$.

4.1. Stationary QAR(p). Like the traditional p -th order autoregressive process, a QAR(p) process

$$(4.2) \quad y_t = a_{0,t} + a_{1,t}y_{t-1} + \cdots + a_{p,t}y_{t-p} + u_t,$$

can be expressed as an p -dimensional vector autoregression process of order 1

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

with

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

where $A_{p-1,t} = [a_{1,t}, \dots, a_{p-1,t}]$, $Y_t = [y_t, \dots, y_{t-p+1}]^\top$, and 0_{p-1} is the $(p-1)$ -dimensional vector of zeros. In the Appendix, we show that under regularity conditions given in the following Theorem, an \mathcal{F}_t -measurable solution for (4.2) can be found.

Theorem 4.1 *Let $E(A_t \otimes A_t) = \Omega_A$, if $Eu_t^2 = \sigma^2 < \infty$ and the eigenvalues of Ω_A have moduli less than unity, the time series y_t given by (4.2) is covariance stationary and satisfies a central limit theorem*

$$\frac{1}{\sqrt{n}} \sum_{t=1}^n y_t \Rightarrow N(\mu_y, \omega_y^2), \text{ where } \omega_y^2 = \lim n^{-1} E\left[\sum_{t=1}^n (y_t - \mu_y)\right]^2.$$

If we focus our attention on the largest autoregressive root $\alpha_{1,t}$ in the ADF type regression (4.1) and consider the special case that $\alpha_{j,t} = \alpha_j = \text{constant}$ for $j = 2, \dots, p$, then, similar to the QAR(1) model, if $\alpha_{1,t} \leq 1$ and $|\alpha_{1,t}| < 1$ with positive probability, the time series y_t given by (4.1) is covariance stationary.

Corollary 4.1 *If $|\alpha_{1,t}| \leq 1$ and $|\alpha_{1,t}| < 1$ with positive probability, $Eu_t^2 = \sigma^2 < \infty$, and the p -th order polynomial $A(L)$ has all its roots outside the unit circle, then the time series y_t given by (4.1) is covariance stationary and satisfies a central limit theorem.*

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

$$\Omega_0 = \begin{bmatrix} 1 & \mu_y & 0_q^\top \\ \mu_y & \gamma_0 & \Gamma^\top \\ 0_q & \Gamma & \Omega_{00} \end{bmatrix}$$

where

$$\Gamma = (\gamma_0 - \gamma_1, \gamma_1 - \gamma_2, \dots, \gamma_{q-1} - \gamma_q)^\top$$

$$\Omega_{00} = \begin{bmatrix} 2\delta_0 & \delta_1 - \delta_0 & \cdots & \delta_{q-1} - \delta_{q-2} \\ \delta_1 - \delta_0 & 2\delta_0 & \cdots & \delta_{q-2} - \delta_{q-3} \\ \vdots & \vdots & \ddots & \vdots \\ \delta_{q-1} - \delta_{q-2} & \delta_{q-2} - \delta_{q-3} & \cdots & 2\delta_0 \end{bmatrix}$$

and 0_q is the q -dimensional vector of zeros. In the special case that $Ey_t = 0$,

$$\Omega_0 = \text{diag} \left[1, \begin{bmatrix} \gamma_0 & \Gamma^\top \\ \Gamma & \Omega_{00} \end{bmatrix} \right].$$

We have the following asymptotic results on the p -th order quantile autoregression.

Lemma 4.1 *Under assumptions given in Theorem 4.1,*

$$n^{-1/2} \sum_{t=1}^n x_t \psi_\tau(u_{t\tau}) \Rightarrow N(0, \tau(1-\tau)\Omega_0),$$

$$\sum_{t=1}^n \int_0^{(n^{-1/2}v)^\top x_t} \{I(u_{t\tau} \leq s) - I(u_{t\tau} < 0)\} ds \Rightarrow \frac{1}{2} v^\top \Omega_1 v$$

where $\Omega_1 = E[f_{t-1}[F_{t-1}^{-1}(\tau)]x_t x_t^\top] = \lim n^{-1} \sum_{t=1}^n f_{t-1}[F_{t-1}^{-1}(\tau)]x_t x_t^\top$.

Theorem 4.2 *Under the assumptions given in Theorem 4.1,*

$$\sqrt{n}(\hat{\alpha}(\tau) - \alpha(\tau)) \Rightarrow N(0, \tau(1-\tau)\Omega_1^{-1}\Omega_0\Omega_1^{-1}).$$

4.2. Unit Root Case. In the unit root case, $\alpha_{1,t} = \alpha_1 = 1$, and (4.1) reduces to the conventional ADF regression,

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

Let $\hat{\alpha}(\tau) = (\hat{\alpha}_0(\tau), \hat{\alpha}_1(\tau), \dots, \hat{\alpha}_p(\tau))$ and $D_n = \text{diag}(\sqrt{n}, n, \sqrt{n}, \dots, \sqrt{n})$, then the analysis of $\hat{\alpha}(\tau)$ follows a similar procedure to that of Section 3.

Denote $w_t = \Delta y_t$, then, for each τ , the partial sums of the process $(w_t, \psi_\tau(u_{t\tau}))$ follow a bivariate invariance principle

$$n^{-1/2} \sum_{t=1}^{\lfloor nr \rfloor} (w_t, \psi_\tau(u_{t\tau}))^\top \Rightarrow (B_w(r), B_\psi^\tau(r))^\top = \underline{\Sigma}(\tau)^{1/2} (W_1(r), W_2(r))^\top = BM(0, \underline{\Sigma}(\tau))$$

where $W_1(r)$ and $W_2(r)$ are independent standard Brownian motions and

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

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

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

Notice again that u_t are uncorrelated with y_{t-1} , we have

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

In addition, for convenience of our analysis, we denote $[1, B_w(r)]^\top$ as $\overline{B}_w(r)$.

We also need to consider the limiting distribution of

$$(4.3) \quad \begin{bmatrix} \frac{1}{\sqrt{n}} \sum_t \Delta y_{t-1} \psi_\tau(u_{t\tau}) \\ \vdots \\ \frac{1}{\sqrt{n}} \sum_t \Delta y_{t-q} \psi_\tau(u_{t\tau}) \end{bmatrix}.$$

If we denote that $E[w_t w_{t-j}] = \nu_j$, it can be shown that (4.3) converges to a q -dimensional normal variate $\Phi = [\Phi_1, \dots, \Phi_q]^\top$ 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 $\int_0^1 \overline{B}_w dB_\psi^\tau$.

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

Theorem 4.3 *Under the unit root assumption,*

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

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

Remark 4.1 As an immediate by-product of Theorem 4.3, 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 4.2 *Under the unit root assumption,*

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

In particular,

$$(4.4) \quad n(\hat{\alpha}_1(\tau) - 1) \Rightarrow \frac{1}{f(F^{-1}(\tau))} \left[\int_0^1 \underline{B}_w^2 \right]^{-1} \int_0^1 \underline{B}_w dB_\psi^\tau,$$

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

5. INFERENCE ON THE QAR PROCESS

The asymptotic theory we developed in the previous sections facilitates statistical inference on the quantile autoregression process. In this section, we turn our attention to inference in QAR models. We are particularly interested in three types of models: The first model is the QAR model that allows different persistent effects at different quantiles. In relatively “short” range such a time series may display “local persistency” but over longer time horizon it behave like a stationary system. The second case of interest is the classical constant-coefficient stationary autoregressive model and the third case is the unit root model that has attracted a great deal of research attention in recent years. We shall study quantile regression inference of these models. In particular, although other inference problems can be analyzed, we consider here the following three types of inference problems that are of paramount interest in many applications: (1) $H_{01} : \alpha_1(\tau) = \bar{\alpha} < 1$, with known $\bar{\alpha}$; (2) $H_{02} : \text{Constancy of } \alpha_1(\tau)$, i.e., $\alpha_1(\tau) = \bar{\alpha} < 1$, for $\tau \in \mathcal{T}$ and with unknown $\bar{\alpha}$; (3) The unit root hypothesis $H_{03} : \alpha_1(\tau) = 1$. For inference problems (1) and (3), we consider both the case at specific quantiles τ (say, median, lower quartile, upper quartile) or over a range of quantiles $\tau \in \mathcal{T}$.

5.1. Testing $H_{01} : \alpha_1(\tau) = \bar{\alpha}$, with known $|\bar{\alpha}|$; The hypothesis H_{01} may be treated as an important special case of a more general inference problem in the form of a linear hypothesis $H_{01}^* : R\alpha(\tau) = r$ (at $\tau = \tau_0$ or, $\tau \in \mathcal{T}$) and y_t is stationary, where R denotes an $m \times p$ -dimensional matrix.

Under the assumptions of Theorem 3.1 or Theorem 4.1, we have

$$\sqrt{n}R(\hat{\alpha}(\tau) - \alpha(\tau)) \Rightarrow [R\Omega_1^{-1}\Omega_0\Omega_1^{-1}R^\top]^{1/2} \mathcal{N}(0, \tau(1-\tau)I_{p+1})$$

Therefore, a regression Wald statistic (if we consider $\tau = \tau_0$) or process (if we consider $\tau \in \mathcal{T}$) can be constructed as

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

where $\hat{\Omega}_1$ and $\hat{\Omega}_0$ are consistent estimators of Ω_1 and Ω_0 . If we are interested in testing $R\alpha(\tau) = r$ at a particular quantile $\tau = \tau_0$, a Chi-square test can be conducted based on the statistic $W_n(\tau_0)$. If we are interested in testing $R\alpha(\tau) = r$ over $\tau \in \mathcal{T}$, we may consider, say, a Kolmogorov-Smirnov (KS) type sup-Wald test $\sup_{\tau \in \mathcal{T}} W_n(\tau)$. The limiting distributions are summarized in the following Theorem.

Theorem 5.1 *Under the assumptions of Theorem 1 or 4, and under the hypothesis H_{01}^* ,*

$$W_n(\tau_0) \Rightarrow \chi_m^2, \text{ and } \sup_{\tau \in \mathcal{T}} W_n(\tau) \Rightarrow \sup_{\tau \in \mathcal{T}} Q_m^2(\tau),$$

where $Q_m(\tau)$ is a Bessel process of order m . For any fixed τ , $Q_m^2(\tau) \sim \chi_m^2$ is a centered Chi-square random variable with m -degrees of freedom.

For the special case $H_{01} : \alpha_1(\tau) = \bar{\alpha}_1$, since we only have one restriction, t-test may be considered. Let

$$V_n(\tau) = \frac{\sqrt{n}(\hat{\alpha}_1(\tau) - \bar{\alpha}_1)}{\hat{\omega}_{22}}$$

where ω_{22}^2 is the $(2, 2)$ -element in $\hat{\Omega}_1^{-1} \hat{\Omega}_0 \hat{\Omega}_1^{-1}$. To test $\alpha_1(\tau) = \bar{\alpha}_1$ at τ_0 , we consider $t_n(\tau) = V_n(\tau) / \sqrt{\tau(1-\tau)}$. Under the null hypothesis, $t_n(\tau) \Rightarrow \mathcal{N}(0, 1)$, providing an asymptotic normal test. To test the hypothesis $\alpha_1(\tau) = \bar{\alpha}$ for $\tau \in \mathcal{T}$, we may consider a Kolmogorov-Smirnov statistic $\sup_{\tau \in \mathcal{T}} |V_n(\tau)|$, which converges (under the null) weakly to $\sup_{\tau \in \mathcal{T}} |G_1(\tau)|$, where $G_1(\tau)$ is a standard Brownian bridge; or a Cramer-von-Mises (CM) type test of the form $\int_{\tau \in \mathcal{T}} V_n(\tau)^2 d\tau$, converging to $\int_{\tau \in \mathcal{T}} G_1(\tau)^2 d\tau$ under the null hypothesis..

For the estimation of $\hat{\Omega}_1$, see, inter alia, Koenker and Bassett(1982), Koenker (1994), Powell (1987), and Koenker and Machado (1999) for related discussions.

5.2. Constancy of $\alpha_1(\tau)$. We consider testing the hypothesis of constancy of $\alpha_1(\tau) = \alpha_1 < 1$ over $\tau \in \mathcal{T}$ (stationary case) in the presence of estimated nuisance parameters. This can be treated as a test for the asymmetric dynamics in time series y_t . Natural candidates for testing constancy of $\alpha_1(\tau)$ over $\tau \in \mathcal{T}$ are again, the KS type test $\sup_{\tau \in \mathcal{T}} |V_n(\tau)|$ or CM test $\int_{\tau \in \mathcal{T}} V_n(\tau)^2 d\tau$, where

$$(5.1) \quad V_n(\tau) = \sqrt{n}(\hat{\alpha}_1(\tau) - \alpha_1) / \hat{\omega}_{22}.$$

However, we do not know α_1 and thus the test $\sup_{\tau \in \mathcal{T}} |V_n(\tau)|$ is infeasible. Under the null, it is possible to estimate α_1 by, say, $\hat{\alpha}_1$, at rate root- n , thus we may replace α_1 by $\hat{\alpha}_1$ in (5.1) and consider testing constancy of $\alpha_1(\tau)$ based on $\sup_{\tau \in \mathcal{T}} |\hat{\alpha}_1(\tau) - \hat{\alpha}_1|$. Unfortunately, the estimation of α_1 brings nuisance parameters into the limiting distributions of

$$\hat{V}_n(\tau) = \frac{\sqrt{n}(\hat{\alpha}_1(\tau) - \hat{\alpha}_1)}{\hat{\omega}_{22}}.$$

As shown in (5.2) below, the necessity of estimating α_1 introduces a drift component ($[f(F^{-1}(\tau)) \lim \sqrt{n}(\hat{\alpha}_1 - \alpha_1) / \Omega_{0.22}]$) in addition to the simple Brownian bridge process, invalidating the distribution-free character of the original tests $\sup_{\tau \in \mathcal{T}} |V_n(\tau)|$ and $\int_{\tau \in \mathcal{T}} V_n(\tau)^2 d\tau$.

To restore the asymptotically distribution free nature of inference, we employ a martingale transformation proposed by Khmaladze (1981) over the process $\hat{V}_n(\tau)$.

Notice that under H_{02} , $\Omega_1^{-1}\Omega_0\Omega_1^{-1} = [f(F^{-1}(\tau))]^{-2}\Omega_0^{-1}$, thus

$$(5.2) \quad \begin{aligned} \widehat{V}_n(\tau) &= \sqrt{n}((\widehat{\alpha}_1(\tau) - \alpha_1) - (\widehat{\alpha}_1 - \alpha_1))/\widehat{\omega}_{22} \\ &\Rightarrow G_1(\tau) - [f(F^{-1}(\tau))Z \end{aligned}$$

where $G_1(\tau)$ is a standard Brownian bridge and $Z = \lim \sqrt{n}(\widehat{\alpha}_1 - \alpha_1)/\Omega_{0.22}$, where $\Omega_{0.22}$ is the (2,2)-element of the matrix Ω_0^{-1} . We construct the following transformation over $\widehat{V}_n(\tau)$:

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

where $\dot{g}_n(s)$ and $C_n(s)$ are uniformly consistent estimators of $\dot{g}(r) = (1, (\dot{f}/f)(F^{-1}(r)))^\top$ and $C(s) = \int_s^1 \dot{g}(r)\dot{g}(r)^\top dr$. The transformed process $\widetilde{V}_n(\tau)$ converges to a standard Brownian motion. For estimation of $\dot{g}_n(s)$ and more discussions of quantile regression inference based on the martingale transformation approach, see, Koenker and Xiao (2002) and references therein.

Theorem 5.2 *Under H_{02} ,*

$$\widetilde{V}_n(\tau) \Rightarrow W(\tau), \quad \sup_{\tau \in \mathcal{T}} |\widetilde{V}_n(\tau)| \Rightarrow \sup_{\tau \in \mathcal{T}} |W(\tau)|, \quad \text{and} \quad \int_{\mathcal{T}} V_n(\tau)^2 \Rightarrow \int_{\mathcal{T}} W(\tau)^2,$$

where $W(r)$ is a standard Brownian motion.

Remark 5.1 Again, we can consider a more general inference problem $H_{02}^* : R\alpha(\tau) = r$ and y_t is stationary. R denotes an $m \times p$ -dimensional matrix and r is an m -dimensional vector with some or all elements unknown but estimable. The hypothesis H_{02} can then be treated as the special case where r equals the true value of α_1 (which has to be estimated) and R is an $(p+1)$ -dimensional row vectors with the second element being one and other elements being zeros. The vector r is unknown but can be estimated.

5.3. Unit Root Tests. The unit root hypothesis has been frequently examined in time series applications in recent years. In this Section, we consider unit root tests based on the quantile autoregression. The proposed autoregressive quantile regression provides a robust approach that examines the unit root property not only at the median, but also at other quantiles of the process. We expect that at least for certain cases, say in the presence of non-Gaussian innovations or asymptotic dynamics, the quantile regression based unit root tests may have some advantages over the traditional testing procedure.

We express the unit root hypothesis in terms of $H_{03} : \alpha_1(\tau) = 1$. To test the unit root hypothesis, we may test $\alpha_1(\tau) = 1$ at some selected representative quantiles (say, $\{\tau_j\}_{j=1}^J$) (see discussions later in this section). Alternatively, we can construct

Kolmogorov-Smirnov or Cramer-von-Mises type tests based on the regression quantile process for $\tau \in \mathcal{T}$. We define the coefficient-based process

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

and let $t_n(\tau)$ be the t-ratio statistic of $\widehat{\alpha}_1(\tau)$:

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

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 is the projection matrix onto the space orthogonal to $(1, \Delta y_{t-1}, \dots, \Delta y_{t-q})$. By the results in previous sections, we have that under the unit root hypothesis

$$(5.3) \quad U_n(\tau) \Rightarrow U(\tau) = \frac{1}{f(F^{-1}(\tau))} \left[\int_0^1 \underline{B}_w^2 \right]^{-1} \int_0^1 \underline{B}_w dB_\psi^\tau,$$

and

$$(5.4) \quad t_n(\tau) \Rightarrow t(\tau) = \frac{1}{\sqrt{\tau(1-\tau)}} \left[\int_0^1 \underline{B}_w^2 \right]^{-1/2} \int_0^1 \underline{B}_w dB_\psi^\tau,$$

The above two processes are quantile regression counterparts of the coefficient and t-ratio based statistics of the conventional Augmented Dickey-Fuller tests. 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:

$$(5.5) \quad QKS_\alpha = \sup_{\tau \in \mathcal{T}} |U_n(\tau)|, \quad QKS_t = \sup_{\tau \in \mathcal{T}} |t_n(\tau)|,$$

and

$$(5.6) \quad QCM_\alpha = \int_{\tau \in \mathcal{T}} U_n(\tau)^2 d\tau, \quad QCM_t = \int_{\tau \in \mathcal{T}} t_n(\tau)^2 d\tau.$$

In practice, we may calculate $U_n(\tau)$ and $t_n(\tau)$ at, say, $\{\tau_i = i/n\}_{i=1}^n$, and thus the test statistics are obtained by taking maximum over τ_i (for the KS type tests) or using numerical integration (for the CM type tests).

It is clear from (5.3) and (5.4) that the above distributions are not standard and depend on nuisance parameters. If we test H_{03} at selected quantiles, modifications are possible to remove the nuisance parameters (see our discussion later in this section). However, Monte Carlo experiments indicate that the finite sample performance based on the modified statistics may be poor, especially in the presence of normal innovations [also see Thompson (2002) for related discussion]. For this reason, we may consider generating critical values for the unmodified statistics using simulation or resampling methods.

We first consider the following resampling procedure:

1. Let $w_t = \Delta y_t$, ($t = 2, \dots, n$), estimate β_1, \dots, β_q by a q -th order autoregression¹:

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

and obtain residuals \hat{u}_t .

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

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

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

3. Generate y_t^* under the unit root null: $y_t^* = y_{t-1}^* + w_t^*$, with $y_1^* = y_1$.
4. Estimate the following p -th order autoregressive quantile regression

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

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

$$U_n^*(\tau) = n(\hat{\alpha}_1^*(\tau) - 1),$$

and

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

and calculate QKS_α^* , QKS_t^* , QCM_α^* , and QCM_t^* based on $U_n^*(\tau)$ and $t_n^*(\tau)$.

In the above procedure, we generate y_t^* under the null hypothesis of unit root. The limiting null distribution of the test statistics can then be approximated by repeating steps 2-4 many times. 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.,

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

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

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

Alternatively, instead of using resampling methods, we may directly simulate the Brownian motions. Notice that \underline{B}_w and B_ψ^τ are Brownian motions and can be approximated by sums of Gaussian random variables, the limiting distribution of the

¹We may also use the Yule-Walker method, which is asymptotically equivalent to the OLS method, to estimate the autoregression.

quantity $\left[\int_0^1 \underline{B}_w^2\right]^{-1} \int_0^1 \underline{B}_w dB_\psi^\tau$ may be approximated by simulation. In particular, we may replace step 4 by directly approximating $\int_0^1 \underline{B}_w^2$ and $\int_0^1 \underline{B}_w dB_\psi^\tau$ using

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

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

$$\frac{n}{f(\widehat{F^{-1}}(\tau))} \left[\sum_t (y_t^* - \bar{y}^*)^2 \right]^{-1} \left[\sum_t (y_t^* - \bar{y}^*) \psi_\tau(u_{t\tau}^*) \right],$$

and

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

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

To test the unit root hypothesis, we may also consider testing $\alpha_1(\tau) = 1$ at some selected representative quantiles $\{\tau_s\}_{s=1}^S$ (say, quartiles or deciles). Again, we may consider our test using the coefficient based statistic $U_n(\tau_s)$ or the t -ratio statistic $t_n(\tau_s)$ and use critical values generated by simulation or resampling methods. In this case, modifications are possible to restore the distributional free properties of the test statistics at quantile τ .

In particular, we can decompose $\int_0^1 \underline{B}_w dB_\psi^\tau$ as

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

where $\lambda_{w\psi}(\tau) = \sigma_{w\psi}(\tau)/\sigma_w^2$ 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 with \underline{B}_w . The limiting distribution of $n(\hat{\alpha}_1(\tau) - 1)$ can be decomposed into

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

Let $\bar{y} = n^{-1} \sum_t y_t$, and $f(\widehat{F^{-1}}(\tau))$, $\hat{\sigma}_w^2$ and $\hat{\sigma}_{w\psi}(\tau)$ be consistent nonparametric estimators of $f(F^{-1}(\tau))$, σ_w^2 and $\sigma_{w\psi}(\tau)$, and $\hat{\lambda}_{w\psi}(\tau) = \hat{\sigma}_{w\psi}(\tau)/\hat{\sigma}_w^2$, then under the null,

$$nf(\widehat{F^{-1}}(\tau))(\hat{\alpha}_1(\tau) - 1) - \frac{n \sum (y_t - \bar{y})(\Delta y_t) \hat{\lambda}_{w\psi}(\tau)}{\sum (y_t - \bar{y})^2} \Rightarrow \frac{\int \underline{B}_w dB_{\psi.w}^\tau}{\int_0^1 \underline{B}_w^2}.$$

We may construct the following modified statistics

$$Q_1(\tau) = \frac{\widehat{\sigma}_w}{\widehat{\sigma}_{\psi.w}(\tau)} \left[n f(\widehat{F^{-1}}(\tau)) (\widehat{\alpha}_1(\tau) - 1) - \frac{n \sum (y_t - \bar{y})(\Delta y_t) \widehat{\lambda}_{w\psi}(\tau)}{\sum (y_t - \bar{y})^2} \right],$$

and

$$\begin{aligned} Q_2(\tau) &= \left[\frac{1}{\widehat{\sigma}_w^2 n^2} \sum_t (y_t - \bar{y})^2 \right]^{1/2} Q_1(\tau) \\ &= \frac{f(\widehat{F^{-1}}(\tau))}{\widehat{\sigma}_{\psi.w}(\tau)} \left[\sum_t (y_t - \bar{y})^2 \right]^{1/2} (\widehat{\alpha}_1(\tau) - 1) - \frac{\widehat{\lambda}_{w\psi}(\tau) \sum (y_t - \bar{y})(\Delta y_t)}{\widehat{\sigma}_{\psi.w}(\tau) [\sum_t (y_t - \bar{y})^2]^{1/2}} \end{aligned}$$

Under the unit root hypothesis,

$$Q_1(\tau) \Rightarrow \left[\int_0^1 \underline{W}_1^2 \right]^{-1} \int \underline{W}_1 dW_2, \text{ and } Q_2(\tau) \Rightarrow N(0, 1).$$

The statistic $Q_2(\tau)$ provides a normal test for the unit root hypothesis at the τ -th quantile.

Remark 5.2 Other test statistics that are asymptotically equivalent to $Q_1(\tau)$ and $Q_2(\tau)$ can be constructed following Herce (1996). The decomposition of $\int_0^1 \underline{B}_w dB_\psi^\tau$ may be written as

$$\sigma_{w\psi}(\tau) \int \underline{W}_1 dW_1 + \Delta(\tau)^{1/2} \int \underline{W}_1 dW_2,$$

where $\Delta(\tau) = \det(\underline{\Sigma}(\tau)) = \sigma_w^2 \sigma_\psi^2(\tau) - \sigma_{w\psi}^2(\tau)$, and $\underline{W}_1(r) = W_1(r) - \int_0^1 W_1$. In addition, if we consider the ordinary least squares ADF regression

$$(5.7) \quad y_t = \alpha_0 + \alpha_1 y_{t-1} + \sum_{j=1}^q \alpha_{j+1} \Delta y_{t-j} + u_t$$

and denote the estimator of α_1 as $\widetilde{\alpha}_1$, it is well known that

$$n(\widetilde{\alpha}_1 - 1) \Rightarrow \frac{\sigma_u \int_0^1 \underline{W}_1 dW_1}{\sigma_w \int_0^1 \underline{W}_1^2}.$$

Following Herce (1996), we may construct the following modified statistic

$$L_1(\tau) = \frac{f(\widehat{F^{-1}}(\tau)) \widehat{\sigma}_w^2}{\widehat{\Delta}(\tau)^{1/2}} n(\widehat{\alpha}_1(\tau) - 1) - \frac{\widehat{\sigma}_w \widehat{\sigma}_{w\psi}(\tau)}{\widehat{\sigma}_u \widehat{\Delta}(\tau)^{1/2}} n(\widetilde{\alpha}_1 - 1)$$

and

$$L_2(\tau) = \left[\frac{1}{\widehat{\sigma}_w^2 n^2} \sum_t (y_t - \bar{y})^2 \right]^{1/2} L_1(\tau)$$

where $\hat{\sigma}_u^2 = n^{-1} \sum_{t=1}^n \hat{u}_t^2$, with \hat{u}_t being the residual of the ordinary ADF regression, and $\hat{\Delta}(\tau) = \tau(1 - \tau)\hat{\sigma}_w^2 - \hat{\sigma}_{w\psi}(\tau)^2$. Under the unit root assumption,

$$L_1(\tau) \Rightarrow \left[\int_0^1 \underline{W}_1^2 \right]^{-1/2} N(0, 1), \text{ and } L_2(\tau) \Rightarrow N(0, 1).$$

6. NUMERICAL RESULTS

6.1. Monte Carlo Results. We conducted a limited Monte Carlo experiment to examine the effectiveness of the QAR method. In particular, we examine the size and power properties of the proposed tests. Given the empirical interests on unit root and persistency properties, we report our simulation results on the tests of the unit root hypothesis.

The data were generated from

$$(6.1) \quad y_t = \alpha_t y_{t-1} + u_t,$$

where u_t are i.i.d. random variables. We consider both the case where u_t are standard normal variates and the case that u_t are student- t distributed variables with 3 degrees of freedom. For the tests, we considered both the Kolmogorov-Smirnov type test (denoted as QKS in the tables) and the Cramer-von-Mises type test (denoted as QCM in the tables), and compared them with the traditional ADF test. The coefficient based tests and t-ratio based tests gave similar results and we report those of the coefficient based tests, i.e. QKS_α , QCM_α , and ADF_α tests. For the quantile regression based tests, we use critical values obtained by resampling procedures. The number of repetitions in the resampling process is 2000. For each test, the number of repetition is 100. We also choose $\mathcal{T} = [0.1, 0.9]$. The sample size is $n = 100$.

When $\alpha_t = 1 = \text{constant}$, y_t is a unit root process and the empirical rejection rates gives the size of tests. For the choice of alternatives, we considered both the conventional constant autoregression with $\alpha_t = \alpha = 0.95, 0.9$, and 0.85 , and the case with asymmetric dynamics. In this case, we considered the following choices of α_t ,

$$\begin{cases} \alpha_t = 1 & u > 0 \\ \alpha_t = \underline{\alpha} & u \leq 0. \end{cases}, \text{ with } \underline{\alpha} = 0.95, 0.9, 0.85.$$

Table 1 reports the empirical size and power for the case with Gaussian innovations and table 2 reports results for student- t innovations. The following general conclusion can be draw from the Monte Carlo results: (1) The proposed QAR method has reasonable performance relative to the conventional procedures in the presence of Gaussian innovations, and has higher power in the presence of non-Gaussian innovations. (2) In the presence of asymmetric dynamics under the alternatives, the QAR method has in general better performance. (3). The Cramer-von-Mises type test has relatively better finite sample performance than the Kolmogorov-Smirnov type test.

Table 1: Empirical Rejection Rates with Gaussian Innovations				
	ADF _t	ADF _α	QKS	QCM
$\alpha = 1$	0.06	0.05	0.05	0.06
$\alpha = 0.95$	0.10	0.12	0.11	0.12
$\alpha = 0.90$	0.22	0.26	0.21	0.23
$\alpha = 0.85$	0.60	0.68	1	1
$\left\{ \begin{array}{l} \alpha = 1 \quad u > 0 \\ \alpha = 0.95 \quad u \leq 0. \end{array} \right.$	0.06	0.09	0.1	0.1
$\left\{ \begin{array}{l} \alpha = 1 \quad u > 0 \\ \alpha = 0.9 \quad u \leq 0. \end{array} \right.$	0.12	0.15	0.16	0.23
$\left\{ \begin{array}{l} \alpha = 1 \quad u > 0 \\ \alpha = 0.85 \quad u \leq 0. \end{array} \right.$	0.36	0.41	0.45	0.47

Table 2: Empirical Rejection Rates with $t(3)$ -Distributed Innovations				
	ADF	ADF	QKS	QCM
$\alpha = 1$	0.06	0.08	0.05	0.05
$\alpha = 0.95$	0.16	0.23	0.18	0.24
$\alpha = 0.9$	0.37	0.45	0.42	0.60
$\alpha = 0.85$	0.54	0.68	0.61	0.84
$\left\{ \begin{array}{l} \alpha = 1 \quad u > 0 \\ \alpha = 0.95 \quad u \leq 0. \end{array} \right.$	0.08	0.08	0.07	0.12
$\left\{ \begin{array}{l} \alpha = 1 \quad u > 0 \\ \alpha = 0.9 \quad u \leq 0. \end{array} \right.$	0.15	0.20	0.23	0.27
$\left\{ \begin{array}{l} \alpha = 1 \quad u > 0 \\ \alpha = 0.85 \quad u \leq 0. \end{array} \right.$	0.22	0.26	0.32	0.44

7. U.S. INTEREST RATE DYNAMICS

There have been many claims and observations that some economic time series are asymmetric over the business cycle. For example, it has been observed that increases in the unemployment rate are much sharper than declines; firms are more apt to increase than to decrease in prices. It has also been argued that positive shocks to the economy may be more persistent than negative shocks. For this reason, studies have been conducted on the existence of asymmetric behavior in these series. If an economic time series displays asymmetric dynamics systematically, then appropriate theoretical models are needed to incorporate such behavior.

In this section, we apply the QAR model to several US interest rate series. The short-term interest rate is central to much of theoretical and empirical macroeconomics and finance. However, there is still no consensus on the dynamics of short

term interest rate. In this section, we examine interest rate dynamics using our proposed procedures. We focus on the interest rate itself and do not consider multifactor (term structure) models.

The data that we consider are one month, three month, and annual rates of interest in the US. Following the literature, we focus on nominal rates to avoid the data problems that would be created by attempting to define real rates. In particular, we looked at (seasonally adjusted) monthly observations of the 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.

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. Evidence based on the traditional unit root tests has accumulated suggesting that there is a unit root in interest rates. See, inter alia, Nelson and Plosser 1982, Schotman and Van Dijk 1991, El-Jahel et al. 1997, Ball and Torous, 1996).

We 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 ADF regression estimates of the largest autoregressive roots of the three interest series are all very close to unity (see the last rows of Tables 3A, 4A, 5A). Tables 3A, 4A, and 5A 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 the 5% level of significance, leading to the conclusion that the interest rate series exhibit unit roots.

We re-visit these interest rate series using the proposed QAR methods. We first test the unit root hypothesis in these series using the Kolmogorov type statistic QKS_α and the Cramer-von-Mises type test QCM_α . The tests were constructed over $\tau \in \mathcal{T} = [0.1, 0.9]$. The last two columns of Tables 3A, 4A, and 5A report the calculated statistics and the 5% level critical values that were calculated based on the resampling procedure given in Section 5. 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 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 pure unit root process.

Tables 3B, 4B, and 5B provide a more detailed examination on the interest rate series at selected quantiles. In particular, we investigate the behavior of these series at each decile. The second column in each of these tables report the estimates of the largest autoregressive root at each specified quantile. These estimates indicate

that there exist asymmetry in the persistency. The largest autoregressive coefficient estimate $\hat{\alpha}_1(\tau)$ has different values over different quantiles, displaying asymmetric dynamics over the business cycle. In particular, we find that $\hat{\alpha}_1(\tau)$ monotonically 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. The interest rate has an 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 Fed. That is, it might be more acceptable for the Fed to lower rates by a large amount and quickly 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, the interest rates are more persistent in the presence of positive shocks than the negative ones.

We also consider tests for the unit root hypothesis based the autoregression estimates $\hat{\alpha}_1(\tau)$ at selected quantiles. The third columns in Tables 3B, 4B, 5B 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 4 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 in Section 5. 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 during an expansion.

The QAR method also provides an alternative explanation to the conditional heteroskedasticity in the short term interest rate time series. There have been many empirical studies on volatility of interest rate data. Among existing studies, one of the most popular classes of empirical models for the study of short-term interest rate volatility are the continuous time model, in which volatility is parameterized as a function of previous interest rates levels (y_{t-1}). A partial listing of these type models are Merton (1973), Brennan and Schwartz (1980), Cox, Ingersoll and Ross (1985), Longstaff and Schwartz (1992), including the square root model of Cox, Ingersoll and Ross (1985). Chan, Karolyi, Longstaff, and Sanders (1992) considered several generalizations of these continuous time models and studied their interest rate dynamics. They conclude that one of the most important features of the short-term interest rate dynamics is the relationship between interest rate volatility and the previous level of interest rate. Also see empirical studies of interest rate using ARCH/GARCH type

Table 3A				
	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: $\hat{\alpha}_1 = 0.976$

Table 3B						
Quantiles	$\hat{\alpha}_1(\tau)$	$U_n(\tau)$	2.5% c.v.	5% c.v.	95% c.v.	97.5% c.v.
0.1	0.886	-41.4**	-22.75	-18.06	2.88	4.19
0.2	0.929	-26.4**	-12.31	-10.11	1.62	2.40
0.3	0.961	-14.3**	-7.48	-5.94	1.05	1.55
0.4	0.981	-7.06*	-3.69	-3.10	0.49	0.75
0.5	0.994	-2.08	-3.61	-2.85	0.49	0.74
0.6	1.014	5.39#	-5.65	-4.51	0.76	1.17
0.7	1.029	11.13##	-7.75	-6.30	1.12	1.67
0.8	1.055	20.47##	-11.17	-9.14	1.67	2.36
0.9	1.111	41.39##	-18.83	-15.11	2.49	3.73

OLS Estimator: $\hat{\alpha}_1 = 0.976$

TABLE 1. T

those values denoted by a (*) are significant the 5% level against the alternative $\alpha_1(\tau) < 1$, and those with (**) are significant at the 1% level against this alternative. Similarly, the values denoted by a (#) are significant the 5% level against the alternative is $\alpha_1(\tau) > 1$, and those with a (##) are significant at the 1% level.

models where the volatilities are affected by shocks (unexpected “news”), see, inter alia, Engle, Lilien and Robins (1987), Evans (1989).

The QAR model provides a simple alternative model of conditional heteroskedasticity. In this model, the interest rate volatility is still affected by the previous level, but also adjusted by unexpected “news”. Our empirical analysis find strong evidence of asymmetry in the business cycle dynamics of short term interest rate. These empirical results suggests that the conditional heteroskedasticity in short term interest rate series may (at least partially) be caused by the asymmetric response of the Fed.

Table 4A				
	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: $\hat{\alpha}_1 = 0.977$				

Table 4B						
Quantiles	$\hat{\alpha}_1(\tau)$	$U_n(\tau)$	2.5% c.v.	5% c.v.	95% c.v.	97.5% c.v.
0.1	0.884	-43.03	-19.99	-16.51	2.54	3.88
0.2	0.926	-27.55	-9.45	-7.62	1.24	1.84
0.3	0.959	-15.22**	-6.59	-5.50	1.03	1.47
0.4	0.984	-5.83	-4.39	-3.50	0.62	0.86
0.5	0.991	-3.51	-4.03	-3.26	0.51	0.76
0.6	1.012	4.82	-5.53	-4.48	0.76	1.19
0.7	1.034	12.62	-7.79	-6.44	1.06	1.52
0.8	1.065	23.91	-10.95	-8.78	1.60	2.21
0.9	1.107	39.73	-20.41	-15.99	2.93	4.13
OLS Estimator: $\hat{\alpha}_1 = 0.977$						

Table 5A				
	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: $\hat{\alpha}_1 = 0.974$				

Table 5B						
Quantiles	$\hat{\alpha}_1(\tau)$	$U_n(\tau)$	2.5% c.v.	5% c.v.	95% c.v.	97.5% c.v.
0.1	0.829	-14.64	-26.89	-22.13	3.33	4.95
0.2	0.865	-11.61	-12.12	-9.15	1.24	2.02
0.3	0.965	-2.99	-6.00	-4.50	0.89	1.25
0.4	0.981	-1.64	-5.24	-4.12	0.92	1.22
0.5	1.004	0.39	-5.52	-4.40	0.90	1.35
0.6	1.052	4.44	-6.88	-5.48	1.17	1.74
0.7	1.126	10.84	-9.53	-7.69	1.61	2.45
0.8	1.165	14.25	-13.32	-11.11	2.22	3.36
0.9	1.126	10.83	-17.88	-14.65	3.06	4.75
OLS Estimator: $\hat{\alpha}_1 = 0.974$						

APPENDIX A. PROOFS

A.1. Proof of Theorem 3.1. To find an \mathcal{F}_t -measurable solution to (3.1), we first express y_t in terms of measurable functions on \mathcal{F}_t by repeating (3.1), the existence and properties of the solution to (3.1) can then be reduced to those of a moving average process with random coefficients:

$$\sum_{j=0}^{\infty} \beta_{tj} u_{t-j},$$

where $\beta_{t0} = 1$, $\beta_{tj} = \prod_{i=0}^{j-1} \alpha_{t-i}$, for $j \geq 1$. To show that $y_t = \sum_{j=0}^{\infty} \beta_{tj} u_{t-j}$ is a Doubly Stochastic Moving Average Process, we need to show that, for each t , the infinite series defining y_t converges in mean or in L^2 (Doob 1953, p155). Denote $E(\alpha_t) = \mu_\alpha$ and $E(\alpha_t)^2 = \omega_\alpha^2$, then $|\mu_\alpha| < 1$ and $\omega_\alpha^2 < 1$. It's easy to verify that

$$\sum_{j=0}^{\infty} E[\beta_{tj}]^2 = \sum_{j=0}^{\infty} \omega_\alpha^{2j} = \frac{1}{1 - \omega_\alpha^2} < \infty.$$

In addition, since u_t (and thus α_t) are iid, $E[\beta_{tj} u_{t-j} \beta_{tl} u_{t-l}] = 0$ for any $j \neq l$. Consequently, for each t the sequence $\{\beta_{tj} u_{t-j}\}_{j=0}^{\infty}$ is an orthogonal sequence with $E[\beta_{tj} u_{t-j}]^2 = \omega_\alpha^{2j} \sigma^2$. As a result, the limit $\sum_{j=0}^{\infty} \beta_{tj} u_{t-j}$ exists in mean square, and thus in probability. By routine computations, we can obtain the following results:

$$E[y_t] = 0, \quad E[y_t y_{t+h}] = \sigma^2 \frac{\mu_\alpha^{|h|}}{1 - \omega_\alpha^2}.$$

Following the approach of Gordin (1969) and Hannan (1973), we write

$$y_t = \sum_{j=0}^{m-1} \{E(y_t | \mathcal{F}_{t-j}) - E(y_t | \mathcal{F}_{t-j-1})\} + E(y_t | \mathcal{F}_{t-m}).$$

From the above analysis we have $E(y_t | \mathcal{F}_{t-m}) \rightarrow 0$ in quadratic mean as $m \rightarrow \infty$, thus if we denote $\xi_{t,j} = E(y_t | \mathcal{F}_{t-j}) - E(y_t | \mathcal{F}_{t-j-1})$, and we have

$$y_t = \sum_{j=0}^{\infty} \xi_{t,j}.$$

Since $\{\xi_{t,j}, \mathcal{F}_{t-j}\}$ is a martingale difference sequence and $\sum_{j=0}^{\infty} (\text{var}(\xi_{0j}))^{1/2} < \infty$, an application of Gordin(1969), gives

$$\frac{1}{\sqrt{n}} \sum_{t=1}^n y_t \Rightarrow N(0, \omega_y^2), \text{ where } \omega_y^2 = \frac{(1 + \mu_\alpha)\sigma^2}{(1 - \mu_\alpha)(1 - \omega_\alpha^2)}.$$

■

The proofs of Lemma 3.1, Lemma 3.2 and Theorem 3.2 are largely simplified versions of Lemma 4.1 and Theorem 4.2, which are contained in the proof of Theorem 4.2. The proofs of Lemma 3.3 and Theorem 3.3 are similar to the proof of Theorem 4.3.

A.2. Proof of Theorem 4.1. We consider the general form of a p -th order autoregression model with an intercept term given by (4.2). Denote $E(a_{j,t}) = \mu_j$, and assume that $1 - \sum \mu_j \neq 0$. Let $\mu = \alpha_0 / (1 - \sum_{j=1}^p \mu_j)$, and $\underline{y}_t = y_t - \mu$, we have

$$(A.1) \quad \underline{y}_t = a_{1,t}\underline{y}_{t-1} + \cdots + a_{p,t}\underline{y}_{t-p} + v_t,$$

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

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

and the $p \times p$ random matrix

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

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

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

and the original process can be written as

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

By substitution, we have $\underline{Y}_t = \underline{Y}_{t,m} + R_{t,m}$, where

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

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

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

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

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

If we denote

$$A = E[A_t] = \begin{bmatrix} \bar{\mu}_{p-1} & \mu_p \\ I_{p-1} & 0_{p-1} \end{bmatrix},$$

where $\bar{\mu}_{p-1} = [\mu_1, \dots, \mu_{p-1}]$, then $A_t = A + \Xi_t$, where $E(\Xi_t) = 0$, and

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

then

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

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

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

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

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

For the special case that $\alpha_{j,t} = \alpha_j = \text{constant}$ ($j = 2, \dots, p$) in (4.1) and $|\alpha_{1,t}| \leq 1$ and $|E\alpha_{1,t}| < 1$, if $A(L)$ has all its roots outside the unit circle, the eigenvalues of Ω_A have moduli less than unity. Thus, the time series y_t given by (4.1) is covariance stationary. ■

A.3. Proof of Theorem 4.2. 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 $u \neq 0$,

$$\begin{aligned} \rho_\tau(u - v) - \rho_\tau(u) &= -v\psi_\tau(u) + (u - v)\{I(0 > u > v) - I(0 < u < v)\} \\ \text{(A.2)} \quad &= -v\psi_\tau(u) + \int_0^v \{I(u \leq s) - I(u < 0)\} ds. \end{aligned}$$

Let $u_{t\tau} = y_t - x_t^\top \alpha(\tau)$, and denote $\hat{v} = \sqrt{n}(\hat{\alpha}(\tau) - \alpha(\tau))$, then $\rho_\tau(y_t - \hat{\alpha}(\tau)^\top x_t) = \rho_\tau(u_{t\tau} - (n^{-1/2}\hat{v})^\top x_t)$. If \hat{v} is a minimizer of $Z_n(v) = \sum_{t=1}^n [\rho_\tau(u_{t\tau} - (n^{-1/2}v)^\top x_t) - \rho_\tau(u_{t\tau})]$, we have $\hat{v} = \sqrt{n}(\hat{\alpha}(\tau) - \alpha(\tau))$.

Notice that $E[\psi_\tau(u_{t\tau}) | \mathcal{F}_{t-1}] = 0$. Using the identity (A.2), the objective function of minimization problem can be written as

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

The asymptotics for $n^{-1/2} \sum_{t=1}^n x_t \psi_\tau(u_{t\tau})$ is straightforward. Note that $x_t \in \mathcal{F}_{t-1}$ so since $E[\psi_\tau(u_{t\tau})|\mathcal{F}_{t-1}] = 0$, $x_t \psi_\tau(u_{t\tau})$ is a martingale difference sequence and by the properties of y_t

$$n^{-1/2} \sum_{t=1}^n x_t \psi_\tau(u_{t\tau}) \Rightarrow N(0, \tau(1-\tau)\Omega_0).$$

We now consider the limit of

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

For convenience of asymptotic analysis, we denote

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

We further define

$$\bar{\xi}_t(v) = E\{\xi_t(v)|\mathcal{F}_{t-1}\}, \quad \text{and} \quad \bar{W}_n(v) = \sum_{t=1}^n \bar{\xi}_t(v),$$

then $\{\xi_t(v) - \bar{\xi}_t(v)\}$ is a martingale difference sequence.

Denote the conditional distribution function $F_{t-1}(\bullet) = \Pr[y_t < \bullet|\mathcal{F}_{t-1}]$, and its derivative as $f_{t-1}(\bullet)$, *a.s.*, notice that

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

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

Notice that $f_{t-1}(s_n)$ is uniformly integrable for any sequence $s_n \rightarrow F_{t-1}^{-1}(\tau)$,

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

By our assumptions and stationarity of $\{f_{t-1}[F_{t-1}^{-1}(\tau)]\}$ and y_t , we have

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

By the Asymptotic Equivalence Lemma, the limiting distribution of $\sum_t \xi_t(v)$ is the same as that of $\sum_t \bar{\xi}_t(v)$. As a result,

$$\begin{aligned} Z_n(v) &= -\sum_{t=1}^n (n^{-1/2}v)^\top x_t \psi_\tau(u_{t\tau}) + \sum_{t=1}^n \int_0^{(n^{-1/2}v)^\top x_t} \{I(u_{t\tau} \leq s) - I(u_{t\tau} < 0)\} ds. \\ &\Rightarrow -v^\top N(0, \tau(1-\tau)\Omega_0) + \frac{1}{2}v^\top \Omega_1 v = Z(v) \end{aligned}$$

By the convexity Lemma of Pollard (1991) and arguments of Knight (1989), notice that $Z_n(v)$ and $Z(v)$ are minimized at $\hat{v} = \sqrt{n}(\hat{\alpha}(\tau) - \alpha(\tau))$ and $N(0, \tau(1-\tau)\Omega_1^{-1}\Omega_0\Omega_1^{-1})$ respectively, by Lemma A of Knight (1989) we have,

$$\sqrt{n}(\hat{\alpha}(\tau) - \alpha(\tau)) \Rightarrow N(0, \tau(1-\tau)\Omega_1^{-1}\Omega_0\Omega_1^{-1})$$

■

A.4. Proof of Theorem 4.3. Again we follow the approach of Knight (1989). The proof is similar to that of Theorem 5 but with changes to accommodate the nonstationarity. Using this approach, Hecce (1996) derives the asymptotic distribution of the least absolute deviation estimators. We use a similar argument here. If we denote $v = D_n(\theta - \theta(\tau))$, where $D_n = \text{diag}(\sqrt{n}, n, \sqrt{n} \cdot \cdot, \sqrt{n})$, the minimization is equivalent to

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

Using identity (A.2), we have

$$\begin{aligned} &\sum_{t=1}^n [\rho_\tau(u_{t\tau} - v^\top D_n^{-1}x_t) - \rho_\tau(u_{t\tau})] \\ &= -\sum_{t=1}^n v^\top D_n^{-1}x_t \psi_\tau(u_{t\tau}) + \sum_{t=1}^n (u_{t\tau} - v^\top D_n^{-1}x_t) \{I(0 > u_{t\tau} > v^\top D_n^{-1}x_t) - I(0 < u_{t\tau} < v^\top D_n^{-1}x_t)\}. \end{aligned}$$

For the first term,

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

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

$$\sum_{t=1}^n (u_{t\tau} - v^\top D_n^{-1} x_t) I(0 < u_{t\tau} < v^\top D_n^{-1} x_t).$$

For convenience of asymptotic analysis, we denote

$$U_n(v) = \sum_{t=1}^n z_t(v), \text{ where } z_t(v) = (v^\top D_n^{-1} x_t - u_{t\tau}) I(0 < u_{t\tau} < v^\top D_n^{-1} x_t).$$

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

$$\begin{aligned} U_{nm}(v) &= \sum_{t=1}^n z_{tm}(v), \\ z_{tm}(v) &= (v^\top D_n^{-1} x_t - u_{t\tau}) I(0 < u_{t\tau} < v^\top D_n^{-1} x_t) M_t \\ M_t &= I(0 \leq v_2 n^{-1/2} y_{t-1} \leq m). \end{aligned}$$

We further define

$$\bar{z}_{tm}(v) = E\{(v^\top D_n^{-1} x_t - u_{t\tau}) I(0 < u_{t\tau} < v^\top D_n^{-1} x_t) M_t | \mathcal{F}_{t-1}\},$$

and

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

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

$$\begin{aligned} \bar{U}_{nm}(v) &= \sum_{t=1}^n E\{(v^\top D_n^{-1} x_t - u_{t\tau}) I(0 < u_{t\tau} < v^\top D_n^{-1} x_t) M_t | \mathcal{F}_{t-1}\} \\ &= \sum_{t=1}^n \int_{F_u^{-1}(\tau)}^{[v^\top D_n^{-1} x_t + F_u^{-1}(\tau)] M_t} [(v^\top D_n^{-1} x_t + F_u^{-1}(\tau)) M_t - r] f_u(r) dr \\ &= \sum_{t=1}^n \int_{F_u^{-1}(\tau)}^{[v^\top D_n^{-1} x_t + F_u^{-1}(\tau)] M_t} \left[\int_r^{[v^\top D_n^{-1} x_t + F_u^{-1}(\tau)] M_t} ds \right] f_u(r) dr \\ &= \sum_{t=1}^n \int_{F_u^{-1}(\tau) \leq s \leq [v^\top D_n^{-1} x_t + F_u^{-1}(\tau)] M_t} \int_{F_u^{-1}(\tau) \leq r \leq s} f_u(r) dr ds \\ &= \sum_{t=1}^n \int_{F_u^{-1}(\tau)}^{[v^\top D_n^{-1} x_t + F_u^{-1}(\tau)] M_t} [s - F_u^{-1}(\tau)] \left[\frac{F_u(s) - F_u(F_u^{-1}(\tau))}{s - F_u^{-1}(\tau)} \right] ds \end{aligned}$$

Thus

$$\begin{aligned} \bar{U}_{nm}(v) &= \sum_{t=1}^n \int_{F_u^{-1}(\tau)}^{[v^\top D_n^{-1} x_t + F_u^{-1}(\tau)] M_t} [s - F_u^{-1}(\tau)] f_u[F_u^{-1}(\tau)] ds + o_p(1) \\ &= \sum_{t=1}^n f_u[F_u^{-1}(\tau)] \left\{ \frac{[s - F_u^{-1}(\tau)]^2}{2} \Big|_{F^{-1}(\tau)}^{[v^\top D_n^{-1} x_t + F^{-1}(\tau)] M_t} \right\} + o_p(1) \\ &= \frac{1}{2} \sum_{t=1}^n f_u[F^{-1}(\tau)] v^\top [D_n^{-1} x_t x_t^\top D_n^{-1}] v M_t + o_p(1). \end{aligned}$$

Thus

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

where

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

We now follow the arguments of Pollard (1984, p171), notice that $(v^\top D_n^{-1} x_t) I(0 \leq v_2 n^{-1/2} y_{t-1} \leq m) \xrightarrow{P} 0$ uniformly in t ,

$$\sum_{t=1}^n E[z_{tm}(v)^2 | \mathcal{F}_{t-1}] \leq \max\{(v^\top D_n^{-1} x_t) I(0 \leq v_2 n^{-1/2} y_{t-1} \leq m)\} \sum_t \bar{z}_{tm}(v) \rightarrow 0.$$

thus the following summation of martingale difference sequence

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

converges to zero in probability. Thus the limiting distribution of $\sum_t z_{tm}(v)$ is the same as that of $\sum_t \bar{z}_{tm}(v)$, i.e., $U_{nm}(v) \Rightarrow \eta_m$. Let $m \rightarrow \infty$, we have

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

and

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

By a similar argument as Hecce (1996), we can show that

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

Similarly, we show that $\sum_{t=1}^n (u_{t\tau} - (D_n^{-1} v)^\top x_t) \{I(0 > u_{t\tau} > (D_n^{-1} v)^\top x_t)\}$ converges to

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

with

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

Thus,

$$\sum_{t=1}^n (u_{t\tau} - (D_n^{-1} v)^\top x_t) \{I(0 > u_{t\tau} > (D_n^{-1} v)^\top x_t) - I(0 < u_{t\tau} < (D_n^{-1} v)^\top x_t)\} \Rightarrow f(F^{-1}(\tau)) v^\top \Psi v$$

where

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

As a result,

$$\begin{aligned}
Z_n(v) &= -\sum_{t=1}^n (D_n^{-1}v)^\top z_t \psi_\tau(u_{t\tau}) \\
&+ \sum_{t=1}^n (u_{t\tau} - (D_n^{-1}v)^\top z_t) \{I(0 > u_{t\tau} > (D_n^{-1}v)^\top z_t) - I(0 < u_{t\tau} < (D_n^{-1}v)^\top z_t)\} \\
&\Rightarrow -v^\top \Phi^* + f(F^{-1}(\tau))v^\top \Psi v
\end{aligned}$$

■

REFERENCES

- [1] Ball, C.A. and W.N. Torous, 1996, Unit Roots and the estimation of Interest rate dynamics, *Journal of Empirical Finance*, 3, 215-238.
- [2] Bassett, G., and R. Koenker, 1982, An Empirical Quantile Function for linear models with iid errors, *Journal of the American Statistical Association*, Vol. 77, 407-415.
- [3] Billingsley, P, 1961, The Lindeberg-Levy Theorem for Martingales, *Proc. Amer. Math. Soc.*, 12, 788-792.
- [4] Brandt, A., 1986, The stochastic equation $Y_{n+1} = A_n Y_n + B_n$ with stationary coefficients, *Adv. Applied Probability*, 18, 211-220.
- [5] Brennan, M., and E. Schwartz, Analyzing Convertible Bonds, *Journal of Financial and Quantitative Analysis*, 15, 1980, 907-929.
- [6] Chan, K.C., G.A. Karolyi, F.A. Longstaff, and A. Sanders, 1992, An empirical comparison of alternative models of the short-term interest rate, *Journal of Finance*, V.47, 1209-1227.
- [7] Cox, J., J. Ingersoll, and S. Ross, 1985, A theory of term structure of interest rates, *Econometrica*, 53, 145-166.
- [8] Davis, R., K. Knight, and J. Liu, 1992, M-estimation for autoregressions with infinite variance, *Stochastic Processes and Their Applications*, 40, 145-180.
- [9] Denneberg, D. 1994, *Non-additive measure and integral*, Kluwer Academic Publishers.
- [10] Dickey, D.A., and W.A. Fuller, 1979, Distribution of Estimators for Autoregressive Time Series with a Unit Root, *Journal of the American Statistical Association*, Vol. 74, 427-431.
- [11] Doob, J., 1953, *Stochastic Processes*, John Wiley, New York.
- [12] El-Jahel, L., H. Lindberg, and W. Perraudin, 1997, Interest rate distributions, yield curve modelling and monetary policy, *Mathematics of Derivative Securities*, eds. Dempster and Pliska, Cambridge University Press, 1-35.
- [13] Engle, R., D. Lilien, and R. Robins, 1987, Estimating time varying risk premia in the term structure: the ARCH-M model, *Econometrica*, 55, 391-407.
- [14] Evans, M. 1989, Interpreting the term structure using the intertemporal capital asset pricing model, *Manuscript*, New York University.
- [15] Gordin, M.I., 1969, The central limit theorem for stationary process, *Soviet Mathematics-Doklady*, 10:1174-1176.
- [16] Hannan, E.J., 1973, Central limit theorems for time series regression, *Z. Wahrsch. Verw. Gebiete* 26: 157-170.
- [17] Hasan, M.N. and R. Koenker, 1997, Robust rank tests of the unit root hypothesis, *Econometrica* 65, No.1, 133-161.
- [18] Hecce, M., 1996, Asymptotic Theory of LAD estimation in a unit root process with finite variance errors, *Econometric Theory*, 12, 129-153.
- [19] Jureckova and Hallin, 1999, Optimal Tests for Autoregressive Models Based on Autoregression Rank Scores, *The Annals of Statistics*, 27, 1385-1414,
- [20] Karlsen, H.A., 1990, Existence of Moments in a stationary stochastic difference equation, *Adv. Applied Probability*, 22, 129-146.
- [21] Khmaladze, E., 1981, Martingale Approach to the goodness of fit tests, *Theory Probab. Appl.*, 26, 246-265.
- [22] Knight, K., 1989, Limit Theory for Autoregressive-parameter estimates in an infinite -variance random walk, *The Canadian Journal of Statistics*, V. 17, 261-278.
- [23] Knight, K., 1998, Asymptotics for L1 regression estimates under general conditions, *Annals of Statistics*.
- [24] Koenker, R. and G. Bassett, 1978, Regression Quantiles, *Econometrica*, V46, 33-49.

- [25] Koenker, R. and G. Bassett, 1982, Robust Tests for heteroskedasticity based on regression quantiles, *Econometrica*, 50, 43-61.
- [26] Koenker, R. and J. Machado, 1999, Goodness of fit and related inference processes for quantile regression, *Journal of the American Statistical Association*, Vol. 81, 1296-1310.
- [27] Koenker, R. 2000, Galton, Edgeworth, Frisch and prospects for quantile regression in econometrics, *J. of Econometrics*, 95, 347-374.
- [28] Koenker, R. and Z. Xiao, 2002, Inference on the Quantile Regression Processes, *Econometrica*, 70, 1583-1612.
- [29] Koul, H., and A. K. Saleh, 1995, Autoregression quantiles and related rank-scores processes, *The Annals of Statistics*, V. 23, No. 2, 670-689.
- [30] McCabe, B. and A. Tremayne, 1995, Testing a time series for difference stationarity, *The Annals of Statistics*, 23, 1015-1028.
- [31] Merton, R., 1973, Theory of Rational Option Pricing, *Bell Journal of Economics and Management Science*, 4, 141-183.
- [32] Nelson, C.R. and C.I. Plosser, 1982, Trends and random walks in macroeconomic time series: some evidence and implications, *Journal of Monetary Economics*, 10, 139-162.
- [33] Nicholls, D.F., and B.G. Quinn, 1982, *Random Coefficient Autoregressive Models: An Introduction*, Springer-Verlag.
- [34] Phillips, P.C.B., and S. Durlauf, 1986, Multiple time series regression with integrated processes, *Review of Economic Studies*, 53, 473-495.
- [35] Pollard, D., 1984, *Convergence of Stochastic Processes*, Springer-Verlag, New York.
- [36] Pollard, D., 1991, Asymptotics for Least Absolute Deviation Regression Estimators, *Econometric Theory*, 7, 186-199.
- [37] Rissanen, J., 1978, Modelling by shortest data description, *Automatica*, 14, 465-471.
- [38] Schmeidler, D., 1996, Integral representation without additivity, *Proceedings of Amer. Math. Society*, 97, 255-261.
- [39] Schotman, P., and H.K. van Dijk, 1991, On the Bayesian routes to unit root, *Journal of Applied Econometrics*, 6, 387-402.
- [40] Schwarz, G., 1978, Estimating the dimension of a model, *The Annals of Statistics*, 6, 461-464.
- [41] Tong, H., 1990, *Nonlinear Time Series: A Dynamical Approach*, Oxford University Press.
- [42] Weiss, A., 1987, Estimating Nonlinear Dynamic Models Using Least Absolute Error Estimation, *Econometric Theory*, 7, 46-68.

UNIVERSITY OF ILLINOIS AT URBANA-CHAMPAIGN

UNIVERSITY OF ILLINOIS AT URBANA-CHAMPAIGN