Quantile Cointegrating Regression

Zhijie Xiao*

Department of Economics, Boston College, USA.

January, 2009.

Abstract

Quantile regression has important applications in risk management, portfolio optimization, and asset pricing. The current paper studies estimation, inference and financial applications of quantile regression with cointegrated time series. In addition, a new cointegration model with varying coefficients is proposed. In the proposed model, the value of cointegrating coefficients may be affected by the shocks and thus may vary over the innovation quantile. The proposed model may be viewed as a stochastic cointegration model which includes the conventional cointegration model as a special case. It also provides a useful complement to cointegration models with (G)ARCH effects. Asymptotic properties of the proposed model and limiting distribution of the cointegrating regression quantiles are derived. In the presence of endogenous regressors, fully-modified quantile regression estimators and augmented quantile cointegrating regression are proposed to remove the second order bias and nuisance parameters. Regression Wald test are constructed based on the fully modified quantile regression estimators. An empirical application to stock index data highlights the potential of the proposed method.

JEL: C22, G1.

Key Words: ARCH/GARCH, Cointegration, Portfolio Optimization, Quantile Regression, Time Varying.

1   Introduction

Since Granger (1981) and Engle and Granger (1987), cointegration has become a common econometric tool for empirical analysis in numerous areas (see, inter alia, Phillips and Ouliaris 1988; Johansen 1995; and Hsiao 1997, among others), especially in macroeconomic and financial applications. Well-known financial applications of

cointegration include Campbell and Shiller (1987) in the study of bubbles in asset prices, Cochrane (1994) and Lettau and Ludvigson (2001) on the predictability of stock prices, Hall, Anderson and Granger (1992) on term structure of interest rates, Pindyck and Rothemberg (1992), Lucas (1997) and Alexander (1999) on portfolio allocation. Also see, inter alia, Evans (1991), Campbell, Lo, and MacKinley (1997), Cerchi and Havenner (1988), Chowdhury (1991), Hendry (1996), on other applications in finance.

In applications of portfolio management, cointegration measures long run co-movements in prices, leading to hedging methodologies that may be more effective than traditional correlation analysis-based approach in the long term. Recent research (e.g. Alexander (1999)) indicate that mis-pricing and over-hedging can occur if cointegration is ignored. When portfolios are allocated using risk criteria such as the conditional value at risk (CVaR), the optimization problem leads to a quantile regression.

Quantile regression method has recently attracted an increasing amount of research attention in finance. Taylor (1999) applies quantile regression approach to estimating the distribution of multiperiod returns. Engle and Manganelli (2004) proposes estimating value at risk (VaR) using quantile regression. Quantile regression is now an important tool in modern risk management operations. For example, the popular risk measure, value at risk (VaR), is simply a concept of quantile and can be naturally estimated using quantile regression method. In recent years, motivated by regulatory reasons in the financial sector, an influential axiomatic foundation raised by Artzner, Delbaen, Eber, and Heath (1999) is the concept of "coherence". A widely used coherent risk measure is the *conditional value at risk*[1] (CVaR) (Rockafellar and Uryasev (2000)) (or, in different names, *Expected Shortfall*, Acerbi and Tasche (2002); *tail conditional expectation*, Artzner, Delbaen, Eber, and Heath (1999)). Bassett, Koenker and Kordas (2004) recently show that, when the portfolio risk is measured by CVaR, the managers operation can be formulated as a quantile regression of cointegrated time series. Quantile cointegrating regression also provide a robust method of index tracking in portfolio management. See, Koenker and Zhao (1996), Chernozhukov and Umanstev (2001), Christoffersen, Hahn and Inoue (2001), Giacomini and Komunjer (2005), for more studies on financial applications of quantile regression.

Although there has been a large amount of recent attempts in applying quantile regression to financial time series models, there is little investigation on the statistical validity and properties of these methods and models. The first contribution of this paper is to study the statistical properties of quantile regression estimation and inference of cointegrated time series. Limiting distribution of the regression quantiles is derived. In the presence of endogenous regressors, a fully-modified quantile regression

---

[1]VaR is not a coherent risk measure.

estimator is proposed to remove the second order bias and nuisance parameters. We develop statistical inference based on the quantile regression estimators. Asymptotic properties of the proposed fully-modified quantile regression estimator and testing procedure based on this estimator are studied.

Although cointegration has gained great popularity in the last 20 years, absence of cointegration has been frequently discovered in applications using traditional analysis on time series that are seemly cointegrated. One explanation of these empirical findings is the existence of varying cointegrating coefficients - the coefficients that characterizes their long-run relationship may vary over time, although these economic variables still move together in the long run. We try to address this issue in the current paper. The second contribution of this paper is to propose a new model of cointegration in which the cointegrating coefficients may be varying over time. The proposed model may be viewed as a stochastic cointegration model which includes the conventional cointegration as a special case. In particular, the value of cointegrating coefficients may be affected by the shocks received in each period, and thus may vary over the innovation quantile. For this reason, we call it *quantile cointegration*. The model 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 cointegration models. The quantile cointegration model may be interpreted as a random coefficient regression model with strongly dependent coefficients. The quantile cointegration model allows for additional volatility of the dependent variables in addition to the regressors, and provides an interesting class of cointegration model with conditional heteroskedasticity. We hope that the proposed model provides a useful complement to traditional models with (G)ARCH effects. We apply the proposed quantile cointegration model to U.S. stock index data. The empirical evidence indicates that the cointegrating coefficients are not constant over time, and asymmetric asset pricing dynamics brings additional volatility into prices in addition to market fundamentals.

In matters of notation, we use "$\Rightarrow$" to signify weak convergence of the associated probability measures, $[nr]$ to signify the integer part of nr, $:=$ to signify definitional equality, and $I(k)$ to denote integration of order k. Continuous stochastic process such as the Brownian motion $B(r)$ on $[0,1]$ are usually written simply as B and integrals $\int$ are understood to be taken over the interval $[0,1]$, unless otherwise specified.

2 Quantile Regression on Cointegration Model

In this section, we consider quantile regression of the following cointegration model:

$$y_t = \alpha + \beta' x_t + u_t = \theta' z_t + u_t, \tag{1}$$

where x_t is a k-dimensional vector of integrated regressors, $z_t = (1, x_t')'$, and u_t is mean zero stationary. The quantile regression estimator of the cointegrating vector can be obtained by solving the problem

$$\widehat{\theta}(\tau) = \arg\min_{\theta \in \mathbb{R}^p} \sum_{t=1}^{n} \rho_\tau(y_t - z_t^\top \theta), \tag{2}$$

where $\rho_\tau(u) = u(\tau - I(u < 0))$ as in Koenker and Bassett (1978). In the special case $\tau = 0.5$, the above quantile regression delivers the least absolute deviation (LAD) estimation of the cointegration model (1).

2.1 Limiting Distribution of the Quantile Regression Estimator

To derive the limiting distribution of the quantile regression estimator of the cointegrating vector we follow the approach of Knight (1991) (also see Herce (1996), Hasan and Koenker (1997), Koenker and Xiao (2006) for related results). Let $f(\cdot)$ and $F(\cdot)$ be the p.d.f. and c.d.f. of u_t, denoting $\psi_\tau(u) = \tau - I(u < 0)$, $\alpha(\tau) = \alpha + F^{-1}(\tau)$, $\theta(\tau) = (\alpha(\tau), \beta')'$, and

$$u_{t\tau} = y_t - \theta(\tau)'z_t = u_t - F^{-1}(\tau),$$

we have $Q_{u_{t\tau}}(\tau) = 0$, where $Q_{u_{t\tau}}(\tau)$ is the τ-th quantile of $u_{t\tau}$, and

$$E\psi_\tau(u_{t\tau}) = 0.$$

To facilitate the asymptotic analysis, we make the following assumptions.

Assumption A: Let $v_t = \Delta x_t$, $\{u_t, v_t\}$ is a zero-mean, stationary sequence of $(k+1)$-dimensional random vectors. The partial sums of the vector process $(\psi_\tau(u_{t\tau}), v_t)$ follow a multivariate invariance principle

$$n^{-1/2} \sum_{t=1}^{[nr]} \begin{bmatrix} \psi_\tau(u_{t\tau}) \\ v_t \end{bmatrix} \Rightarrow \begin{bmatrix} B_\psi(r) \\ B_v(r) \end{bmatrix} = BM(0, \Omega)$$

where Ω is the covariance matrix of the Brownian motion $(B_\psi(r), B_v(r)')'$.

Assumption B: The distribution function of u_t, $F(u)$, has a continuous density $f(u)$ with $f(u) > 0$ on $\{u : 0 < F(u) < 1\}$.

Assumption C: The conditional distribution function $F_{t-1}(u) = \Pr[u_t < u | u_{t-j}, j \geq 1; v_{t-k}, k \geq 0]$ has derivative $f_{t-1}(\bullet)$, a.s., and $f_{t-1}(s_n)$ is uniformly integrable for any sequence $s_n \to F^{-1}(\tau)$, and $E[f_{t-1}^\delta(F^{-1}(\tau))] < \infty$ for some $\delta > 1$.

Conformable to $(\psi_\tau(u_{t\tau}), v_t)$, we partition Ω into

$$\Omega = \begin{bmatrix} \omega_\psi^2 & \Omega_{\psi v} \\ \Omega_{v\psi} & \Omega_{vv} \end{bmatrix}$$

The asymptotic distribution of the quantile regression estimator is closely related to the asymptotic behavior of $n^{-1} \sum_{t=1}^{n} x_t \psi_\tau(u_{t\tau})$. Under Assumption A, it is easy to verify that

$$n^{-1} \sum_{t=1}^{n} x_t \psi_\tau(u_{t\tau}) \Rightarrow \int_0^1 B_v dB_\psi + \lambda_{v\psi},$$

where $\lambda_{v\psi}$ is the one-sided long-run covariance between v_t and $\psi_\tau(u_{t\tau})$.

Due to the nonstationarity of x_t, the two components in $\widehat{\theta}(\tau) = (\widehat{\alpha}(\tau), \widehat{\beta}(\tau)')'$ have different rates of convergence. In particular, the estimate of cointegrating vector $\widehat{\beta}(\tau)$ converges at rate n, while the intercept $\widehat{\alpha}(\tau)$ converges at rate $\sqrt{n}$. Thus, we introduce the standardization matrix $D_n = diag(\sqrt{n}, nI_k)$, where I_k is a $k \times k$ identity matrix. The limiting distribution of the quantile regression estimator for the cointegration model is summarized in the following Theorem.

THEOREM 1. *Under Assumptions A, B, and C,*

$$D_n(\widehat{\theta}(\tau) - \theta(\tau)) \Rightarrow \frac{1}{f(F^{-1}(\tau))} \left[\int_0^1 \overline{B}_v \overline{B}_v^\top \right]^{-1} \left[\int_0^1 \overline{B}_v dB_\psi + \Delta_{v\psi} \right],$$

where $\overline{B}_v(r) = (1, B_v(r)')'$, and $\Delta_{v\psi} = (0, \lambda'_{v\psi})'$. In particular,

$$n(\widehat{\beta}(\tau) - \beta) \Rightarrow \frac{1}{f(F^{-1}(\tau))} \left[\int_0^1 \underline{B}_v \underline{B}_v^\top \right]^{-1} \left[\int_0^1 \underline{B}_v dB_\psi + \lambda_{v\psi} \right] \qquad (3)$$

where $\underline{B}_v(r) = B_v(r) - rB_v(1)$ is a k-dimensional demeaned Brownian motion.

The above limiting result is very similar to that of the conventional cointegrating regression estimators: (i) The quantile regression estimator of the cointegrating vector is consistent at the usual $O(n)$ rate. (ii) Like OLS, the quantile regression estimator suffers from second order bias ($\lambda_{v\psi}$) coming from the correlation between the regressor x and the residual u. (iii) In addition, the Brownian motions $B_v(r)$ and $B_\psi(r)$ are in general correlated (as long as $\Omega_{\psi v} \neq 0$ in Ω). (iv) Similar to the usual limit theory for the LAD estimator in both stationary and nonstationary time series regression, the limiting distribution (3) depends on the sparsity function $1/f(F^{-1}(\tau))$. In the special case when $\Delta_{v\psi} = 0$ and $\Omega_{\psi v} = 0$ (x_t and u_s are independent), the limiting distribution (3) is a mixed normal.

2.2 A Fully-Modified Quantile Regression Estimator

We are interested in developing estimation and inference procedures based on the quantile regression in cointegration models. As will become clear in later analysis, the asymptotic behavior of quantile regression-based inference procedures depends on the limiting distribution of $\widehat{\theta}(\tau)$. However, as shown by Theorem 1, the limiting

5

processes $B_v(r)$ and $B_\psi(r)$ are correlated Brownian motions whenever contemporaneous correlation between v_t and $\psi_\tau(u_{t\tau})$ exists. Despite super-consistency, $\widehat{\theta}(\tau)$ is second-order biased and the miscentering effect in the limit distribution is reflected in $\Delta_{v\psi}$. Consequently, the distribution of the test based on the quantile regression residual will be dependent on nuisance parameters.

To restore the asymptotic nuisance parameter free property of inference procedure, we need to modify the original quantile regression estimator so that we obtain a mixed normal limiting distribution. In this paper, we consider two approaches to achieve this goal: (1) Nonparametric fully-modification on the original quantile regression estimator and (2) Parametrically augmented quantile regression using leads and lags. We propose a nonparametric fully-modified quantile regression estimator to deal with the endogeneity problem in this section. In Section 3, we introduce a parametrically augmented quantile regression using leads and lags and extend the conventional cointegration model to the case with varying-coefficients.

We develop a fully-modified quantile cointegrating regression estimator in the spirit of Phillips and Hansen (1990). We first decompose the limiting distribution (3) into the following two components:

$$\frac{1}{f(F^{-1}(\tau))} \left[\int_0^1 \underline{B}_v \underline{B}_v^\top \right]^{-1} \int_0^1 \underline{B}_v dB_{\psi.v}, \text{ and}$$

$$\frac{1}{f(F^{-1}(\tau))} \left[\int_0^1 \underline{B}_v \underline{B}_v^\top \right]^{-1} \left[\int_0^1 \underline{B}_v dB_v^\top \Omega_{vv}^{-1} \Omega_{v\psi} + \lambda_{v\psi} \right]$$

where $B_{\psi.v}(r) = B_\psi(r) - \Omega_{\psi v} \Omega_{vv}^{-1} B_v(r)$ is Brownian motion with variance $\omega_{\psi.v}^2 = \omega_\psi^2 - \Omega_{\psi v} \Omega_{vv}^{-1} \Omega_{v\psi}$. Notice that $B_{\psi.v}(r)$ is independent of $B_v(r)$ and the first term in the above decomposition, $\left[\int_0^1 \underline{B}_v \underline{B}_v^\top \right]^{-1} \int_0^1 \underline{B}_v dB_{\psi.v}$, is a mixed Gaussian variate.

The basic idea of fully-modification on $\widehat{\beta}(\tau)$ (or $\widehat{\theta}(\tau)$) is to construct a nonparametric correction to remove the second term in the above decomposition. To facilitate the nonparametric correction, we consider the following kernel estimates of Ω_{vv}, $\Omega_{v\psi}$, $\lambda_{v\psi}$, λ_{vv}:

$$\widehat{\lambda}_{v\psi} = \sum_{h=0}^{M} k(\frac{h}{M}) C_{v\psi}(h), \widehat{\lambda}_{vv} = \sum_{h=0}^{M} k(\frac{h}{M}) C_{vv}(h),$$

$$\widehat{\Omega}_{v\psi} = \sum_{h=-M}^{M} k(\frac{h}{M}) C_{v\psi}(h), \widehat{\Omega}_{vv} = \sum_{h=-M}^{M} k(\frac{h}{M}) C_{vv}(h),$$

where $k(\cdot)$ is the lag window defined on $[-1, 1]$ with $k(0) = 1$, and M is the bandwidth parameter satisfying the property that $M \to \infty$ and $M/n \to 0$ (say $M = O(n^{1/3})$ for many commonly used kernels, as in Andrews, 1991) as the sample size $n \to \infty$. The quantities $C_{v\psi}(h)$ and $C_{vv}(h)$ are sample covariances defined by $C_{v\psi}(h) =$

$n^{-1} \sum' v_t \psi_\tau(\widehat{u}_{t+h,\tau})$, $C_{vv}(h) = n^{-1} \sum' v_t v'_{t+h}$, where $\sum'$ signifies summation over $1 \le t, t+h \le n$. Candidate kernel functions can be found in standard texts (e.g., Hannan, 1970; Brillinger, 1980; and Priestley, 1981). Let $f(\widehat{F^{-1}(\tau)})$ be a nonparametric sparsity estimator of $f(F^{-1}(\tau))$ (see, e.g., Siddiqui (1960), Bofinger (1975)), we define the following nonparametric fully modified quantile regression estimators:

$$\widehat{\theta}(\tau)^+ = \begin{pmatrix} \widehat{\alpha}(\tau) \\ \widehat{\beta}(\tau)^+ \end{pmatrix}$$

where

$$\widehat{\beta}(\tau)^+ = \widehat{\beta}(\tau) - \frac{1}{f(\widehat{F^{-1}(\tau)})} \left[\sum_t x_t x'_t \right]^{-1} \left[\sum_t x_t v'_t \widehat{\Omega}_{vv}^{-1} \widehat{\Omega}_{v\psi} + n \widehat{\lambda}_{v\psi}^+ \right] \tag{4}$$

and

$$\widehat{\lambda}_{v\psi}^+ = \widehat{\lambda}_{v\psi} - \widehat{\lambda}_{vv} \widehat{\Omega}_{vv}^{-1} \widehat{\Omega}_{v\psi}.$$

Like the fully modified OLS estimators, the fully modified quantile regression estimator of the cointegrating vector has a mixed normal distribution in limit.

THEOREM 2. *Under Assumptions A, B, and C,*

$$
\begin{aligned}
D_n \left(\widehat{\theta}(\tau)^+ - \theta(\tau) \right) &\Rightarrow \frac{1}{f(F^{-1}(\tau))} \left[\int_0^1 \overline{B}_v \overline{B}_v^\top \right]^{-1} \int_0^1 \overline{B}_v dB_{\psi.v} \\
&\equiv MN \left(0, \frac{\omega_{\psi.v}^2}{f(F^{-1}(\tau))^2} \left[\int_0^1 \overline{B}_v \overline{B}_v^\top \right]^{-1} \right).
\end{aligned}
$$

In particular

$$
\begin{aligned}
n(\widehat{\beta}(\tau)^+ - \beta) &\Rightarrow \frac{1}{f(F^{-1}(\tau))} \left[\int_0^1 \underline{B}_v \underline{B}_v^\top \right]^{-1} \int_0^1 \underline{B}_v dB_{\psi.v} \\
&\equiv MN \left(0, \frac{\omega_{\psi.v}^2}{f(F^{-1}(\tau))^2} \left[\int_0^1 \underline{B}_v \underline{B}_v^\top \right]^{-1} \right).
\end{aligned}
$$

2.3 Regression Wald Test

The fully modified quantile regression estimator and resulting asymptotic mixture normal distribution facilitates statistical inference based on quantile cointegrating regression. In this section, we consider the classical inference problem of linear restrictions on the cointegrating vector β:

$$H_0 : R\beta = r,$$

where R denotes an $q \times k$-dimensional matrix and r is an q-dimensional vector.

Under the null hypothesis $H_0 : R\beta = r$ and the assumptions of our previous theorem, we have

$$\frac{f(F^{-1}(\tau))}{\omega_{\psi.v}} \left[R \left(\int_0^1 \underline{B}_v \underline{B}_v^\top \right)^{-1} R^\top \right]^{-1/2} n(R\widehat{\beta}^+(\tau) - r) \Rightarrow N(0, I_q), \qquad (5)$$

where $N(0, I_q)$ represents a q-dimensional standard Normal. Therefore, let

$$M_X = \sum_{t=1}^n (x_t - \overline{x})(x_t - \overline{x})',$$

a regression Wald statistic can be constructed as

$$W_n(\tau) = \frac{f(\widehat{F^{-1}(\tau)})}{\widehat{\omega}_{\psi.v}}(R\widehat{\beta}^+(\tau) - r)^\top \left[R M_X^{-1} R^\top \right]^{-1} (R\widehat{\beta}^+(\tau) - r),$$

where $f(\widehat{F^{-1}(\tau)})$ and $\widehat{\omega}_{\psi.v}$ are consistent estimators of $f(F^{-1}(\tau))$ and $\omega_{\psi.v}$. The limiting distribution of the Wald statistic is summarized in the following Theorem.

THEOREM 3. *Under the assumptions of Theorem 2 and the linear restriction H_0,*

$$W_n(\tau) \Rightarrow \chi_q^2,$$

where χ_q^2 is a centered Chi-square random variable with q-degrees of freedom.

3 Cointegration with Varying Coefficients

3.1 Time-Varying Cointegration Models

Cointegration measures long run equilibrium relationship. For example, many empirical studies on asset pricing consider the rational expectations model for stock prices

$$P_t = (1 + \gamma)^{-1} E_t(P_{t+1} + D_{t+1}), \qquad (6)$$

which is a first-order expectational difference equation, where P_t is the real stock price at t, γ is the real rate of return, and D_t is the dividend. In empirical analyses, dividends is usually characterized as an integrated process (random walk) with drift. A forward-looking solution to the above equation suggests that stock prices and market fundamentals should be cointegrated. Based on such a cointegration relationship, there is a large collection of empirical study on asset pricing. [See, inter alia, Campbell and Shiller (1988), Diba and Grossman (1988), Evans (1991), Campbell, Lo, and MacKinley (1997), Cerchi and Havenner (1988), Chowdhury (1991), Hendry (1996).]

In the traditional cointegration model of Engle and Granger (1987), the cointegrating vector β is constant. However, many financial and economic applications

8

suggest that the cointegrating vector might be varying. Application of cointegration in investment analysis shows that frequent rebalancing is necessary to keep the portfolio in line with the index, indicating the value of cointegrating vector is changing over time. Although the present value model suggests that asset prices are cointegrated with market fundamentals, it is also well known that stock prices are much more volatile than market fundamentals such as dividends, a plausible source of this additional volatility comes from varying cointegrating vector.

In this section, we attempt to extend the traditional cointegration model to a more general class of models in which the cointegrating coefficients are allowed to be varying over time.[2] In particular, we wish to consider the cointegrating regression (1) where the value of cointegrating coefficients $\beta = (\beta_1, \cdots, \beta_k)$ may be affected by the shocks. In recent years, a lot of research effort has been devoted to modifications of the traditional models to incorporate the effect of different types of shocks into one time series model. It is widely acknowledged that many important economic variables may display asymmetric adjustment paths (e.g. Neftci (1984), Enders and Granger (1998), Beaudry and Koop (1993)). In this paper, we introduce the effect of shocks into cointegration models. However, subtle issues arise due to endogeneity of the cointegration model. If we simply consider $\beta = \beta_t$ as functions of u_t, it is difficult to identify β_t from $\beta_t' x_t$ because of correlation between u_t and Δx_t. For this reason, we decompose the residual term u_t into a pure innovation component (denoted as ε_t) and a component related to (and thus can be represented as leads and lags of) Δx_t, and model the varying cointegrating coefficients β_t as a function of the pure innovation component. In particular, we consider the following model which is an extension of (1),

$$y_t = \alpha + \beta_t' x_t + u_t, \tag{7}$$

and make the following assumptions:

Assumption A':   Let $v_t = \Delta x_t$, $\{u_t, v_t\}$ is a zero-mean, stationary sequence of $(k+1)$-dimensional random vectors and for some K, u_t has the following representation

$$u_t = \sum_{j=-K}^{K} v_{t-j}' \Pi_j + \varepsilon_t, \tag{8}$$

where ε_t is a stationary process such that

$$E(v_{t-j}\varepsilon_t) = 0, \text{ for any } j.$$

The partial sums of the vector process $(\psi_\tau(\varepsilon_{t\tau}), v_t)$ follow a multivariate invariance

---

[2]Park and Hahn (1999) studied another type of cointegration model with time-varying coefficients where the coefficients is a function of deterministic time trend t: $\beta_t = \beta(t/T)$.

principle

$$n^{-1/2} \sum_{t=1}^{[nr]} \left[\begin{array}{c} \psi_\tau(\varepsilon_{t\tau}) \\ v_t \end{array} \right] \Rightarrow B(r) = \left[\begin{array}{c} B_\psi^*(r) \\ B_v(r) \end{array} \right] = BM(0, \Omega^*)$$

Assumption B′: The distribution function of ε_t, $F_\varepsilon(\varepsilon)$, has a continuous density $f_\varepsilon(\varepsilon)$ with $f_\varepsilon(\varepsilon) > 0$ on $\{\varepsilon : 0 < F_\varepsilon(\varepsilon) < 1\}$.

Assumption C′: The conditional distribution function $F_{t-1}(u) = \Pr[u_t < u | u_{t-j}, j \geq 1; v_{t-k}, k \geq 0]$ has derivative $f_{t-1}(\bullet)$, $a.s.$, and $f_{t-1}(s_n)$ is uniformly integrable for any sequence $s_n \to F^{-1}(\tau)$, and $E[f_{t-1}^\delta(F^{-1}(\tau))] < \infty$ for some $\delta > 1$.

Assumption D: Let $\beta_t = (\beta_{1t}, \cdots, \beta_{kt})$, the cointegrating coefficients β_{it} are monotone functions of the innovation process ε_t.

The idea of using leads and lags to deal with endogeneity in traditional cointegration model was proposed by Saikkonen (1991). It can be verified that, under Assumption A',

$$f_{\varepsilon\varepsilon}(\lambda) = f_{uu}(\lambda) - f_{uv}(\lambda) f_{vv}(\lambda)^{-1} f_{vu}(\lambda)$$

where $f_{\varepsilon\varepsilon}(\lambda)$, $f_{uu}(\lambda)$, $f_{vv}(\lambda)$ are spectral densities of ε, u, v, and $f_{uv}(\lambda)$ is the cross spectral of u and v, implying that the long run variance of ε is $\omega_{\varepsilon\varepsilon}^2 = \omega_{uu}^2 - \Omega_{uv} \Omega_{vv}^{-1} \Omega_{vu}$. Notice that the Brownian motion $B_\psi^*(r)$ is now independent with $B_v(r)$. We partition the covariance matrix (of the Brownian motion $B(r)$) Ω^* into

$$\Omega^* = \left[\begin{array}{cc} \omega_\psi^{*2} & 0 \\ 0 & \Omega_{vv} \end{array} \right].$$

Under Assumption A′, the original cointegrating regression (7) can be re-written as:

$$y_t = \alpha + \beta_t' x_t + \sum_{j=-K}^{K} \Delta x_{t-j}' \Pi_j + \varepsilon_t.$$

If we denote the τ-th quantile of ε_t as $Q_\varepsilon(\tau)$, let $\mathcal{F}_t = \sigma\{x_t, \Delta x_{t-j}, \forall j\}$, then, conditional on $\mathcal{F}_t$, the τ-th quantile of y_t is given by

$$Q_{y_t}(\tau|\mathcal{F}_t) = \alpha + \beta(\tau)' x_t + \sum_{j=-K}^{K} \Delta x_{t-j}' \Pi_j + F_\varepsilon^{-1}(\tau), \tag{9}$$

where $F_\varepsilon(\cdot)$ is the c.d.f. of ε_t. Let Z_t be the vector of regressors consisting $z_t = (1, x_t)$ and $(\Delta x_{t-j}', j = -K, \cdots, K)$, $\Theta = (\alpha, \beta_t', \Pi_{-K}', \cdots, \Pi_K')'$, and

$$\Theta(\tau) = (\alpha(\tau), \beta(\tau)', \Pi_{-K}', \cdots, \Pi_K')'$$

where $\alpha(\tau) = \alpha + F_\varepsilon^{-1}(\tau)$, then, we can re-write the above regression as

$$y_t = \Theta' Z_t + \varepsilon_t$$

and

$$Q_{y_t}(\tau|\mathcal{F}_t) = \Theta(\tau)'Z_t. \tag{10}$$

Let $\varepsilon_{t\tau} = \varepsilon_t - F_\varepsilon^{-1}(\tau)$, then

$$Q_{\varepsilon_{t\tau}}(\tau) = 0.$$

In the above model, the value of the cointegration coefficients are affected by the innovation received at each period. Consequently the cointegrating vector can vary over the quantiles and thus may be quantile (τ) dependent. The conditioning variables not only shift the location of the distribution of y_t, but also may alter the scale and shape of the conditional distribution. We will refer to this model as the Quantile Cointegration model. Of course, the quantile cointegration model includes the conventional cointegration model of Engle and Granger (1987) as a special case where $\beta(\tau)$ is a vector of constants. In this special case,

$$y_t = \alpha + \beta'x_t + \sum_{j=-K}^{K} \Delta x'_{t-j}\Pi_j + \varepsilon_t, \tag{11}$$

and

$$Q_{y_t}(\tau|x_t) = \alpha + \beta'x_t + \sum_{j=-K}^{K} \Delta x'_{t-j}\Pi_j + F_\varepsilon^{-1}(\tau).$$

We now consider the following modified quantile cointegrating regression:

$$\widehat{\Theta}(\tau) = \arg\min_{\theta} \sum_{t=1}^{n} \rho_\tau(y_t - \Theta'Z_t), \tag{12}$$

Denote $G_n = diag(D_n, \sqrt{n}, \cdots, \sqrt{n}) = diag(\sqrt{n}, n, \cdots, n, \sqrt{n}, \cdots, \sqrt{n})$. Conformable with $\Theta(\tau)$, we partition $\widehat{\Theta}(\tau)$ as follows:

$$\widehat{\Theta}(\tau)' = \left[\begin{array}{cccc} \widehat{\alpha}(\tau), & \widehat{\beta}(\tau)', & \widehat{\Pi}_{-K}(\tau)', & \cdots, & \widehat{\Pi}_K(\tau)' \end{array} \right].$$

Given $\widehat{\Theta}(\tau)$, the τ-th conditional quantile function of y_t, conditional on x_t, can be estimated by,

$$\hat{Q}_{y_t}(\tau|\mathcal{F}_t) = z_t^\top \widehat{\Theta}(\tau),$$

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

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

for some appropriately chosen sequence of τ's.

The limiting distribution of this estimator is given in the following Theorem.

THEOREM 4. *Under Assumptions A′, B′, C′, and D,*

$$G_n(\widehat{\Theta}(\tau) - \Theta(\tau)) \Rightarrow \frac{1}{f_\varepsilon(F_\varepsilon^{-1}(\tau))} \begin{bmatrix} \int_0^1 \overline{B}_v \overline{B}_v^\top & 0 \\ 0 & \Gamma \end{bmatrix}^{-1} \begin{bmatrix} \int_0^1 \overline{B}_v dB_\psi^* \\ \Psi \end{bmatrix}.$$

In particular

$$n(\widehat{\beta}(\tau) - \beta(\tau)) \Rightarrow \frac{1}{f_\varepsilon(F_\varepsilon^{-1}(\tau))} \left[\int_0^1 \underline{B}_v \underline{B}_v^\top \right]^{-1} \int_0^1 \underline{B}_v dB_\psi^*,$$

and where $\overline{B}_v(r)$ and $\underline{B}_v(r)$ are the same as those defined in Theorem 1, $\Gamma = E(V_t V_t')$ and $V_t = (\Delta x'_{t-K}, \cdots, \Delta x'_{t+K})'$, and Ψ is a multivariate normal with dimension conformable with $(\Pi_{-K}(\tau)', \cdots, \Pi_K(\tau)')'$.

It is straightforward to extend the model to allow the coefficients Π_j to be quantile dependent:

$$y_t = \alpha + \beta_t' x_t + \sum_{j=-K}^{K} \Delta x'_{t-j} \Pi_{jt} + \varepsilon_t.$$

REMARK: Notice that the time varying coefficient model may be re-written as constant coefficient model with conditional heteroskedasticity. If we denote $E(\beta_t) = \beta$, $E(\Pi_{jt}) = \Pi_j$, we may write the above random coefficient model as:

$$y_t = \alpha + \beta' x_t + \sum_{j=-K}^{K} \Delta x'_{t-j} \Pi_j + w_t.$$

where the error term:

$$w_t = \varepsilon_t + (\beta_t - \beta)' x_t + \sum_{j=-K}^{K} \Delta x'_{t-j} (\Pi_{jt} - \Pi_j).$$

Since that β_t and Π_{jt} are functions of the innovation term ε_t, this is a cointegration model with conditional heteroskedasticity. The conditional heteroskedasticity comes from the varying-parameters and may display asymmetric dynamics. In this sense, the proposed model may be viewed as an useful alternative of the widely used ARCH or GARCH models, and has the advantage of computational simplicity and allow for certain type of asymmetric behavior in the multivariate system.

3.2 Inference on Quantile Cointegration Models

Notice that the limiting distribution of $\widehat{\beta}(\tau)$ is mixture normal, statistical inference procedures can be constructed based on the above augmented quantile regression. If we consider again the inference problem in Section 3, $H_0 : R\beta(\tau) = r$, let $\widehat{\beta}(\tau)$ be estimated from the augmented quantile regression, and $f_\varepsilon(\widehat{F_\varepsilon^{-1}(\tau)})$ and $\widehat{\omega}_\psi^*$ are

consistent estimators of $f_\varepsilon(F_\varepsilon^{-1}(\tau))$ and ω_ψ^*, we may construct the following regression Wald statistic:

$$W_n(\tau) = \frac{\widehat{f_\varepsilon(F_\varepsilon^{-1}(\tau))}^2}{\widehat{\omega}_\psi^*}(R\widehat{\beta}(\tau) - r)^\top \left[RM_X^{-1}R^\top\right]^{-1}(R\widehat{\beta}(\tau) - r)$$

where M_X is defined as in Section 3, then, we obtain a similar result as Theorem 3.

THEOREM 5. *Under the assumptions of Theorem 4 and the linear restriction H_0,*

$$W_n(\tau) \Rightarrow \chi_q^2,$$

where χ_q^2 is a centered Chi-square random variable with q-degrees of freedom.

Another interesting inference problem in the quantile cointegration model is the hypothesis test on constancy of the cointegrating vector β. In particular, we are interested in the hypothesis $H_{02} : \beta(\tau) = \overline{\beta}$, over $\tau \in \mathcal{T}$, where $\overline{\beta}$ is a vector of unknown constants.

A natural preliminary candidate for testing constancy of the cointegrating vector is a standardized version of $\widehat{\beta}(\tau) - \overline{\beta}$. Under the null,

$$n\left(\widehat{\beta}(\tau) - \overline{\beta}\right) \Rightarrow \frac{1}{f_\varepsilon(F_\varepsilon^{-1}(\tau))}\left[\int_0^1 \underline{B}_v\underline{B}_v^\top\right]^{-1}\int_0^1 \underline{B}_v dB_\psi^*,$$

by the result of Theorem 4. In practice, the vector of constants $\overline{\beta}$ is unknown and appropriate estimator of $\overline{\beta}$ is needed. In many econometrics applications, a n-consistent preliminary estimator of $\overline{\beta}$ is available. Denote $\widehat{\beta}$ as a preliminary estimator of $\overline{\beta}$, we look at the process

$$\widehat{V}_n(\tau) = n(\widehat{\beta}(\tau) - \widehat{\beta}).$$

Under H_{02},

$$\widehat{V}_n(\tau) \Rightarrow \frac{1}{f_\varepsilon(F_\varepsilon^{-1}(\tau))}\left[\int_0^1 \underline{B}_v\underline{B}_v^\top\right]^{-1}\int_0^1 \underline{B}_v dB_\psi^* - p\lim n\left(\widehat{\beta} - \overline{\beta}\right)$$

which depends on the preliminary estimation of β. If $\widehat{\beta}$ is the OLS estimator of in (11), under H_{02},

$$\sup_\tau \left|\widehat{V}_n(\tau)\right| \Rightarrow \sup_\tau \left|\frac{1}{f_\varepsilon(F_\varepsilon^{-1}(\tau))}\left[\int_0^1 \underline{B}_v\underline{B}_v^\top\right]^{-1}\int_0^1 \underline{B}_v d\left(B_\psi^* - f_\varepsilon(F_\varepsilon^{-1}(\tau))B_\varepsilon^*\right)\right|$$

where $B_\varepsilon^*(\cdot)$ is the limit of partial sum of ε_t. Thus, we mat test varying-coefficient behavior based on the Kolmogoroff-Smirnoff statistic $\sup_\tau \left|\widehat{V}_n(\tau)\right|$.

The necessity of estimating $\overline{\beta}$ introduces a drift component $(p\lim n\left(\widehat{\beta} - \overline{\beta}\right))$ in addition to the limit of $n\left(\widehat{\beta}(\tau) - \overline{\beta}\right)$. We may generate critical values for the statistic

$\sup_\tau \left| \widehat{V}_n(\tau) \right|$ using simulation or resampling methods. Using the usual notation $*$ to signify the bootstrap samples and P^* for the probability conditional on the original sample, we may consider the following resampling procedure:

(1) First, obtain estimates $\widehat{\beta}(\tau)$ and $\widehat{\beta}$ by quantile regression and OLS regression respectively from

$$y_t = \alpha + \beta' x_t + \sum_{j=-K}^{K} \Delta x'_{t-j} \Pi_j + \varepsilon_t.$$

Construct $\widehat{V}_n(\tau) = n(\widehat{\beta}(\tau) - \widehat{\beta})$, and obtain residuals

$$\widehat{u}_t = y_t - \widehat{\alpha} - \widehat{\beta}' x_t, \ t = 1, ..., n,$$

(2) Define $\widehat{w}_t = (v_t, \widehat{u}_t)$, $v_t = \Delta x_t$, apply a sieve (autoregression) estimation on $\widehat{w}_t$

$$\widehat{w}_t = \sum_{j=1}^{q} \widehat{B}_j \widehat{w}_{t-j} + \widehat{e}_t, \ t = q+1, ..., n,$$

and get fitted residuals $\widehat{e}_t = \widehat{w}_t - \sum_{j=1}^{q} \widehat{B}_j \widehat{w}_{t-j}$, $t = q+1, ..., n$.

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

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

with $w_j^* = \widehat{w}_j$ for $j = 1, ..., q$.

(4) Define $w_t^* = (v_t^*, u_t^*)$ in conformable with $\widehat{w}_t = (v_t, \widehat{u}_t)$, and generate x_t^* from: $x_t^* = x_{t-1}^* + v_t^*$, with $x_1^* = x_1$. Generate

$$y_t^* = \widehat{\alpha} + \widehat{\beta}' x_t^* + u_t^*$$

Thus, we obtain the bootstrapped samples (y_t^*, x_t^*).

(5) We now construct bootstrap version of $\widehat{\beta}(\tau)$, $\widehat{\beta}$, and $\widehat{V}_n(\tau)$ using the bootstrapped samples (y_t^*, x_t^*). We first calculate $\widehat{\beta}^*(\tau)$ and $\widehat{\beta}^*$ from quantile and OLS regression on

$$y_t^* = \alpha + \beta' x_t^* + \sum_{j=-K}^{K} \Delta x_{t-j}^{*\prime} \Pi_j + \varepsilon_t^*,$$

then, we construct

$$\widehat{V}_n^*(\tau) = n(\widehat{\beta}^*(\tau) - \widehat{\beta}^*).$$

In the above procedure, to make the subsequent bootstrap test valid, we generate y_t^* under the null hypothesis of constant β. The limiting null distribution of the test

statistics can then be approximated by repeating steps 2-5 many times. Let $C_t^*(\tau, \theta)$ be the (100θ)-th quantiles, i.e.,

$$\mathrm{P}^* \left[\sup_\tau \left| \widehat{V}_n^*(\tau) \right| \le C_t^*(\tau, \theta) \right] = \theta,$$

then the hypothesis of constant cointegrating coefficients will be rejected at the $(1-\theta)$ level if $\sup_\tau \left| \widehat{V}_n(\tau) \right| \le C_t^*(\tau, \theta)$.

Alternatively, instead of using resampling methods, we may directly simulate the Brownian motions

$$\frac{1}{f_\varepsilon(F_\varepsilon^{-1}(\tau))} \left[\int_0^1 \underline{B}_v \underline{B}_v^\top \right]^{-1} \int_0^1 \underline{B}_v d \left(B_\psi^* - f_\varepsilon(F_\varepsilon^{-1}(\tau)) B_\varepsilon^* \right).$$

In particular, we may replace the regressions in step 5 by, say, directly approximating $\int_0^1 \underline{B}_v \underline{B}_v^\top$ and $\int_0^1 \underline{B}_v dB_\psi^*$ using

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

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

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

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

3.3   A Robust Test For Cointegration

Quantile cointegrating regression not only provides a robust method for many financial applications such as portfolio management, but also expands the modeling options for economic time series. The proposed method can be used to develop new tools for improved inference on cointegrated time series.

Denoting $\psi_\tau(u) = \tau - I(u < 0)$, and consider the quantile regression residual

$$\varepsilon_{t\tau} = y_t - Q_{y_t}(\tau | \mathcal{F}_t) = y_t - \Theta(\tau)' Z_t = \varepsilon_t - F_\varepsilon^{-1}(\tau),$$

then we have $Q_{\varepsilon_{t\tau}}(\tau) = 0$, where $Q_{\varepsilon_{t\tau}}(\tau)$ signifies the τ-th quantile of $\varepsilon_{t\tau}$, and

$$E\psi_\tau(\varepsilon_{t\tau}) = 0.$$

The cointegration relationship may be tested by directly looking at the fluctuation in the residual process $\varepsilon_{t\tau}$ from the quantile cointegrating regression. In the case of

15

cointegration, the residual process should be stable and fluctuations in the residuals reflect only equilibrium errors. Otherwise, the fluctuations in the residuals can be expected to be of a larger order of magnitude. Thus, cointegration can be tested based on $\varepsilon_{t\tau}$. If we consider the following partial sum process

$$Y_n(r) = \frac{1}{\omega_\psi^* \sqrt{n}} \sum_{j=1}^{[nr]} \psi_\tau(\varepsilon_{j\tau}),$$

where ω_ψ^{*2} is the long run variance of $\psi_\tau(\varepsilon_{j\tau})$, under appropriate assumptions, the partial sum process follow an invariance principle and converges weakly to a standard Brownian motion $W(r)$. Choosing a continuous functional $h(\cdot)$ that measures the fluctuation of $Y_n(r)$, notice that $\psi_\tau(\varepsilon_{j\tau})$ is indicator-based, a robust test for cointegration can be constructed based on $h(Y_n(r))$. By the continuous mapping theorem, under regularity conditions and the null of cointegration,

$$h(Y_n(r)) \Rightarrow h(W(r)).$$

In principle, any metric that measures the fluctuation in $Y_n(r)$ is a natural candidate for the functional h. The classical Kolmogoroff-Smirnoff type or Cramer-von Mises type measures are of particular interest. Under the alternative of no cointegration, the statistic diverges to ∞.

In practice, we estimate $\Theta(\tau)$ by $\widehat{\Theta}(\tau)$ using (12), and obtain the residuals

$$\widehat{\varepsilon}_{t\tau} = y_t - \widehat{\Theta}(\tau)' Z_{t,K}$$

A robust test for cointegration can then be constructed based on

$$\widehat{Y}_n(r) = \frac{1}{\widehat{\omega}_\psi^* \sqrt{n}} \sum_{j=1}^{[nr]} \psi_\tau(\widehat{\varepsilon}_{j\tau}).$$

where $\widehat{\omega}_\psi^{*2}$ is a consistent estimator of ω_ψ^{*2}. Under regularity assumptions and the hypothesis of cointegration,

$$\widehat{Y}_n(r) \Rightarrow \widetilde{W}(r) = W_1(r) - \left[\int_0^1 dW_1 \overline{W}_2' \right] \left[\int_0^1 \overline{W}_2 \overline{W}_2' \right]^{-1} \int_0^r \overline{W}_2(s),$$

where $\overline{W}_2(r) = (1, W_2(r)')'$, W_1 and W_2 are independent 1 and k-dimensional standard Brownian motions. (see Tyurin and Xiao (2006) for more discussion on robust tests for cointegration.)

4   Monte Carlo Results

A Monte Carlo experiment was conducted to examine the finite sample performance of quantile regression with cointegrated time series. We focus on two important issues

in our Monte Carlo study: (1) The efficiency gain of robust method such as quantile regression over OLS in cointegration models with non-Gaussian innovations; and (2) Application of quantile cointegrating regression on the study of time varying behavior in cointegration.

For case (1), the data were generated from the following bivariate regression model

$$y_t = \alpha + \beta x_t + u_t, \text{ with } \alpha = 0, \ \beta = 1,$$

where

$$\Delta x_t = v_t, t = 1, ..., n.$$

We compare the OLS estimator of β with the Median regression estimator for different data generating processes. In particular, we report results for various cases when u_t and v_t are iid Normal or student-t with degrees of freedom 2, 3, 4. The initial values are all set to be zero.

Tables 1A and 1B report the standard errors (STD) and mean-squared-errors (MSE) of the median regression ($\tau = 0.5$) vs. OLS estimation of β for sample sizes $T = 100$, and $T = 200$. Number of repetitions is 500.

Table 1A: OLS v.s. Median Regression Estimation of β, $T = 100$		OLS		Median Reg.	
		STD	MSE	STD	MSE
$v_t \sim N(0,1)$	$u_t \sim t(3)$	0.0656	0.0043	0.041	0.0017
	$u_t \sim t(4)$	0.0535	0.0028	0.0402	0.0016
$v_t \sim t(4)$	$u_t \sim t(3)$	0.0408	0.0017	0.0316	0.00099
	$u_t \sim t(4)$	0.0331	0.0011	0.0299	0.00089
$v_t \sim t(3)$	$u_t \sim t(3)$	0.0346	0.0012	0.0283	0.00079
Table 1B: OLS v.s. Median Regression Estimation of β, $T = 200$		OLS		Median Reg.	
		STD	MSE	STD	MSE
$v_t \sim N(0,1)$	$u_t \sim t(3)$	0.0267	0.00071	0.020	0.00040
$v_t \sim N(0,1)$	$u_t \sim t(4)$	0.0206	0.00042	0.0188	0.00035
$v_t \sim t(4)$	$u_t \sim t(3)$	0.02	0.000398	0.0154	0.000235
$v_t \sim t(4)$	$u_t \sim t(4)$	0.0162	0.000262	0.0151	0.000228
$v_t \sim t(3)$	$u_t \sim t(3)$	0.0179	0.000321	0.0142	0.000200

We then study cointegration with time-varying cointegrating coefficient using quantile regression. We consider the following model

$$y_t = \alpha + \beta_t x_t + u_t.$$

where β_t may vary over different quantile of the innovation distribution. Conditional on $\mathcal{F}_{t-1}$, the τ-th quantile of y_t is given by

$$Q_{y_t}(\tau | \mathcal{F}_{t-1}) = \alpha + \beta(\tau)' x_t + F_u^{-1}(\tau),$$

We are interested in testing the hypothesis that $\beta(\tau) = $ constant over τ. We consider the following two choices of β_t:

$$
\begin{aligned}
\text{(i). } \beta_t &= 1, \\
\text{(ii). } \beta_t &= \beta(u_t) = \left\{ \begin{array}{ll} 1, & u_t \geq 0, \\ -1, & u_t < 0, \end{array} \right. .
\end{aligned}
$$

When $\beta_t = 1$, it is constant over all quantiles and thus the empirical rejection rates corresponds to the empirical size. In the second choice, $\beta_t = \beta(u_t)$, the cointegrating vector takes different values over different quantiles of the error distribution and thus the rejection rates corresponds to the empirical power.

The data in our second experiment were generated from different distributions of $\{u_t\}$ and $\{v_t\}$. Again, we consider u_t and v_t being i.i.d. random variables of Normal and student-t with different degrees of freedoms. The bootstrap based procedure introduced in Section 3.2. is conducted to test the varying-coefficient behavior in the cointegration system for different sample sizes ($T = 100$ and 200). The number of repetitions is 500. Representative results of the empirical size and power of the test are reported in Table 2.

Table 2: Testing for Time Varying Cointegrating Parameter					
		$T = 100$		$T = 200$	
		Size	*Power*	*Size*	*Power*
$v_t \sim N(0,1)$	$u_t \sim N(0,1)$	7%	54%	6.6%	87.5%
$v_t \sim N(0,1)$	$u_t \sim t(3)$	5.6%	75%	5.2%	90%
$v_t \sim N(0,1)$	$u_t \sim t(4)$	4.5%	60%	5%	89.5%
$v_t \sim t(3)$	$u_t \sim t(3)$	8%	62%	5.2%	90%
$v_t \sim t(4)$	$u_t \sim t(4)$	6.7%	56%	6.5%	95%

Information in Tables 1A and 1B indicates that efficiency gain can be achieved from a robust cointegrating regression in the presence of non-normal distributed data. From the Monte Carlo results in Table 2, we can see that the quantile regression based tests for varying-coefficients have reasonable size and good power in finite sample. We can also see improved sampling performance as the sample size increases, corroborating the asymptotic theory.

5 An Empirical Application to Asset Pricing Model

In this section, we apply the quantile cointegration model to stock index data from the U.S. In particular, we collected price and dividend yield data for the Standard and Poor (S&P) 500 Index from January 1974 to September 1998. The source of the data is the on-line service of Datastream. We analyze the relationship between prices and market fundamentals using the quantile cointegrating regression.

If we consider the standard rational expectations model (6) for stock prices, a forward-looking solution to this model indicates that stock prices (P_t) and market fundamentals (D_t) should be cointegrated. However, there has been concern about a direct regression based on this specification: violation of limited liability. For instance, if the conditional distribution of the prices is normal, then there will always be a positive probability of obtaining a negative price (see, e.g., Campbell, Lo and MacKinlay (1997, p32)). For this reason, many researchers consider the above rational expectations model in terms of logarithms of price and dividend. Following Campbell and Shiller (1988), we write the log linear approximation of (6) as

$$p_t + q = \kappa + \delta E_t p_{t+1} + (1 - \delta)E_t d_{t+1} \tag{13}$$

where p_t and d_t are logarithms of P_t and D_t, q is the log gross return rate, δ is the average ratio of the stock price to the sum of the stock price and the dividend ($0 < \delta < 1$), and κ is a function of δ. Under the transversality condition that

$$\lim_{k \to \infty} \delta^k E_t p_{t+k} = 0, \tag{14}$$

the unique forward-looking market fundamental solution to (13) is given by

$$p_t = \eta + (1 - \delta) \sum_{j=0}^{\infty} \delta^j E_t d_{t+1+j}. \tag{15}$$

Since d_t appears to be nonstationary in empirical analyses, it is usually characterized as an integrated process with drift:

$$d_{t+1} = \mu + d_t + \varepsilon_t, \tag{16}$$

where ε_t is an I(0) process of innovations with $E(\varepsilon_t) = 0$. Combining (15) and (16), we have

$$p_t = \alpha + \beta d_t. \tag{17}$$

Thus p_t is also an integrated process with drift. Although both p_t and d_t are nonstationary, there exists a long run equilibrium relationship between p_t and d_t, and the linear combination of p_t and d_t ($p_t - \alpha - \beta d_t$) is I(0). Fluctuations in the residual process $p_t - \alpha - \beta d_t$ are simply equilibrium errors and thus are covariance stationary. In other words, p_t and d_t are cointegrated.

Regression model (17) has been examined by many empirical researchers based on OLS technique. A very important feature from the previous analysis is that stock prices are much more volatile than market fundamentals such as dividends. Estimate of the cointegrating parameter β based on OLS regression over (17) displays a lot of variability. Consequently, in portfolio management, frequent rebalancing is needed to keep the portfolio in line with the index. These empirical observations suggest

that the value of cointegrating parameter is changing over time. In this section, we examine the relationship between prices and market fundamentals using a quantile cointegrating regression.

To examine various specifications in modeling asset prices, we consider the following two models.

Model 1. The first model is built on (17). Using leads and lags to absorb the endogeneity, we have

$$p_t = \alpha + \beta d_t + \sum_{j=-K}^{K} \pi_j \Delta d_{t-j} + \varepsilon_t.$$

Our first empirical model is an extension of the above model (to cointegration with varying cointegrating coefficients). Notice that in the above model the market price is characterized by (6) with a constant rate of return. Indeed, many empirical applications in asset pricing are based on OLS regression on this cointegration relationship [see, e.g. Campbell and Shiller (1987), Gordon (1962), Evans (1991), among others]. Relaxing the assumption of a constant rate of return will substantially complicate the forward-looking solution to the rational expectation model. In general, there is no simple analytical solution unless we impose additional assumptions on the associated conditional expectation.

If we consider the general model which allows γ to change over time:

$$P_t = E_t \left\{ \frac{P_{t+1} + D_{t+1}}{1 + \gamma_t} \right\}. \tag{18}$$

Solving equation (18) recursively and denoting the growth rate of real dividend as g_t, we obtain the following expression for the fundamental value of asset prices:

$$P_t = E_t \left\{ \sum_{j=0}^{\infty} \left(\prod_{i=0}^{j} \left[\frac{1 + g_{t+i}}{1 + \gamma_{t+i}} \right] \right) \right\} D_t.$$

Gordon (1962) assumes that γ and g are constant and thus we obtain the conventional cointegration model (16). In order to generalize the conventional cointegration model (16) and allow for time varying coefficient, appropriate simplification has to be introduced. In particular, using the log linear approximation, we have

$$p_t = \alpha + \beta_t d_t + u_t, \tag{19}$$

where β_t is a function of

$$E_t \left\{ \sum_{j=0}^{\infty} \left(\prod_{i=0}^{j} \left[(1 + g_{t+i})/(1 + \gamma_{t+i}) \right] \right) \right\}.$$

20

Appropriate approximation of β_t has to be used. For example, Barsky and DeLong (1993) considered an extension by imposing additional assumptions on the construction of g. Donaldson and Kamstra (1996) use similar idea in estimating market fundamental. The first model in our empirical application is built on above idea and considers the following quantile cointegrating regression:

$$p_t = \alpha + \beta_t d_t + \sum_{j=-K}^{K} \pi_{jt} \Delta d_{t-j} + \varepsilon_t. \tag{20}$$

where the cointegrating coefficient β_t (and other coefficients) are time varying, depending on the new information (or shocks) received in the period. Thus, the cointegrating coefficient β is in the form of a function of the innovation process ε_t. Such a model is quantile dependent and captures additional volatility in stock prices P_t. Conditional on past information, the above model has the following quantile domain representation:

$$Q_{p_t}(\tau|\mathcal{F}_t) = \alpha(\tau) + \beta(\tau)d_t + \sum_{j=-K}^{K} \pi_j(\tau) \Delta d_{t-j}. \tag{21}$$

We now apply quantile regression to the above model. Quantile regression estimates of the cointegrating coefficients are reported in Table 1 below.

Table 1: Quantile Cointegration Estimates Based on Model 1							
	$\tau = 0.05$	$\tau = 0.10$	$\tau = 0.15$	$\tau = 0.20$	$\tau = 0.25$	$\tau = 0.30$	$\tau = 0.35$
$\widehat{\beta}(\tau)$	40.7929	43.0539	43.3579	43.9045	44.3159	44.9067	46.2997
	$\tau = 0.40$	$\tau = 0.45$	$\tau = 0.50$	$\tau = 0.55$	$\tau = 0.60$	$\tau = 0.65$	$\tau = 0.70$
$\widehat{\beta}(\tau)$	47.2917	47.4789	47.3677	47.2634	47.2675	48.0674	49.1072
	$\tau = 0.75$	$\tau = 0.80$	$\tau = 0.85$	$\tau = 0.90$	$\tau = 0.95$		
$\widehat{\beta}(\tau)$	50.9336	53.3779	59.6298	68.7037	77.4728		
OLS estimate: $\widehat{\beta} = 57.8924$							

Model 2. Model 1 considers dividends as the major source of market fundamental. Besides dividends, other fundamental sources of stock price may also be accounted for. Consequently, other covariates that help explaining market fundamentals may also be included in the cointegrating regression. For example, another variable that provides useful information might be the short term interest rate r_t. For this reason, we may include r_t as another explanatory variable in the cointegrating regression, then we obtain an extension of (19):

$$p_t = \alpha + \beta_t d_t + \lambda_t r_t + v_t, \tag{22}$$

21

Again, using leads and lags to absorb the endogeneity, we consider the following cointegrating regression:

$$p_t = \alpha + \beta_t d_t + \lambda_t r_t + \sum_{j=-K}^{K} \pi_{jt} \Delta d_{t-j} + \sum_{j=-K}^{K} \gamma_{jt} \Delta r_{t-j} + \varepsilon_t. \qquad (23)$$

where the cointegrating coefficients are allowed to be time varying and thus quantile dependent:

$$Q_{p_t}(\tau | \mathcal{F}_t) = \alpha(\tau) + \beta(\tau) d_t + \lambda(\tau) r_t + \sum_{j=-K}^{K} \pi_j(\tau) \Delta d_{t-j} + \sum_{j=-K}^{K} \gamma_j(\tau) \Delta r_{t-j}.$$

Quantile regression estimates of the cointegrating coefficients based on Model 2 are reported in Table 2 below.

Table 2: Quantile Cointegration Estimates Based on Model 2							
	$\tau = 0.05$	$\tau = 0.10$	$\tau = 0.15$	$\tau = 0.20$	$\tau = 0.25$	$\tau = 0.30$	$\tau = 0.35$
$\widehat{\beta}(\tau)$	34.8271	36.7676	37.2902	37.4863	37.8032	38.0976	38.5612
	$\tau = 0.40$	$\tau = 0.45$	$\tau = 0.50$	$\tau = 0.55$	$\tau = 0.60$	$\tau = 0.65$	$\tau = 0.70$
	39.1491	39.2799	40.0530	40.7666	41.4132	43.0672	47.4614
	$\tau = 0.75$	$\tau = 0.80$	$\tau = 0.85$	$\tau = 0.90$	$\tau = 0.95$		
$\widehat{\beta}(\tau)$	50.1081	52.4722	59.4586	69.6467	73.5931		
OLS estimate: $\widehat{\beta} = 54.4848$							

The evidence based on these point estimates of the cointegrating coefficients at each quantile suggests that the cointegrating coefficients are not constant over time - thus bringing additional volatility into asset prices in addition to market fundamentals. The cointegrating coefficient estimate $\widehat{\beta}(\tau)$ has different values over different quantiles (ranging from 40.7929 (or 34.827) in low quantiles to 77.4728 (or 73.5931) in upper quantiles in model 1 (model 2), displaying asymmetric dynamics over time. In particular, $\widehat{\beta}(\tau)$ increases when we move from lower quantiles to higher quantiles.

Formal tests for varying-coefficient cointegration relationship is also conducted using the bootstrap-based test proposed in Section 3.2.. For Model 1, the calculated test statistic $\sup_\tau \left| \widehat{V}_n(\tau) \right| = 2383.3$, and the 1%, 5%, and 10% bootstrapped critical values are 173.18, 92.43, 67.86, respectively. For Model 2, the calculated test statistic $\sup_\tau \left| \widehat{V}_n(\tau) \right| = 2387.5$, and the 1%, 5%, and 10% bootstrapped critical values are 124.25, 98.37, 71.03, respectively. In both models, the null hypothesis of constant cointegrating coefficients are rejected even at 1% level, displaying a strong evidence of varying-coefficient behavior.

6 Conclusions and Generalizations

Quantile cointegrating regression not only provides a robust method for many financial applications such as portfolio management, but also expands the modeling options for economic time series. The proposed models indicate that there might be important information about cointegration models which are not detectable from the traditional OLS based analysis.

Some important future extensions of the quantile cointegration model can be conducted. First, quantile regression analysis can be extended to cointegration models with infinite variance errors. In this case, the limiting theory will be different. Faster rate of convergence can be found and mixture normal asymptotics can be achieved without fully-modification. Second, the quantile cointegrating regression model may be extended to the case with general functional coefficients $\theta(z_t)$. A quantile cointegrating regression model with general functional coefficients take the following form:

$$Q_{y_t}(\tau|z_t, X_t) = \theta_\tau(z_t)'X_t.$$

We may apply the local polynomial method to the above quantile regression model.

7 Appendix: Proofs

7.1 Proof of Theorem 1

The following results is useful in developing asymptotics for the regression quantile estimates: For $u \neq 0$,

$$\rho_\tau(u - v) - \rho_\tau(u) = -v\psi_\tau(u) + (u - v)\{I(0 > u > v) - I(0 < u < v)\}, \qquad (24)$$

where $\psi_\tau(u) = \tau - I(u < 0)$.

Let $u_{t\tau} = u_t - F^{-1}(\tau)$, then $Q_{u_{t\tau}}(\tau) = 0$! If we denote

$$(\alpha + F^{-1}(\tau), \beta)' = \theta(\tau)$$

then $u_{t\tau} = y_t - \theta(\tau)'z_t$.

Further we denote $\widehat{v} = D_n(\widehat{\theta}(\tau) - \theta(\tau))$, where $D_n = diag(\sqrt{n}, n, \cdots, n)$, $\rho_\tau(y_t - \widehat{\theta}(\tau)'z_t) = \rho_\tau(u_{t\tau} - (D_n^{-1}\widehat{v})'z_t)$. Minimization (2) is equivalent to the following problem:

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

If $\widehat{v}$ is a minimizer of $Z_n(v)$, we have $\widehat{v} = D_n(\widehat{\theta}(\tau) - \theta(\tau))$.

The objective function $Z_n(v) = \sum_{t=1}^n \left[\rho_\tau(u_{t\tau} - (D_n^{-1}v)'z_t) - \rho_\tau(u_{t\tau})\right]$ is a convex random function and is similar to the one of Knight (1989). Knight (1989) (also see similar

argument in Pollard 1991) shows that if the finite-dimentional distributions of $Z_n(\cdot)$ converge weakly to those of $Z(\cdot)$ and $Z(\cdot)$ has a unique minimum, the convexity of $Z_n(\cdot)$ implies that $\widehat{v}$ converges in distribution to the minimizer of $Z(\cdot)$. Notice that $Q_{u_{t\tau}}(\tau) = 0$, we have

$$E\psi_\tau(u_{t\tau}) = 0.$$

In general u_t and Δx_t are correlated and thus B_ψ and B_x are correlated Brownian motions. Under Assumption A, the vector partial sum process $\{\psi_\tau(u_{t\tau}), \Delta x_t'\}$ follow an invariance principle that

$$n^{-1/2} \sum_{t=1}^{[nr]} \begin{bmatrix} \psi_\tau(u_{t\tau}) \\ \Delta x_t \end{bmatrix} \Rightarrow B(r) = BM(\Omega)$$

a Brownian motion with covariance matrix

$$\Omega = \begin{bmatrix} \omega_\psi^2 & \omega_{x\psi} \\ \omega_{x\psi} & \omega_x^2 \end{bmatrix},$$

and

$$n^{-1} \sum_{t=1}^{n} x_t \psi_\tau(u_{t\tau}) \Rightarrow \int_0^1 B_x dB_\psi + \lambda_{x\psi}$$

where $\lambda_{x\psi}$ is the one sided long-run variance between Δx_t and $\psi_\tau(u_{t\tau})$.

Using the result of (24), the objective function of minimization problem can be written as

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

$$= -\sum_{t=1}^{n} (D_n^{-1}v)'z_t \psi_\tau(u_{t\tau})$$

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

For the first term, under Assumptions A and B,

$$D_n^{-1} \sum_{t=1}^{n} z_t \psi_\tau(u_{t\tau}) = \begin{bmatrix} n^{-1/2} \sum_{t=1}^{n} \psi_\tau(u_{t\tau}) \\ n^{-1} \sum_{t=1}^{n} x_t \psi_\tau(u_{t\tau}) \end{bmatrix} \Rightarrow \begin{bmatrix} \int_0^1 dB_\psi \\ \int_0^1 B_x dB_\psi + \lambda_{x\psi} \end{bmatrix}.$$

Next we examine the limit of

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

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

We first consider the limit of

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

If we denote $v = D_n(\theta - \theta(\tau))$, and partition v and $\theta(\tau)$ conformable with $z_t = (1, x_t')'$, we denote

$$v = \left[\begin{array}{c} v_1 \\ v_2 \end{array}\right], \theta(\tau) = \left[\begin{array}{c} \alpha(\tau) \\ \beta \end{array}\right].$$

For convenience of asymptotic analysis, we denote

$$W_n(v) = \sum_{t=1}^{n}(v'D_n^{-1}z_t - u_{t\tau})I(0 < u_{t\tau} < v'D_n^{-1}z_t) = \sum_{t=1}^{n}\xi_t(v),$$

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

To avoid technical problems in taking conditional expectations, we consider truncation of $v'D_n^{-1}z_t$ at some finite number $m > 0$ and denote

$$W_{nm}(v) = \sum_{t=1}^{n}\xi_{tm}(v),$$

$$\xi_{tm}(v) = (v'D_n^{-1}z_t - u_{t\tau})I(0 < u_{t\tau} < v'D_n^{-1}z_t)I(v'D_n^{-1}z_t \le m).$$

Denote the information set upto time t as $F_{t-1} = \sigma\{u_{t-j}, v_{t-j+1}, j \ge 1\}$, then $z_t \in F_{t-1}$. We further define

$$\overline{\xi}_{tm}(v) = E\{(v'D_n^{-1}z_t - u_{t\tau})I(0 < u_{t\tau} < v'D_n^{-1}z_t)I(v'D_n^{-1}z_t \le m)|\mathcal{F}_{t-1}\},$$

and

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

then $\{\xi_{tm}(v) - \overline{\xi}_{tm}(v)\}$ is a martingale difference sequence. Denote the conditional distribution function

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

and its derivative as $f_{t-1}(\bullet)$, $a.s.$, and assume that $f_{t-1}(s_n)$ is uniformly integrable for any

25

sequence $s_n \to F^{-1}(\tau)$.

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

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

$$= \sum_{t=1}^{n} E\left\{(v'D_n^{-1}z_t + F^{-1}(\tau) - u_t)\right.$$

$$\left.I(F^{-1}(\tau) < u_t < v'D_n^{-1}z_t + F^{-1}(\tau))I(v'D_n^{-1}z_t \le m)|\mathcal{F}_{t-1}\right\}$$

$$= \sum_{t=1}^{n} \int_{F^{-1}(\tau)}^{[v'D_n^{-1}z_t+F^{-1}(\tau)]I(v'D_n^{-1}z_t\le m)} [\int_r^{[v'D_n^{-1}z_t+F^{-1}(\tau)]I(v'D_n^{-1}z_t\le m)} ds]f_{t-1}(r)dr$$

$$= \sum_{t=1}^{n} \int_{F^{-1}(\tau)\le s\le[v'D_n^{-1}z_t+F^{-1}(\tau)]I(v'D_n^{-1}z_t\le m)} \int_{F^{-1}(\tau)\le r\le s} f_{t-1}(r)drds$$

$$= \sum_{t=1}^{n} \int_{F^{-1}(\tau)}^{[v'D_n^{-1}z_t+F^{-1}(\tau)]I(v'D_n^{-1}z_t\le m)} [s - F^{-1}(\tau)]\left[\frac{F_{t-1}(s) - F_{t-1}(F^{-1}(\tau))}{s - F^{-1}(\tau)}\right] ds.$$

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

$$\overline{W}_{nm}(v) = \sum_{t=1}^{n} \int_{F^{-1}(\tau)}^{[v'D_n^{-1}z_t+F^{-1}(\tau)]I(v'D_n^{-1}z_t\le m)} [s - F^{-1}(\tau)]f_{t-1}[F^{-1}(\tau)]ds + o_p(1)$$

$$= \sum_{t=1}^{n} f_{t-1}[F^{-1}(\tau)]\left\{\frac{[s - F^{-1}(\tau)]^2}{2} \Big|_{F^{-1}(\tau)}^{[v'D_n^{-1}z_t+F^{-1}(\tau)]I(v'D_n^{-1}z_t\le m)}\right\} + o_p(1)$$

$$= \frac{1}{2}\sum_{t=1}^{n} f_{t-1}[F^{-1}(\tau)][v'D_n^{-1}z_t]^2 I(v'D_n^{-1}z_t \le m) + o_p(1)$$

$$= \frac{1}{2n}\sum_{t=1}^{n} f_{t-1}[F^{-1}(\tau)]v'[\sqrt{n}D_n^{-1}z_t z_t' D_n^{-1}\sqrt{n}]vI(v'D_n^{-1}z_t \le m) + o_p(1)$$

By Assumption C and stationarity of $\{f_{t-1}[F^{-1}(\tau)]\}$, we have

$$\sup_{0\le r\le 1}\left|\frac{1}{n^{1-\varepsilon}}\sum_{t=1}^{[nr]}[f_{t-1}[F^{-1}(\tau)] - f[F^{-1}(\tau)]]\right| \xrightarrow{P} 0$$

for some $\varepsilon > 0$. Thus

$$\overline{W}_{nm}(v) \Rightarrow \frac{1}{2}f[F^{-1}(\tau)]v'\left\{\int_0^1 \begin{bmatrix} 1 & \int_0^1 B_x \\ \int_0^1 B_x & \int_0^1 B_x B_x' \end{bmatrix} I(0 < v_1 + v_2'B_x(s) \le m)\right\}v := \eta_m$$

We now follow the arguments of Pollard (1991), notice that $(v'D_n^{-1}z_t)I(0 \le v'D_n^{-1}z_t \le m) \xrightarrow{P} 0$ uniformly in t,

$$\sum_{t=1}^{n} E[\xi_{tm}(v)^2|\mathcal{F}_{t-1}] \le \max\{(v'D_n^{-1}z_t)I(0 \le v'D_n^{-1}z_t \le m)\}\sum \overline{\xi}_{tm}(v) \to 0.$$

26

thus the following summation of martingale difference sequence

$$\sum_t \{\xi_{tm}(v) - \bar{\xi}_{tm}(v)\}$$

converges to zero in probability. By the Asymptotic Equivalence Lemma, the limiting distribution of $\sum_t \xi_{tm}(v)$ is the same as that of $\sum_t \bar{\xi}_{tm}(v)$, i.e.,

$$W_{nm}(v) \Rightarrow \eta_m$$

Let $m \to \infty$, we have

$$\eta_m \Rightarrow \frac{1}{2} f(F^{-1}(\tau)) v' \left[\int_0^1 B_z B_z' \right] v I(v' B_z(s) > 0) = \eta,$$

where $B_z(s) = (1, B_x(s)')$.

Now we show that

$$\lim_{m \to \infty} \lim \sup_{n \to \infty} \Pr[|W_n(v) - W_{nm}(v)| \geq \varepsilon] = 0.$$

This holds because

$$
\begin{aligned}
&= \Pr[|W_n(v) - W_{nm}(v)| \geq 0] \\
&= \Pr\left[\sum_t (v' D_n^{-1} z_t - u_{t\tau}) I(0 < u_{t\tau} < v' D_n^{-1} z_t) I(v' D_n^{-1} z_t > m) > 0 \right] \\
&\leq \Pr\left[\cup_t \{v' D_n^{-1} z_t > m\} \right] \\
&= \Pr\left[\max_t \{v' D_n^{-1} z_t\} > m \right],
\end{aligned}
$$

and

$$\lim_{m \to \infty} \Pr[\sup_{1 \leq r \leq 1} v' B_z(r) > m] = 0.$$

By Billingsley (1968),

$$W_{nm}(v) \Rightarrow \eta,$$

i.e.

$$\sum_{t=1}^n ((D_n^{-1} v)' z_t - u_{t\tau}) I(0 < u_{t\tau} < (D_n^{-1} v)' z_t) \Rightarrow \frac{1}{2} f(F^{-1}(\tau)) v' \int_0^1 B_z B_z' v$$

Similarly, we can show that

$$\sum_{t=1}^n (u_{t\tau} - (D_n^{-1} v)' z_t) \{ I(0 > u_{t\tau} > (D_n^{-1} v)' z_t) \Rightarrow \frac{1}{2} f(F^{-1}(\tau)) v' \left[\int_0^1 B_z B_z' \right] v I(v' B_z(s) < 0).$$

Thus,

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

$$\Rightarrow f(F^{-1}(\tau)) v' \left[\begin{array}{cc} 1 & \int_0^1 B_x \\ \int_0^1 B_x & \int_0^1 B_x B_x' \end{array} \right] v.$$

27

As a result,

$$Z_n(v)$$

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

$$= -\sum_{t=1}^{n} (D_n^{-1}v)'z_t\psi_\tau(u_{t\tau})$$

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

$$\Rightarrow \quad -v'\begin{bmatrix} \int_0^1 dB_\psi \\ \int_0^1 B_x dB_\psi + \Delta_{x\psi} \end{bmatrix} + f(F^{-1}(\tau))v'\begin{bmatrix} 1 & \int_0^1 B_x \\ \int_0^1 B_x & \int_0^1 B_x B_x' \end{bmatrix} v := Z(v)$$

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

$$\frac{1}{2f(F^{-1}(\tau))}\begin{bmatrix} 1 & \int_0^1 B_x \\ \int_0^1 B_x & \int_0^1 B_x B_x' \end{bmatrix}^{-1}\begin{bmatrix} \int_0^1 dB_\psi \\ \int_0^1 B_x dB_\psi + \lambda_{x\psi} \end{bmatrix}$$

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

$$D_n(\widehat{\theta}(\tau) - \theta(\tau)) \quad \Rightarrow \quad \frac{1}{2f(F^{-1}(\tau))}\begin{bmatrix} 1 & \int_0^1 B_x \\ \int_0^1 B_x & \int_0^1 B_x B_x' \end{bmatrix}^{-1}\begin{bmatrix} \int_0^1 dB_\psi \\ \int_0^1 B_x dB_\psi + \lambda_{x\psi} \end{bmatrix}.$$

7.2   Proof of Theorem 2

By result of Theorem 1, the limiting distribution of $n(\widehat{\beta}(\tau) - \beta(\tau))$ can be written as

$$\frac{1}{f(F^{-1}(\tau))}\left[\int_0^1 \underline{B}_v \underline{B}_v^\top\right]^{-1}\int_0^1 \underline{B}_v dB_{\psi.v} +$$

$$\frac{1}{f(F^{-1}(\tau))}\left[\int_0^1 \underline{B}_v \underline{B}_v^\top\right]^{-1}\left[\int_0^1 \underline{B}_v dB_v^\top \Omega_{vv}^{-1}\Omega_{v\psi} + \lambda_{v\psi}\right]$$

In addition, $\widehat{\Omega}_{vv}$, $\widehat{\Omega}_{v\psi}$, $\widehat{\lambda}_{v\psi}$, and $\widehat{\lambda}_{vv}$ are consistent estimates of Ω_{vv}, $\Omega_{v\psi}$, $\lambda_{v\psi}$, λ_{vv}, thus

$$D_n\left(\widehat{\theta}(\tau)^+ - \theta(\tau)\right)$$

$$= \begin{pmatrix} \sqrt{n}\,[\widehat{\alpha}(\tau) - \alpha(\tau)] \\ n\left[\widehat{\beta}(\tau)^+ - \beta(\tau)\right] \end{pmatrix}$$

$$= \begin{pmatrix} \sqrt{n}\,[\widehat{\alpha}(\tau) - \alpha(\tau)] \\ n\left[\widehat{\beta}(\tau) - \beta(\tau)\right] \end{pmatrix} - \begin{pmatrix} 0 \\ -\frac{1}{f(\widehat{F}^{-1}(\tau))}\left[\frac{1}{n^2}\sum_t x_t x_t'\right]^{-1}\left[\frac{1}{n}\sum_t x_t v_t'\widehat{\Omega}_{vv}^{-1}\widehat{\Omega}_{v\psi} + \widehat{\lambda}_{v\psi}^+\right] \end{pmatrix}$$

$$\Rightarrow \quad \frac{1}{f(F^{-1}(\tau))}\left[\int_0^1 \overline{B}_v \overline{B}_v^\top\right]^{-1}\int_0^1 \overline{B}_v dB_{\psi.v}.$$

7.3 Proof of Theorem 3

This is an immediate consequence of Theorem 2. Under the null hypothesis $H_0 : R\beta = r$ and by the result of Theorem 2, we have

$$n(R\widehat{\beta}^+(\tau) - r) \Rightarrow N(0, \frac{\omega_{\psi.v}^2}{f(F^{-1}(\tau))^2} \left[R \left(\int_0^1 \underline{B}_v \underline{B}_v^\top \right)^{-1} R^\top \right]),$$

and thus the results of Theorem 3 follows immediately.

7.4 Proof of Theorem 4

This is similar to the proof of Theorem 1.

7.5 Proof of Theorem 5

This is similar to the proof of Theorem 3.

References

Acerbi, C., and D. Tasche, 2002, Expected Shortfall: A Natural Coherent Alternative to Value at Risk," *Economic Notes*, 31, 379-388.

Alexander, C.O., 1999, Optimal hedging using cointegration, *Philosophical Transactions of the Royal Society* A 357, pp. 2039-2058.

Andrews, D.W.K., 1991, "Heteroskedasticity and autocorrelation consistent covariance matrix estimation," *Econometrica*, **59**, 817–858.

Artzner, P., F. Delbaen, J. Eber, and D. Heath, 1999, Coherent Measures of Risk, *Mathematical Finance*, 9, 203-228.

Barsky, R., and B. Delong, 1993, "Why Does the Stock Market Fluctuate?" *Quarterly Journal of Economics*, 108, 291-311.

Bassett, G., R. Koenker, and G. Kordas, 2004, Choquet Risk and Portfolio Optimization, *Journal of Financial Econometrics*.

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

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

Brillinger, D. R., 1980, *Time Series: Data Analysis and Theory.* New York: Holt, Rinehart and Winston.

Campbell, J. Y., and R. J. Shiller, 1987, "Cointegration and Tests of Present Value Models," *Journal of Political Economy*, 95(5), 1062-88.

Campbell, J. Y., and R. J. Shiller, 1988, "The Dividend-Price Ratio and Expectations of Future Dividends and Discount Factors," *Review of Financial Studies*, 1, 195-228.

Campbell, J. Y., A. Lo, and MacKinley, 1997, *The Econometrics of Financial Market*, Princeton University Press.

Cerchi, M. and A. Havenner, 1988, Cointegration and Stock Prices, *Journal of Economic Dynamics and Control*, 12, p333-346.

Chernozhukov, V. and L. Umantsev, "Conditional Value-at-Risk: Aspects of Modeling and Estimation," *Empirical Economics*, 26, 271-292.

Chowdhury, A.R., 1991, Futures market efficiency: evidence from cointegration tests, *The Journal of Future Markets* 11:5, pp. 577-589.

Christoffersen, P. J. Hahn, and A. Inoue, 2001, "Testing and Comparing Value-at-Risk Measures," *Journal of Empirical Finance*, 8, 325-342.

Cochrane, J.Y., 1994, Permanent and Transitory components of GNP and Stock Prices, *Quarterly Journal of Economics*, 109, 241-266.

Diba, B. and H. Grossman, 1988, "Explosive Rational Bubbles in Stock Prices," *American Economic Review*, 78, 520-530.

Donaldson, R. and M. Kamstra, 1996, "A New Dividend Forecasting Procedure That Rejects Bubbles in Asset Prices: The Case of 1929's Stock Crash," *Review of Financial Studies*, 9, 333-383.

Engle, R.F. and C.W.J. Granger (1987), "Cointegration and Error Correction: Representation, Estimation, and Testing", *Econometrica* 55, 251 - 276

Engle, R. F., and S. Manganelli, 2004, "CAViaR: Conditional autoregressive value at risk by regression quantiles," *Journal of Business and Economic Statistics, 22, 367-381.*

Enders, W. and C. Granger, 1998, Unit Root tests and asymetric adjustment with an example sing the term structure of interest rates, Journal of Business and Economic Statistics, V. 16, No. 3, 304-311.

30

Evans, G., 1991, "Pitfalls in Testing for Explosive Bubbles in Asset Prices," *American Economic Review*, 81, 4, 922-930.

Giacomini, R. and I. Komunjer, "Evaluation and Combination of Conditional Quantile Forecasts", *Journal of Business and Economic Statistics*, 23, 416-431.

Gordon, M., 1962, *The Investment, Financing and Valuation of Corporations*, Irwin, Homewood, Ill.

Granger, C.W.J. (1981), "Some Properties of Time Series Data and Their Use in Econometric Model Specification", *Journal of Econometrics*, 121 -130

Hall, A.D., H.M. Anderson and C.W.J. Granger, 1992, A cointegration analysis of treasury bill yields, *The Review of Economics and Statistics*, 74, page 116-126.

Hannan, E. J., 1970, Multiple Time Series, Wiley, New York.

Hendry, D., 1996, *Dynamic Econometrics*, Oxford University Press.

Hercé, M., 1996, Asymptotic Theory od LAD estimation in a unit root process with finite variance errors, Econometric Theory, 12, 129-153.

Hsiao, C., 1997, Cointegration and Dynamic Simultaneous Equations Models, *Econometrica*, 65, pp647-670.

Johansen, S. (1995), *Likelihood-based inference in cointegrated vector autoregressive models*, Oxford: Oxford University Press.

Hasan, M.N. and R. Koenker, 1997, Robust rank tests of the unit root hypothesis, *Econometrica* 65, No.1, 133-161.

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

Koenker, R. and G. Bassett, 1978, Regression Quantiles, *Econometrica*, V46, 33-49.

Koenker, R. and Q. Zhao, 1996, Conditional Quantile Estimation and Inference for ARCH models, *Econometric Theory*, 12, 793-812.

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

Koenker, R. and Z. Xiao, 2006, Quantile Autoregression, *Journal of the American Statistical Association*, Vol. 101, 980-990.

Lettau, M. and S. Ludvigson, 2001, Consumption, Aggregate Wealth, and Expected Stock Returns, *Journal of Finance*, 56, 815-850.

Lucas, A., 1997, "Strategic and Tactical Asset Allocation and the Effect of Long-Run Equilibrium Relations", Research Memorandum 1997-42, Vrije Universiteit Amsterdam.

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

Park, J., and S. Hahn, 1999, Cointegrating Regressions with Time Varying Coefficients, *Econometric Theory*, 15, 664-703.

Phillips, P. C. B. and B. E. Hansen (1990). "Statistical inference in instrumental variables regression with I(1) processes," *Review of Economic Studies* 57, 99–125.

Phillips, P. C. B. and S. Ouliaris (1990), "Asymptotic properties of residual-based tests for cointegration," *Econometrica,* 58, 165–193.

Phillips, P.C.B. and V. Solo, 1992, "Asymptotics for linear processes," *Annals of Statistics*, **20**, 971–1001.

Pindyck, R.S. and J.J. Rothemberg, 1992, The comovement of stock prices, *Quarterly Journal of Economics*, pp. 1073-1103.

Ploberger, W., and W. Kramer, 1992, "The CUSUM test with OLS residuals," *Econometrica*, **60**, 271–85.

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

Priestley, M.B., 1981, *Spectral Analysis and Time Series*, Academic Press.

Robinson, P. M. (1991), Automatic frequency domain inference on semiparametric and nonparametric models, *Econometrica* 59, 755-786*.*

Rockafellar, R., and S. Uryasev, 2000, Optimization of Conditional VaR, *Journal of Risk*, 2, 21-41.

Saikkonen, P., 1991, Asymptotically Efficient Estimation of Cointegration Regression, *Econometric Theory*, 7, 1-21.

Shin, Y., 1994, A residual based test of the null of cointegration against the alternative of no cointegration, *Econometric Theory*, 10, 91—115.

Siddiqui, M., 1960, Distribution of Quantiles from A Bivariate Population, Journal of Research of the National Bureau of Standards, 64B, 145-150.

Taylor, J. W., A Quantile Regression Approach to Estimating the Distribution of Multiperiod Returns, *Journal of Derivatives*, Vol. 7, 64-75, 1999.

Tyurin, K., and Z. Xiao, 2006, Risk Management of Index Fund Portfolios: A Quantile Cointegration Approach, Working Paper.

Xiao, Z., 2001, Testing the null hypothesis of stationarity against an autoregressive unit root alternative, *Journal of Time Series Analysis, 22, 87-105*.

Xiao, Z. and P.C.B. Phillips, 2002, A CUSUM test for Cointegration Using Regression Residuals, *Journal of Econometrics*, 108, 43-61.