Department of Economics

Group for Economic Analysis at Reading (GEAR)

Estimating and Testing for Functional Coefficient Quantile Cointegrating Regression

by Haiqi Li, Jing Zhang and Chaowen Zheng

Discussion Paper No. 2023-07

Department of Economics
University of Reading
Whiteknights
Reading
RG6 6AA
United Kingdom

www.reading.ac.uk

Estimating and Testing for Functional Coefficient Quantile Cointegrating Regression

Haiqi Li* Jing Zhang † Chaowen Zheng ‡

June 8, 2023

Abstract

This paper proposes a generalized quantile cointegrating regressive model for nonstationary time series, allowing coefficients to be unknown functions of informative covariates at each quantile level. Using a local polynomial quantile regressive method, we obtain the estimator for the functional coefficients at each quantile level, which is shown to be nonparametrically super-consistent. To alleviate the endogeneity of the model, this paper proposes a fully modified local polynomial quantile cointegrating regressive estimator which is shown to follow a mixed normal distribution asymptotically. We then propose two types of test statistics related to functional coefficient quantile cointegrating model. The first is to test the stability of the cointegrating vector to determine whether the conventional fixed-coefficient cointegration model is appropriate or not. The Second is to test the presence of the varying coefficient cointegrating relationship among the economic variables based on a modified quantile residual cumulative sum (MQCS) statistic. Monte Carlo simulation results show that the two tests perform quite well in finite samples. Finally, by using the proposed functional coefficient quantile cointegrating model, this paper examines the validity of the purchasing power parity (PPP) theory between China, Japan, South Korea and the United States, respectively.

Keywords: Bootstrap method; Functional coefficient quantile cointegrating model; Local polynomial approach; PPP theory;

JEL classification: C12, C13, C22

*College of Finance and Statistics, Hunan University, Changsha, 410006, China; E-mail: lihaiqi00@hnu.edu.cn.

†College of Finance and Statistics, Hunan University, Changsha, 410006, China; E-mail: jingzhang00@hnu.edu.cn.

‡Corresponding author. Department of Economics, University of Reading, Reading, RG6 6EL, UK; E-mail: chaowen.zheng@reading.ac.uk.

1 Introduction

The cointegration, which was advocated by Granger (1981) and Engle and Granger (1987), has become a popular and powerful tool to investigate the long-run equilibrium relationship among economic variables. For example, the cointegration has been used to examine the money demand function, the validity of the purchasing power parity (PPP) theory, and the price discovery between the spot and future markets for asset prices. Despite its attractive properties, less evidence in favor of cointegration is found in empirical application. Possible reasons for this failure include the conflict between the instability of the long-run relationship and the constancy of the cointegrating model coefficients, and the functional form misspecification. Alternative cointegrating models have been proposed to reconcile the conflict. For example, to allow for cointegrating coefficient to vary with different regimes, Balke et al. (1997) and Caner and Hansen (2001) proposed the threshold cointegration, and Saikkonen and Choi (2004) proposed the smooth transitional cointegration. Park and Phillips (2001) examined the general nonlinear cointegration, while Bierens (1997) proposed a consistent test for the nonparametric cointegration. Park and Hahn (1999), Bierens and Martins (2010), and Li et al. (2020) investigated the time-varying cointegation. Xiao (2009a), Cai et al. (2009), Gu and Liang (2014) and Tu and Wang (2019) studied the functional coefficient cointegration models and offered a more flexible structure of cointegration.

Recently, to investigate the long-run equilibrium relationship between the leptokurtic and heay-tailed time series, Xiao (2009b) proposed the quantile cointegrating regression which allowed the cointegrating coefficients to be quantile dependent. The quantile levels for conditional distribution of economic variables can indicate the states of an economy. Therefore, the quantile cointegrating regression enables us to study whether the long-run equilibrium relationship among economic variables varies over the economic states. As a result, quantile cointegrating regression provides a more complete view of long-run relationship among economic variable of interest. Moreover, as Xiao (2009b) pointed out, it could be regarded as a stochastic cointegration model which included the conventional counterpart as a special case. Cho et al. (2015) developed a quantile autoregressive distributed-lag model to jointly study short-run dynamics and long-run cointegrating relationships across a range of quantiles and establish the asymptotic properities. The quantile cointegrating regression has been extensively used in empirical application. For example, by using the quantile cointegration model, Lee and Zeng (2011) investigated the relationship between the spot and futures oil prices of West Texas Intermediate and Burdekin and Siklos (2012) analyzed the contagion between Chinese, US and Asia Pacific equity markets.

Athough the quantile cointegration model developed by Xiao (2009b) and Cho et al. (2015)

enables researcher to examine the quantile dependent long-run relationship among economic variables, it still assumes that the long-run relationship at each quantile level is linear and fixed over the entire sample period. As Granger and Terasvirta (1993) pointed out, however, "it was well known that relationships between major economic variables were nonlinear and that nonlinear models abounded in economic theory." To the best of our knowledge, only a few studies considered the nonlinear or nonparametric quantile cointegrating regression. [1]

To fill the gap, this paper firstly proposes a general functional coefficient quantile cointegrating regressive model. It is more flexible and includes the conventional quantile cointegating model, nonlinear cointegrating model and functional coefficient cointegrating model as special cases. To the best of our knowledge, this paper is the first to investigate the varying coefficient quantile cointegrating model, allowing the cointegrating coefficients at each quantile level to vary with stationary covariates. Suppose that n is the sample size and h is the bandwidth in the local polynomial estimation, the convergence rate for the proposed estimator for the coefficient is $n\sqrt{h}$ while its stationary counterpart converges only at the rate $\sqrt{nh}$ as shown by Cai et al. (2000). In this sense, our estimator is nonparametrically super-consistent.

Secondly, We show that our proposed estimation for the cointegration parameter is $n\sqrt{h}$ consistent which however suffers from asymptotic bias due the endogeneity issue, making the asymptotic distribution of the estimator non-standard. To correct the bias term, the "Fully Modified" procedure by Phillips and Hansen (1990) is implemented and the result estimator is shown to have a mixed normal distribution asymptotically.

Although our proposed method is robust to parameter instability and will always give consistent estimation results, it is however less efficient when the underlying true parameter are constant. Testing the parameter constantcy is therefore an important inference issue. As our third contribution, we propose test statistics that allow us to test parameters instability at a given quantile and over a range of different quantile levels. Another important inference problem specific to cointegration analysis is to test for cointegration relationship. While there are already various tests in the literature, none of them is applicable for our case due to the complex structure of our proposed model and nonparametric estimation, we therefore propose a new cointegration test, which is our final contribution. The asymptotic property of our proposed test statistics for parameter (in)stability and cointegration are also well established. Via substantial Monte Carlo simulations and comparison with other corresponding existing tests,

[1]Both Li et al. (2016) and Uematsu (2019) study the nonlinear quantile cointegration models and restrict the coefficient to be time invariant. Liang et al. (2019) explore the kernel and local linear quantile estimation for functional coefficient regression model with nonstationary covariates. Tu et al. (2021) propose a specificiation test to test the funcitonal form specification in quantile cointegration. To the best of our knowledge, there is no cointegration test for functional coefficient quantile model.

they are shown to have good finite sample performance.

We finally apply our proposed method into the analysis of the PPP theory for three major Asian countries: China, Japan and South Korea, again United States. Our finding are summarised as follows. The estimated relationships between the nominal exchange rate and price level difference calculated according to Producer Price Index (PPI) demonstrate significant asymmetric property across different economic states and time-varying property depending on the value of interest rate spread. In addition, according to the cointegration test results, we find that PPP theory is valid for explaining exchange rate between China and the United States across different economic status. By contrast, PPP theory is only applicable for explaining exchange rate between Japan and the United States at normal or prosperity economic situations, while it cannot explain that between South Korea and the United States only at the extreme recession period.

The remaining of this paper is organized as follows. Section 2 introduces the functional coefficient quantile cointegrating model and the estimating method. Section 3 establishes the asymptotic properties of the model coefficient estimators. Section 4 develops two test statistics, testing the parameters constancy/stability and the existence of long-run relationships between variables. Section 5 presents the results of Monte Carlo simulations for our proposed tests. Section 6 applies our proposed functional coefficient quantile cointegrating model and the two test statistics into examining the validity of the PPP theory. Section 7 concludes. Section 8 provides supplementary simulation results.

For notation convenience, $\xrightarrow{p}$ and $\Rightarrow$ denote convergence in probability and in distribution, respectively. The notation $\lfloor \cdot \rfloor$ denotes the integer part of a real number. $B(\cdot)$ is denoted as a standard Brownian motion.

2 Functional Coefficient Quantile Cointegrating Model

Suppose that y_t is a scalar dependent variable and x_t is a $k \times 1$ independent variable, and both y_t and x_t are $I(1)$, then the conventional linear cointegrating regressive model advocated by Engle and Granger (1987) can be specified as

$$y_t = \boldsymbol{\beta}' x_t + u_t, \tag{1}$$

where, $\boldsymbol{\beta}$ is the $k \times 1$ coefficient vector and u_t is the stochastic error. We say that y_t and x_t are cointegrated if u_t is stationary, and otherwise, the model specified above is likely to be a spurious regression.

Over the past several decades, the linear cointegrating regressive model (1) has been widely employed in analyzing the long-run relationships between non-stationary economic variables, which is assumed to be constant over the entire sample period. However, the assumption of the fixed linear relationship is only an approximation to the potentially nonlinear, asymmetric and time-varying counterpart in reality. Therefore, it might be subject to the misspecification issue and lead to misleading estimating results and conclusion. For these reasons, we propose the following general model:

$$y_t = \beta(U_t, z_t)' x_t + u_t, \tag{2}$$

where for simplicity, z_t is assumed to be a scalar stationary variable and our analysis below could be easily extended to multidimensional case with tedious notations. Moreover, U_t is a uniformly distributed random variable over $[0, 1]$, i.e., $U_t \sim U[0, 1]$. $\beta(U_t, z_t)$ is the cointegrating coefficient that is allowed to vary with both U_t and z_t. Many similar models have been investigated in the literature, for example, Cai et al. (2009) and Xiao and Phillips (2002) examined a model where the cointegration coefficient is only allowed to be vary with z_t (that is, $\beta(\cdot)$ is a unknown function of z_t), and Chen and Hong (2012) considered a time-varying coefficient model where $\beta(\cdot)$ is set as a function of scaled time t/n. Liang et al. (2019) also studied a very similar model with (2), but both y_t and x_t are assumed to be stationary in their model. Moreover, when $\beta(U_t, z_t) \equiv \beta$, then Model (2) will reduce to Model (1).

Denoting $\mathcal{F}_t = \{x_j, z_j : \forall j \leq t\}$ as the information available up to time t, and assuming that the right-hand side of (2) is monotonically increasing with U_t conditional on $\mathcal{F}_t$, we can write the τ-th conditional quantile of y_t as

$$Q_{y_t}(\tau|\mathcal{F}_t) = \beta(\tau, z_t)' x_t + F_t^{-1}(\tau). \tag{3}$$

The cointegration coefficient $\beta(\tau, z_t)$ in Model (3) is a function of both the quantile level τ and the smoothing variable z_t. More specifically, for any given quantile level τ, $\beta(\tau, \cdot)$ is a function of the smoothing variable z_t, and thus, it can characterize the nonlinear and time varying relationships between economic variables. For any given values of z_t, $\beta(\cdot, z)$ is a function of the quantile level τ, and therefore, it can capture the asymmetric relationships between economic variables.

The generality and robustness of Model (3) is attractive and it enables us to model the unknown and complicated long-run relationships between economic variables. However, it imposes a great challenge to the estimation of the coefficient. In this paper, we propose a local polynomial quantile regressive approach.[2] To be specific, for any given fixed quantile level τ, we consider a local polynomial approximation for the unknown coefficient function $\beta(\tau, z_t)$.

[2]Specifically, we use local linear quantile regressive approach. In the literature, the local linear approach is favorable owning to its nice minimax efficiency and better properties in boundary regions compared to the local constant approach.

Suppose that z_t is an arbitrary point in a small neighborhood of z and $\beta(\tau, z)$ is continuously differentiable with respect to z, then by the Taylor expansion, we can obtain

$$\beta(\tau, z_t) \approx \beta(\tau, z) + \beta^{(1)}(\tau, z)(z_t - z) + \cdots + \frac{\beta^{(q)}(\tau, z)}{q!}(z_t - z)^q, \tag{4}$$

where $\beta^{(q)}(\tau, z)$ denotes the q-th order derivative of $\beta(\tau, z)$ with respect to z.

Let $\beta_j(\tau, z) = \frac{\beta^{(j)}(\tau,z)}{j!}, j = 0, 1, 2, \cdots, q$, we can use a local polynomial approach to fit the function $\beta(\tau, z_t)$ as long as the observations in a small neighborhood of z are sufficiently rich. Substituting (4) into Model (3), the local polynomial nonparametric estimator of the coefficient in Model (3) can be obtained by solving the following minimizing problem:

$$\underset{\beta}{min} \sum_{t=1}^{n} \rho_\tau \left(y_t - \sum_{j=0}^{q} (z_t - z)^j x_t' \beta_j(\tau, z) \right) K(z_t - z), \tag{5}$$

where $\rho_\tau(u) \equiv u(\tau - I(u < 0))$ is the well-known check function, $K(z_t - z) \equiv k((z_t - z)/h)$ is a kernel function with a bounded support $[-1, 1]$and h is the bandwidth or smoothing parameter. The resulting estimator $\hat{\beta}(\tau, z)$ is the so-called local polynomial estimator of $\beta(\tau, z_t)$ for z_t near z.

To simplify the notations, define $X_{tz} = (x_t', (z_t - z)x_t', \cdots, (z_t - z)^q x_t')'$ as the new explanatory variable, which is a local polynomial vector depending on location z, and $\boldsymbol{\beta}(\tau, z) = (\beta(\tau, z)', \beta_1(\tau, z)', \cdots, \beta_q(\tau, z)')'$. In matrix notation, the minimization problem (5) could be rewritten as

$$\underset{\beta}{min} \sum_{t=1}^{n} \rho_\tau \left(y_t - X_{tz}' \boldsymbol{\beta}(\tau, z) \right) K(z_t - z), \tag{6}$$

which is similar to the the objective function of the nonparametric local quantile regression model, see Yu and Jones (1998), Spokoiny et al. (2013). Moreover, the minimizing problem (6) can be solved by using existing software packages (e.g, the `quantreg` in *R*.), which could be seen as an advantage of our proposed method from the perspective of computation.

There are two important issues in nonparametric estimation: the selection of kernel function and bandwidth. While different kernel functions could provide different estimation results since they assign different weights to the data points, the results are in general quite similar when sample size is large since kernel estimations is consistent and asymptotic normal. Without lose of generality, we therefore choose to use the Gaussian kernel in this paper.

It is the bandwidth selection that are more important in nonparametric estimation and there is usually a trade-off between estimation bias and variance: while estimators with smaller bandwidth may also have smaller bias, its variance could be very large. Commonly used bandwidth selection method in the literature including the rule of thumb, plug in and cross validation

(CV) methods. When data are normally distributed, rule of thumb is a good choice. However, it may lead to over smoothness when data are asymmetric or multimodal. Plug-in method is based on minimizing the MISE, which could be difficult to calculate in nonprametric estimation. As a data-driven method, CV has good finite sample performance and is widely used in the literature. Cai and Xu (2009) also proposes a bandwidth selection method based on Akaike information criterion, which has good in-sample performance but the out-sample performance is less satisfactory. We therefore apply the CV method for bandwidth selection and the leave one out cross validation (LOOCV) in particular. The algorithm can be described as follows:

(i). Determine the bandwidth range $[h_d, h_u]$ and choose a bandwidth $h_i \in [h_d, h_u]$.

(ii). Estimate (3) using h_i by excluding data point $\{y_j, x_j, z_j\}$ repeatedly for each $j = 1, 2 \ldots, n$, obtaining the predicted $\hat{y}_j$, and calculate the deviation series $\{\hat{u}_{j\tau} = y_j - \hat{y}_j\}_{j=1}^n$.

(iii). Calculate the average predicted bias $B(h_i) = \frac{1}{n}\sum_{j=1}^n \rho_\tau(\hat{u}_{j\tau})$.

(iv). Repeat the above three steps for each $h_i \in [h_d, h_u]$ and the optimal bandwidth is the one producing the minimal $B(h_i)$, that is, $h_{opt} = \underset{h_d \le h_i \le h_u}{min} B(h_i)$.

3 Asymptotic distribution of the estimator

In this section, we establish the asymptotic properties of our proposed estimators defined in minimization problem (5) and (6). First, we make the following assumptions.

Assumption 1. [Regularity Conditions]

(i). The kernel function $K(\cdot) : [-1, 1] \to \mathbb{R}^+$, satisfies that $\int K(u)du = 1$, $\int uK(u)du = 0$ and $|u|_j(k) = \int |u|^j K(u)du < \infty, j \le 2 + \lambda, 0 < \lambda \le 1$.

(ii). As $n \to \infty$, $0 < h \to 0, nh \to \infty$, and $nh^{4+2\lambda} \to 0$.

Assumption 2. [Smoothness Conditions]

(i). For any given quantile level τ, $\beta(\tau, z_t)$ has continous $q + 1$ order derivative at a neighborhood around z and there exists a constant C such that $\|\beta(\tau, z_t) - \beta(\tau, z) - \beta^{(1)}(\tau, z)(z_t - z)\| \le C|z_t - z|^2$.

(ii). The density function of z_t, $f_z(\cdot)$ is bounded uniformly continuous partial derivatives up to the order q.

(*iii*). For any given number c, the stochastic error u_t has continuously differentiable density $f(c) = F'(c)$, where $F(c) = Pr(u_t < c|\mathcal{F}_t)$ is the cumulative distribution function. Moreover, there exists a constant C such that $|f_t(c) - f_t(d)| \leq Cmin\{|c-d|^{\lambda}, 1\}, 0 < \lambda \leq 1$.

Assumption 3. [Technical Conditions]

(*i*). For any t, $x_t/\sqrt{n}$ is uniformally bounded.

(*ii*). Let $v_t = \Delta x_t$, then $\{u_t, v_t, z_t\}$ is a α-mixing process and the α-mixing coefficient satisfies $\Sigma_{t\geq 1}^{\infty} t^{\gamma} \alpha^{(\delta-2)/\delta}(t) < \infty$, where $\delta \geq 2, \gamma > (\delta-2)/\delta$.

Assumption 1 is standard in nonparametric kernel estimation. Assumption 2(i) is needed for implementing the q-th order Taylor expansion to the coefficient function. The rest assumptions are for establishing the asymptotic normality of the proposed estimator. Defining the piecewise derivative of the check function in the quantile regression as $\psi_\tau(u_{t\tau}) = \tau - I(u_{t\tau} < 0), u_{t\tau} = u_t - F^{-1}(\tau)$ and denoting $\underline{K}_t = K(z_t - z) - EK(z_t - z)$, we have the following functional central limit theorem.

Lemma 1. *Under Assumptions 1 and 3, the partial sums of the stochastic process* $(\underline{K}_t \psi_\tau(u_{t\tau}), v_t)$ *follow the functional central limit theorems:*

$$\begin{pmatrix} \frac{1}{\sqrt{nh}} \sum_{t=1}^{\lfloor nr \rfloor} \underline{K}_t \psi_\tau(u_{t\tau}) \\ \frac{1}{\sqrt{n}} \sum_{t=1}^{\lfloor nr \rfloor} v_t \end{pmatrix} \Rightarrow \begin{pmatrix} B_\psi^k(r) \\ B_v(r) \end{pmatrix} = BM(0, \Omega),$$

where $BM(0,\Omega)$ *represents a mixed normal distribution with zero mean and covariance matrix* Ω*, which could be further decomposed as*

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

Notice that $\omega_\psi^2 = \nu_0(K) f_z(z) \tau(1-\tau)$, where $\nu_0(K) = \int K^2(u)du$, we can clearly see that the kernel weighted sequence $\underline{K}_t \psi_\tau(u_{t\tau})$ is asymptotically uncorrelated, and its partial sum process converges to a Brownian Motion with variance proportional to $\tau(1-\tau)$, which is the short-run variance of $\psi_\tau(u_{t\tau})$. This result indicates an inherent robustness of kernel estimator if u_t is weakly dependent. Intuitively, the downward kernel smoothing reduces the serial dependence such that the kernel weighted sequence has an asymptotic variance similar to the asymptotically uncorrelated sequence. In addition, the limiting processes $B_\psi^k(r)$ and $B_v(r)$ may be correlated whenenver contemporaneous correlation between $\psi_\tau(u_{t\tau})$ and v_t exists, i.e., $\Omega_{\psi v} = \Omega_{v\psi} \neq 0$. It also ensures that the covariance matrix of $B_v(r)$, Ω_{vv} is nonsingular.

Based on Lemma 1, we then can establish the asymptotic normality of $\hat{\beta}(\tau, z)$ given in Theorem 1.

Theorem 1. *Suppose that Assumptions 1 – 3 hold. Then as $n \to \infty$,*

$$n\sqrt{h}\left[\hat{\beta}(\tau,z)-\beta(\tau,z)-\mathcal{B}\right] \Rightarrow \frac{1}{f(F^{-1}(\tau))}\left[f_z(z)\int_0^1 B_v B_v'\right]^{-1}\left[\int_0^1 B_v dB_\psi^k + \lambda_{v\psi}\right], \quad (7)$$

where $\mathcal{B} = \dfrac{h^{q+1}}{(q+1)!}\beta^{q+1}(\tau,z)\mu_{q+1}$, $\mu_j = \int u^j K(u)du$, $j \geq q+1$, and $\lambda_{v\psi}$ represents the one side long run variance-covariance matrix between v_t and $K(z_t - z)\psi_\tau(u_{t\tau})$.

Theorem 1 shows that the estimated coefficients are nonparametrically super-consistent in the sense that their convergence rate is faster than the one when data is stationary, which is $\sqrt{nh}$ (see Cai and Xu (2009)). However, due to the nonparametric estimation, the convergence rate is still slower than n, which is the convergence rate of the estimating coefficients in Xiao's (2009b) quantile linear cointegrating regressive model. Similar to Xiao (2009a), there is a bias term $\mathcal{B}$ in the asymptotic distribution of $\lambda_{v\psi}$, which comes from the approximation error of local polynomial estimation. Since we assume $h \to 0$, this bias term is in general small and would be negligible if we further assume $nh^{(2q+3)/2} \to 0$, where q is the order of the local polynomial. Moreover, regardless of the nonparametrical super-consistency, the asymptotic distribution of $\hat{\beta}(\tau,z)$ is non-standard and depends on the nuisance parameter $\lambda_{v\psi}$. Similar to Xiao (2009b), it can attribute to the endogeneity problem, that is, the correlation between $K(z_t - z)\psi_\tau(u_{t\tau})$ and v_t. When they are uncorrelated, $\lambda_{v\psi} = \Omega_{v\psi} = 0$, and thus, the asymptotic distribution of $\hat{\beta}(\tau,z)$ is mixed normal.

In the general case, to obtain a mixed normal limiting distribution, we develop a fully modified local polynomial quantile regressive estimator formulated as

$$\hat{\beta}^m(\tau,z) = \hat{\beta}(\tau,z) - \frac{1}{\hat{f}(F^{-1}(\tau))}\left[\sum_{t=1}^n x_t x_t' K(z_t - z)\right]^{-1}\left[\sqrt{h}\sum_{t=1}^n x_t v_t'\hat{\Omega}_{vv}^{-1}\hat{\Omega}_{v\psi} + n\sqrt{h}\hat{\lambda}_{v\psi}^m\right], \quad (8)$$

where $\hat{\lambda}_{v\psi}^m = \hat{\lambda}_{v\psi} - \hat{\lambda}_{vv}\hat{\Omega}_{vv}^{-1}\hat{\Omega}_{v\psi}$. The terms $\hat{\lambda}_{v\psi}$, $\hat{\lambda}_{vv}$, $\hat{\Omega}_{vv}$ and $\hat{\Omega}_{v\psi}$ are consistent estimators for $\lambda_{v\psi}$, λ_{vv}, Ω_{vv}, $\Omega_{v\psi}$, respectively, which could be obtained by the following kernel estimation:

$$\begin{aligned}\hat{\lambda}_{v\psi} &= \sum_{j=0}^M k\left(\frac{j}{M}\right)\Gamma_{v\psi}(j), \qquad \hat{\lambda}_{vv} = \sum_{j=0}^M k\left(\frac{j}{M}\right)\Gamma_{vv}(j),\\ \hat{\Omega}_{v\psi} &= \sum_{j=-M}^M k\left(\frac{j}{M}\right)\Gamma_{v\psi}(j), \quad \hat{\Omega}_{vv} = \sum_{j=0}^M k\left(\frac{j}{M}\right)\Gamma_{vv}(j), \end{aligned} \quad (9)$$

where $k(\cdot)$ is the kernel function with bounded support $[-1,1]$, and M is the bandwidth parameter satisfying the conditions that $M \to \infty$ and $M/n \to 0$ as $n \to \infty$. Moreover, $\Gamma_{v\psi}(j) = \frac{1}{n}\sum_{t=1}^{n-j} K(z_t - z)v_t\psi_\tau(\hat{u}_{t+j,\tau})$ and $\Gamma_{vv}(j) = \frac{1}{n}\sum_{t=1}^{n-j} v_t v_{t+j}$ are the sample covariances. $\hat{f}(F^{-1}(\tau))$ is consistent estimators of the density function $f(F^{-1}(\tau))$.

Theorem 2 below gives the asymptotic distribution of $\hat{\beta}^m(\tau,z)$.

Theorem 2. *Suppose that Assumptions 1 – 3 hold. Then as $n \to \infty$,*

$$\begin{aligned} n\sqrt{h}\left[\hat{\beta}^m(\tau,z)-\beta(\tau,z)-\mathcal{B}\right] \Rightarrow & \frac{1}{f(F^{-1}(\tau))}\left[f_z(z)\int_0^1 B_vB_v'\right]^{-1}\left[\int_0^1 B_v dB^k_{\psi\cdot v}\right] \\ \Rightarrow & MN\left(0, \frac{\omega^2_{\psi\cdot v}}{f(F^{-1}(\tau))^2}\left[f_z(z)\int_0^1 B_vB_v'\right]^{-1}\right). \end{aligned} \tag{10}$$

where $B^k_{\psi\cdot v}(r) = B^k_\psi(r) - \Omega_{\psi v}\Omega^{-1}_{vv}B_v$.

Note that $B^k_{\psi\cdot v}(r)$ is independent of B_v and has variance $\omega^2_{\psi\cdot v} = \omega^2_\psi - \Omega_{\psi v}\Omega^{-1}_{vv}\Omega_{\psi v}$. Therefore, Theorem 2 shows that $\hat{\beta}^m(\tau,z)$ has a mixed normal limiting distribution by introducing nonparametric corrrection terms. The fully modified local polynomial quantile regression estimator $\hat{\beta}^m(\tau,z)$ generalizes the traditional fully modified regression estimator of Phillips and Hansen (1990) and the fully modified quantile regression estimator of Xiao (2009b).

4 Inference

In this section, we propose two novel test statistics. The first is to test for stability of the cointegrating vector, and the second is to test for cointegration between economic variable. The parameter stability test is important since the traditional fixed-coefficient model could be more efficient under the null hypothesis of stability but inconsistent under the alternative. Moreover, the cointegration test is also of key importance as when cointegration fails y and x, any estimated relationships (either fixed or time-varying) are likely to be spurious.

4.1 Testing for stability of the cointegrating vector

We first investigate the tests for parameter constancy at a given quantile, that is, $H_{01} : \beta(\tau_0, z) = \beta(\tau_0)$ for a given $\tau_0 \in (0,1)$. The alternative hypothesis is that the null hypothesis H_{01} does not hold.

Under the null hypothesis H_{01}, $\beta(\tau_0)$ is a unknown fixed parameter vector at a given quantile τ_0. Xiao (2009b) has developed a super-consistent estimator of $\beta(\tau_0)$ and shows that

$$n[\hat{\beta}(\tau_0) - \beta(\tau_0)] \Rightarrow \frac{1}{f(F^{-1}(\tau_0))^2}\left[\int_0^1 B_vB_v^{'}\right]^{-1}\int_0^1 B_v dB_{\psi\cdot v}.$$

To test H_{01}, one can directly compare the difference between estimated $\hat{\beta}(\tau_0, z)$ and $\hat{\beta}(\tau_0)$ over a range of different values for z. Moreover, when H_{01} is true and by the result of Theorem 2, for any given value of $z = z^*$, we have

$$\begin{aligned} n\sqrt{h}[\hat{\beta}^m(\tau_0, z^*) - \hat{\beta}(\tau_0)] &= n\sqrt{h}[\hat{\beta}^m(\tau_0, z^*) - \beta(\tau_0)] - \sqrt{h} \cdot n[\hat{\beta}(\tau_0, z^*) - \beta(\tau_0)] \\ &= n\sqrt{h}[\hat{\beta}^m(\tau_0, z^*) - \beta(\tau_0)] + o_p(\sqrt{h}) \\ &\Rightarrow MN(0, \Omega(\tau_0, z^*)), \end{aligned}$$

where $\Omega(\tau_0, z^*) = \dfrac{\sigma^2_{\psi\cdot v}}{f(F^{-1}(\tau_0))^2}\left[f_z(z^*)\int_0^1 B_v B_v'\right]^{-1}$.

To further test the null hypothesis H_{01} over a range of different values of z, we treat $\hat{\beta}^m(\tau_0, z)$ as a function of z and propose the following Kolmogorov-Smirnov type test

$$SA \equiv sup_{z\in[\underline{z},\bar{z}]}|n\sqrt{h}[\hat{\beta}^m(\tau_0, z) - \hat{\beta}(\tau_0)]| \equiv sup_{z\in[\underline{z},\bar{z}]}|\hat{V}(\tau_0, z)|. \tag{11}$$

Note that each $\hat{V}(\tau_0, z)$ follows a mixed normal distribution asymptotically and the correlation between $\hat{V}(\tau_0, z_t)$ and $\hat{V}(\tau_0, z_s)$ is zero, and for any $t \neq s$,

$$E\left[\frac{1}{h}K\left(\frac{z_t - z}{h}\right)K\left(\frac{z_s - z}{h}\right)\right] = O(h) \xrightarrow{p} 0,$$

Theorem 3 establishes the asymptotic distribution of SA.

Theorem 3. *Suppose that Assumptions 1-3 hold. Then under the null hypothesis H_{01}, we have,*

$$SA \Rightarrow sup_{z\in[\underline{z},\bar{z}]}|B(\tau_0, z)|, \quad as\ n \to \infty,$$

where $B(\tau_0, z)$ is a vector of independent Brownian bridge processes on $[\underline{z}, \bar{z}]$.

In practice, one can tabulate the critical values for the limiting distribution of SA. Any significantly large values of SA are evidences in favor of the alternative hypothesis.

By SA, one can check the stability of the coefficient at a given quantile. However, one may also be interested in the stability of the coefficient across different quantiles irrespective of the values for z, that is, one may be interested to test H_{02}: $\beta(\tau, z) = \beta$ for all values of τ and z. To this aim, we extend the SA test defined in (11) to a double supremum test. By the results of Theorem 3 and when H_{02} is true, we can obtain

$$\begin{aligned} SM =& sup_{\tau\in\mathcal{T}} sup_{z\in[\underline{z},\bar{z}]}|\hat{V}(\tau, z)| \\ \Rightarrow& sup_{\tau\in\mathcal{T}} sup_{z\in[\underline{z},\bar{z}]}|B(\tau, z)|, \end{aligned} \tag{12}$$

where $\mathcal{T}$ is a closed interval of quantiles, $B(\tau, z)$ is a k-vector of independent Gaussian processes which is often referred to as the Brownian Pillow, and for each pair of fixed (τ^*, z^*), $B(\tau^*, z^*) \sim MN(0, \Omega(\tau^*, z^*))$.

Although the limiting distribution of the proposed test statistics is free of nuisance parameter, the calculation of its critical value could be imprecise especially when sample size is small due to the complicated form of the variance. As an alternative, we use the bootstrap method to obtain the empirical critical values for the proposed tests. Similar to Xiao (2009b), the bootstrap algorithm can be described as follows:

Algorithm 1

(i). Obtain $\hat{\beta}^m(\tau_0, z)$ via our proposed estimating procedure over a range of interested values for $z \in [\underline{z}, \bar{z}]$, and the estimator $\hat{\beta}^m(\tau_0)$ in Xiao (2009b). The test statistic SA can obtained as

$$SA = sup_{z\in[\underline{z},\bar{z}]}|\hat{V}(\tau_0, z)| = sup_{z\in[\underline{z},\bar{z}]}|n\sqrt{h}[\hat{\beta}^m(z, \tau_0) - \hat{\beta}(\tau_0)]|.$$

(ii). Fit a vector autoregressive (VAR) model to the series $H_t = \{v_t, \hat{u}_{t\tau}\}$, where $v_t = \Delta x_t$ and $\hat{u}_{t\tau} = y_t - \hat{\beta}^m(\tau_0)x_t$,

$$H_t = \sum_{i=1}^{p} A_i H_{t-i} + e_t.$$

The lag order p is usually determined by AIC. By fitting the above VAR model, one can get the residual series $\hat{e}_t = H_t - \sum_{i=1}^{p} \hat{A}_i H_{t-i}$.

(iii). Obtain the bootstrapped residual $\{\hat{e}_t^*\}$ by resampling the centered series $\{\hat{e}_t - \sum_{t=1}^{n} \hat{e}_t/(n-p)\}$, and then, get the H_t^* by $H_t^* = \sum_{i=1}^{p} \hat{A}_i H_{t-1}^* + \hat{e}_t^*$, where $H_t^* = \{v_t^*, \hat{u}_{t\tau}^*\}$.

(iv). Obtain $\{y_t^*, x_t^*\}_{t=1}^n$ by $x_t^* = x_{t-1}^* + v_t^*$ with $x_1^* = x_1$, and $y_t^* = \hat{\beta}(\tau_0)x_t^* + \hat{u}_{t\tau}^*$ under the H_{01}.

(v). Obtain the bootstrapped estimator $\hat{\beta}(\tau_0)^*$ and $\hat{\beta}^m(\tau_0, z)^*$ by using the bootstrapped sample $\{y_t^*, x_t^*\}_{t=1}^n$, and get the bootstrapped test statistic

$$SA^* = sup_{z\in[\underline{z},\bar{z}]}|n\sqrt{h}[\hat{\beta}^m(\tau_0, z)^* - \hat{\beta}(\tau_0)^*]|.$$

(vi). Repeat steps $(iii) - (v)$ B times to obtain the empirical distribution for SA^*.

For any $\alpha \in [0, 1]$, let $c(\alpha)$ denote the corresponding $1-\alpha$ bootstrapped critical value of the empirical distribution of SA, then the null hypothesis H_{01} can be rejected at the α significance level if $SA \leq c(\alpha)$.

The above bootstrapped algorithm could be easily modified to obtain the bootstrapped critical values for the SM test defined in (12). In Step (i), we estimate the model (1) by OLS regression and our proposed estimator over a range of values for quantile levels and z. In Step (ii), replace $\hat{u}_{t\tau}$ by the estimated residuals $\hat{u}_t = y_t - \hat{\beta}x_t$ and generate y_t^* in Step (iii) but under H_{02} via $y_t^* = \hat{\beta}x_t^* + \hat{u}_t^*$, where $\hat{\beta}$ is the OLS estimator for the cointegration parameter.

4.2 Test for cointegration

In this section, we propose a new test for functional coefficient cointegration between y_t and x_t at the τ-th quantile. Similar to Xiao (2009b), we examine the fluctuation of the residual and consider the following partial sum process

$$Y_n(r) = \sqrt{\frac{h}{n}} \sum_{j=1}^{\lfloor nr \rfloor} \psi_\tau(\hat{u}_{t\tau}^m), \tag{13}$$

where $\psi_\tau(u) = \tau - I(u < 0), \hat{u}_{t\tau}^m = y_t - x_t' \hat{\beta}^m(\tau, z_t)$, and it is easy to show that $E[\psi_\tau(\hat{u}_{t\tau}^m)] = 0$. As argued in Xiao (2009b), the above process $Y_n(r)$ would follow an invariance principle and converges weakly to a standard Brownian motion when there exists cointegration relationship between y_t and x_t. Compared to Xiao (2009b), $Y_n(r)$ defined in Equation (13) has an additional scale term $\sqrt{h}$ which accounts for nonparametric estimation effect. Similar to Xiao (2009b), we propose a robust *Kolmogorov-Smirnov* type test for cointegration as follows

$$MQCS(\tau) = \underset{j=1,\cdots,n}{max} \frac{\sqrt{h}}{\hat{\omega}_{\psi\cdot v}\sqrt{n}} \left| \sum_{t=1}^{j} \psi_\tau(\hat{u}_{t\tau}^m) \right|. \tag{14}$$

Theorem 4 provides the limiting distribution of $MQCS(\tau)$.

Theorem 4. *Let $\psi_\tau(\hat{u}_{t\tau}^m) = \tau - \mathrm{I}(y_t \le x_t'\hat{\beta}^m(\tau, z_t))$. Then, under Assumptions 1-3 and irrespective of parameter (in)stability, if there exists a long-run relationship (cointegration) between y_t and x_t as specified in model (3) at τ-th quantile, we have*

$$MQCS(\tau) = \underset{j=1,\cdots,n}{max} \frac{\sqrt{h}}{\hat{\omega}_{\psi\cdot v}\sqrt{n}} \left| \sum_{t=1}^{j} \psi_\tau(\hat{u}_{t\tau}^m) \right| \Rightarrow \underset{0\le r\le 1}{sup} |\widetilde{W}(r)|, \tag{15}$$

where $\widetilde{W}(r) = -(\int_0^r W_v(s)ds)(\int_0^1 W_v(s)W_v'(s)ds)^{-1}(\int_0^r W_v(s)dW_{\psi\cdot v}^k(s))$, $W_{\psi\cdot v}^k(s)$ and $W_v(s)$ are standard Brownian motion independent of each other.

Theorem 4 shows that, the limiting distribution of $MQCS(\tau)$ is free of nuisance parameter. Note that if we obtain the test statistic $MQCS(\tau)$ by using $\hat{\beta}(\tau, z_t)$ rather than the fully modified estimator function coefficient estimator $\hat{\beta}^m(\tau, z_t)$, its limiting distribution would depend on the nuisance parameter. This is because the limiting distribution of $\hat{\beta}(\tau, z_t)$ is not mixed normal.

It is well known that the residual-based cointegration test with stationary null hypothesis could suffer from size distortion (e.g., Shin (1994), Xiao and Phillips (2002)). To improve the finite sample performance, we follow Phillips (2010) and Gu and Liang (2014) and propose a bootstrap procedure for calculating the p-value of the test statistics as follows:

Algorithm 2

(i). Estimate Model (1) at the desired quantile level τ_0 by our proposed procedure and save the residual series $\hat{u}_{t\tau} = y_t - x_t'\hat{\beta}^m(\tau_0, z_t)$. Calculate our proposed test statistics $MQCS(\tau_0)$.

(ii). Construct bootstrapped $\{x_t^*\}$ by using the continuous moving block bootstrap: define the block size as b and let $n_1, \ldots, n_m$ be drawn independently and uniformly from $\{0, 1, \ldots, n-b\}$ with $m = [n/b]$. For $k = 1, \ldots, b.$, generate $x_{sb+k}^* = x_k^* = x_1 + (x_{n_1+k} - x_{n_1})$ when $s = 0$, and $x_{sb+k}^* = x_{sb}^* + x_{n_{s+1}+k} - x_{n_{s+1}}$ when $s = 1, \ldots, m-1$.

(iii). Generate the bootstrap error by $e_t^* = v_t\hat{u}_{t\tau}$ where $\{v_t\}_{t=1}^n$ is a stationary process independent of all the other variables, with $E(v_t) = 0, Var(v_t) = 1$ for $t = 1, \ldots, n, Cov(v_t, v_s) = a((t-s)/l)$, for $t \neq s$, where $a(x) = 1 - |x|$ for $|x| \leq 1$ and $a(x) = 0$ otherwise, where $l = I_n$ is a bandwidth parameter, and $1/l + l/n = o(1)$, as $n \to \infty$. Let $y_t^* = x_t'\hat{\beta}^m(\tau_0, z_t) + e_t^*$

(iv). Calculate the bootstrapped test statistics $MQCS^*(\tau)$ based on the sample $\{y_t^*, z_t, x_t^*\}_{t=1}^n$.

(v). Repeat steps (iii) to (iv) B times to obtain the empirical distribution for $MQCS(\tau)$.

Based on the above procedure, we calculate the p-value of the test statistics, $P^* = 1 - G^*(MQCS(\tau))$ with $G^*(\cdot)$ denoting the empirical cumulative distribution of the test statistic obtained by the above bootstrap algorithm. The null hypothesis of functional cointegration will be rejected at the α significance level if $P^* \leq \alpha$.

5 Monte Carlo simulation

In this section, Monte Carlo experiments are conducted to investigate the finite sample performance of the inference procedures proposed in Section 4.

5.1 Test for Parameter Stability

We first examine the finite properties of test for parameter stability defined in (11) and (12). The Data Generating Process (DGP) is as follows:

$$
\begin{aligned}
DGP1: \quad y_t &= \beta(u_t, z_t)x_t + u_t, \\
x_t &= x_{t-1} + v_t, \quad t = 1, 2, \cdots, n, \\
\begin{pmatrix} u_t \\ v_t \end{pmatrix} &\sim F\left(\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 & \rho \\ \rho & 1 \end{pmatrix}\right).
\end{aligned}
$$

where z_t is generated from $IIDU[0,1]$, and u_t and v_t are generated from standard normal or t distributions with different degrees of freedoms (2 and 4). Unlike Xiao (2009b), we allow u_t

and v_t to be correlated and the correlation coefficient, ρ is set as 0.5 and 0.8, for moderate and strong degree of the endogeneity, respectively. More importantly, we consider the following three cases for the cointegration vector $\beta(u_t, z_t)$:

S1. $\beta(u_t, z_t) = 1$.

P1. $\beta(u_t, z_t) = \begin{cases} e^{z_t} - 1, & u_t < 0, \\ 1, & u_t \geq 0. \end{cases}$

P2. $\beta(u_t, z_t) = [0.5 + \Phi(u_t)] cos\left(\sqrt{2}\pi z_t\right)$.

In case S1, the cointegration coefficient is a constant over all quantiles and irrelevant to other variables, thus it is for investigating the empirical size of both the SA and SM tests. In case P1, the cointegration vector is a constant when the residual is positive but changing with z_t when the residual is negative. Thus it could be used to investigate the power performances of the SM test. Finally for case P2, $\beta(u_t, z_t)$ is affected by informative covariates at each given quantile, thus we use it to investigate the power performances of SA test. For all the above mentioned three cases, the bootstrap procedure introduced in Section 4.1 is conducted to obtain the critical values and the warp-speed method by Giacomini et al. (2013) is employed to speed up the computation.[3]

The proposed data-driven bandwidth selection rule in Section 2 is employed when estimating our proposed estimator defined in (6), and for simplicity, we use nonparametric local linear estimator (i.e., $p = 1$). However, to obtain the modified estimator defined in (8), we also need to estimate various variance-covariance terms as in (9), which involves the selection of the bandwidth, M. Since bandwidth selection plays a key role, we examine the effects of various bandwidths: $M_1 = 1, M_2 = [\tau(1-\tau)/\phi(\Phi^{-1}(\tau))]n^{1/3}, M_3 = 4(n/100)^{1/4}$, where M_1 is a fixed constant, M_2 is of order $n^{1/3}$ that also depends on the quanitle level,[4] and M_3 is a commonly used bandwidth in the literature, e.g. Kwiatkowski et al. (1992), Xiao and Phillips (2002). Without losing generality, we use *Bartlett* kernel function to ensure positive semi-definite of the long-run variance-covariance matrix for all sample sizes.

Moreover, the modified estimator defined in (8) also depends on the sparsity function $f(F^{-1}(\tau))$. Following Koenker and Xiao (2006), we estimate it by

$$\hat{f}(F^{-1}(\tau)) = \frac{2h_n}{\hat{F}_n^{-1}(\tau + h_n) - \hat{F}_n^{-1}(\tau - h_n)},$$

[3]While in standard bootstrap method, a large number, B bootstrap resamples are usually drawn for any given sample size, say K, in the wrap-speed bootstrap, we draw $B = 1$ bootstrap resample and compute the statistics. A sequence of K points are then inverted to calculate the bootstrap critical values.

[4]As shown in Andrews (1991), $O(n^{1/3})$ is the optimal bandwidth for *Bartlett* kernel under the correlated and heteroscedastic error assumption.

where $\hat{F}_n^{-1}(\tau)$ is the empirical distribution of the error term u_t, and the bandwidth h_n is set according the rule proposed in Hall and Sheather (1988). The nominal significance level is set as 5% and we replicate the simulations 1000 times for sample size $n = 200$ and 500, respectively.

Table 1 reports simulation results on the size and power of SM test with coefficient generated according to $S1$ and $P1$, respectively. When the errors are weakly correlated($\rho = 0.5$), we can see that SM test has empirical size very close to the nominal 5% level, and the power also increases quickly and approached 1 with sample size. As ρ becomes larger which leads to stronger serial correlation and endogeneity issue, there is slight size distortion and power loss especially in the case of infinite variance ($t(2)$ distribution) when $n = 200$. They are, however, both much improved as sample size increases. We could also see that our proposed SM test is robust to various settings of bandwidths and error distributions.

Table 1: Finite-sample performance for SM test

		$N(0,1)$		$t(2)$		$t(4)$	
		Size	Power	Size	Power	Size	Power
$n = 200$							
$\rho = 0.5$	M_1	0.040	0.422	0.042	0.262	0.036	0.420
	M_2	0.048	0.282	0.046	0.108	0.048	0.420
	M_3	0.062	0.460	0.050	0.200	0.036	0.410
$\rho = 0.8$	M_1	0.070	0.468	0.026	0.008	0.026	0.198
	M_2	0.036	0.336	0.032	0.048	0.028	0.212
	M_3	0.036	0.416	0.030	0.056	0.028	0.206
$n = 500$							
$\rho = 0.5$	M_1	0.052	0.820	0.052	0.663	0.047	0.873
	M_2	0.050	0.819	0.053	0.666	0.051	0.874
	M_3	0.052	0.819	0.054	0.666	0.050	0.874
$\rho = 0.8$	M_1	0.062	0.813	0.029	0.647	0.049	0.832
	M_2	0.048	0.811	0.032	0.640	0.051	0.832
	M_3	0.060	0.808	0.032	0.643	0.051	0.831

Note: The size of the test is calculated using the data with the true coefficient specified in S1, and the power of the test uses the data with the true coefficient specified in P1. The nominal significance level is set as 5%.

Table 2: Finite-sample performance for SA test.

		$N(0,1)$		$t(2)$		$t(4)$	
		Size	Power	Size	Power	Size	Power
$n = 200$							
$\rho = 0.5$	$\tau = 0.25$	0.062	0.602	0.100	0.795	0.060	0.756
	$\tau = 0.50$	0.052	0.926	0.072	0.754	0.070	0.780
	$\tau = 0.75$	0.062	0.752	0.098	0.762	0.052	0.794
$\rho = 0.8$	$\tau = 0.25$	0.066	0.490	0.102	0.594	0.064	0.730
	$\tau = 0.50$	0.082	0.856	0.108	0.870	0.082	0.924
	$\tau = 0.75$	0.058	0.634	0.102	0.722	0.074	0.774
$n = 500$							
$\rho = 0.5$	$\tau = 0.25$	0.074	0.946	0.088	0.906	0.056	0.946
	$\tau = 0.50$	0.052	1.000	0.052	1.000	0.048	1.000
	$\tau = 0.75$	0.060	0.962	0.086	0.974	0.056	0.992
$\rho = 0.8$	$\tau = 0.25$	0.070	0.950	0.112	0.944	0.066	0.968
	$\tau = 0.50$	0.052	1.000	0.074	1.000	0.110	1.000
	$\tau = 0.75$	0.064	0.972	0.074	0.952	0.048	0.988

Note: The size of the test is calculated using the data with the true coefficient specified in S1, and the power of the test uses the data with the true coefficient specified in P2. The nominal significance level is set as 5%.

Table 2 reports the finite sample performance of SA test at 0.25, 0.50 and 0.75 quantile levels. For all three cases, we only report results for $M = M_2$ since the results are insensitive to bandwidth selection similar to those reported in Table 1. We see that the empirical size and power results are better at 0.5 quantile level than that at tail quantiles. This is not surprising due to insufficient data in the tail of the distribution making the estimation for the sparsity function less precise. Moreover, the empirical size is slightly inflated when the innovations are generated from $t(2)$ distribution when $n = 200$. As sample size increases, the proposed test statistic has both good size and power performance, corroborating the asymptotic theory.

5.2 Test for cointegration

In this section, we examine the finite sample performance of the proposed $MQCS$ test defined in (15).

5.2.1 Test at Median Quantile

While our proposed test could be applied for testing of cointegration over various quantile levels, we first focus on median ($\tau = 0.5$) and compare its performance with many existing mean cointegration tests, including the $\mathcal{T}_n$ statistic by Xiao (2009a) and *CUSUM* test by Xiao and Phillips (2002). The *SupY* test by Xiao (2009b) for testing cointegration in constant-coefficient quantile regression is also conducted for comparison.

Consider the following DGP:

$$\begin{aligned} DGP2: \quad y_t &= \beta(u_t, z_t)x_t + u_t, \quad t = 1, \cdots, n, \\ x_t &= x_{t-1} + v_t, \\ u_t &= \gamma u_{t-1} + \epsilon_t, \\ \begin{pmatrix} \epsilon_t \\ v_t \end{pmatrix} &\sim F\left(\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 & \rho \\ \rho & 1 \end{pmatrix} \right). \end{aligned}$$

where the variables z_t, ϵ_t and v_t are generated from the same stochastic process as in *DGP1*. Notice that we now have an additional process for u_t, and the parameter γ determines whether the cointegration relationship between variables exists or not: when $\gamma < 1$, u_t is stationary, and y_t and x_x are said to be cointegrated; when $\gamma = 1$, u_t is nonstationary, and the relationship between y_t and x_x is then spurious. In our simulations, we consider $\gamma = 0, 0.2, 0.4, 0.6, 0.8$ and 1. While $\gamma = 1$ is to investigate the power of the test, the rest settings are for the size performance. Regarding other parameters, we consider the following three specifications for coefficient function:

1). $\beta(u_t, z_t) = 1$;

2). $\beta(u_t, z_t) = sin\left(\sqrt{2}\pi z_t\right)$;

3). $\beta(u_t, z_t) = [0.5 + \Phi(u_t)]\, cos\left(\sqrt{2}\pi z_t\right)$;

where the coefficient is allowed to be time-varying in the last two cases, and also quantile dependent in the third case.

To save space, we only report the size and power performance for bandwidth $M = M_2$ for $\rho = 0.5$ in Table 3 and $\rho = 0.8$ in Table 4 with $n = 200$. Results for other specifications are similar and presented in Section 8.[5]

Since the results for $\rho = 0.5$ and $\rho = 0.8$ are quite similar, we focus on analysing results in Table 4 and summarize the results as follows. First, as expected, the size of all four test

[5] Additional simulation results for $p = 2$ when fitting nonparametric regression are also provided.

Table 3: Empirical size and power ($n = 200, \rho = 0.5, M = M_2$)

	MQCS				*SupY*				$\mathcal{T}_n$				*CUSUM*			
	$N(0,1)$	$t(2)$	$t(4)$	$N(0,\sigma_t)$	$N(0,1)$	$t(2)$	$t(4)$	$N(0,\sigma_t)$	$N(0,1)$	$t(2)$	$t(4)$	$N(0,\sigma_t)$	$N(0,1)$	$t(2)$	$t(4)$	$N(0,\sigma_t)$
$\beta(u_t, z_t) = 1$																
$\gamma = 0.0$	0.045	0.077	0.012	0.054	0.035	0.045	0.036	0.031	0.005	0.005	0.023	0.012	0.031	0.033	0.037	0.041
$\gamma = 0.2$	0.063	0.102	0.018	0.051	0.037	0.036	0.037	0.038	0.006	0.004	0.019	0.014	0.030	0.026	0.039	0.027
$\gamma = 0.4$	0.060	0.138	0.062	0.051	0.055	0.043	0.024	0.038	0.011	0.012	0.021	0.014	0.050	0.043	0.042	0.031
$\gamma = 0.6$	0.097	0.132	0.041	0.089	0.049	0.070	0.071	0.056	0.012	0.006	0.039	0.016	0.044	0.045	0.063	0.045
$\gamma = 0.8$	0.096	0.145	0.060	0.094	0.075	0.065	0.084	0.076	0.038	0.034	0.082	0.041	0.099	0.101	0.105	0.093
$\gamma = 1.0$	0.780	0.828	0.926	0.808	0.070	0.078	0.068	0.086	0.342	0.370	0.396	0.380	0.754	0.680	0.736	0.746
$\beta(u_t, z_t) = sin\left(\sqrt{2}\pi z_t\right)$																
$\gamma = 0.0$	0.036	0.044	0.082	0.031	0.045	0.055	0.0042	0.053	0.007	0.008	0.029	0.015	0.030	0.028	0.027	0.030
$\gamma = 0.2$	0.062	0.040	0.045	0.047	0.055	0.056	0.041	0.044	0.008	0.010	0.021	0.012	0.040	0.030	0.033	0.035
$\gamma = 0.4$	0.042	0.056	0.101	0.045	0.059	0.034	0.032	0.040	0.010	0.011	0.031	0.020	0.027	0.038	0.037	0.041
$\gamma = 0.6$	0.073	0.065	0.085	0.088	0.050	0.043	0.048	0.047	0.015	0.014	0.036	0.017	0.049	0.050	0.052	0.043
$\gamma = 0.8$	0.171	0.162	0.178	0.179	0.049	0.055	0.053	0.051	0.040	0.031	0.095	0.041	0.115	0.097	0.093	0.121
$\gamma = 1.0$	0.838	0.794	0.826	0.854	0.064	0.070	0.066	0.084	0.410	0.388	0.404	0.370	0.760	0.680	0.746	0.792
$\beta(u_t, z_t) = [0.5 + \Phi(u_t)]cos\left(\sqrt{2}\pi z_t\right)$																
$\gamma = 0.0$	0.0335	0.030	0.021	0.037	0.055	0.047	0.044	0.068	0.351	0.343	0.387	0.347	0.122	0.411	0.169	0.129
$\gamma = 0.2$	0.036	0.053	0.039	0.028	0.058	0.053	0.045	0.069	0.307	0.343	0.388	0.344	0.154	0.437	0.180	0.150
$\gamma = 0.4$	0.049	0.065	0.049	0.036	0.065	0.063	0.043	0.056	0.307	0.321	0.376	0.334	0.168	0.427	0.219	0.155
$\gamma = 0.6$	0.059	0.078	0.079	0.060	0.057	0.052	0.034	0.063	0.303	0.326	0.353	0.337	0.218	0.384	0.234	0.227
$\gamma = 0.8$	0.094	0.105	0.168	0.115	0.052	0.076	0.064	0.052	0.266	0.275	0.335	0.271	0.362	0.472	0.365	0.339
$\gamma = 1.0$	0.786	0.598	0.688	0.642	0.064	0.084	0.094	0.078	0.474	0.534	0.536	0.464	0.872	0.886	0.884	0.902

Note: The *MQCS* is our proposed test, *SupY*, $\mathcal{T}_n$ and *CUSUM* are the tests proposed by Xiao (2009b), Xiao (2009a) and Xiao and Phillips (2002), respectively. $M_2 = [\tau(1-\tau)/\phi(\Phi^{-1}(\tau))]n^{1/3}$.

Table 4: Empirical size and power ($n = 200, \rho = 0.8, M = M_2$)

	MQCS				*SupY*				$\mathcal{T}_n$				*CUSUM*			
	$N(0,1)$	$t(2)$	$t(4)$	$N(0,\sigma_t)$	$N(0,1)$	$t(2)$	$t(4)$	$N(0,\sigma_t)$	$N(0,1)$	$t(2)$	$t(4)$	$N(0,\sigma_t)$	$N(0,1)$	$t(2)$	$t(4)$	$N(0,\sigma_t)$
$\beta(u_t, z_t) = 1$																
$\gamma = 0.0$	0.056	0.015	0.075	0.070	0.051	0.034	0.039	0.032	0.014	0.021	0.009	0.013	0.026	0.031	0.033	0.029
$\gamma = 0.2$	0.064	0.038	0.083	0.113	0.040	0.021	0.036	0.038	0.023	0.016	0.019	0.012	0.030	0.042	0.029	0.030
$\gamma = 0.4$	0.061	0.041	0.108	0.101	0.059	0.029	0.035	0.062	0.011	0.025	0.012	0.009	0.029	0.049	0.042	0.016
$\gamma = 0.6$	0.086	0.045	0.106	0.118	0.056	0.080	0.052	0.066	0.020	0.033	0.052	0.020	0.034	0.051	0.044	0.038
$\gamma = 0.8$	0.097	0.049	0.105	0.133	0.086	0.085	0.098	0.089	0.041	0.071	0.052	0.033	0.051	0.108	0.090	0.062
$\gamma = 1.0$	0.758	0.836	0.794	0.746	0.078	0.050	0.082	0.068	0.402	0.438	0.412	0.458	0.570	0.692	0.716	0.464
$\beta(u_t, z_t) = sin\left(\sqrt{2}\pi z_t\right)$																
$\gamma = 0.0$	0.050	0.074	0.051	0.043	0.052	0.034	0.050	0.060	0.032	0.023	0.014	0.033	0.029	0.033	0.027	0.028
$\gamma = 0.2$	0.052	0.071	0.061	0.054	0.056	0.035	0.039	0.056	0.024	0.022	0.010	0.027	0.034	0.038	0.032	0.032
$\gamma = 0.4$	0.063	0.076	0.057	0.066	0.051	0.037	0.028	0.025	0.027	0.028	0.016	0.019	0.026	00.28	0.038	0.027
$\gamma = 0.6$	0.080	0.099	0.081	0.069	0.046	0.042	0.040	0.039	0.021	0.046	0.014	0.022	0.037	0.042	0.050	0.047
$\gamma = 0.8$	0.176	0.200	0.162	0.168	0.043	0.031	0.037	0.038	0.038	0.073	0.041	0.036	0.049	0.100	0.102	0.059
$\gamma = 1.0$	0.802	0.840	0.840	0.772	0.044	0.098	0.056	0.044	0.396	0.416	0.410	0.448	0.642	0.682	0.706	0.530
$\beta(u_t, z_t) = [0.5 + \Phi(u_t)]cos\left(\sqrt{2}\pi z_t\right)$																
$\gamma = 0.0$	0.031	0.030	0.034	0.043	0.040	0.041	0.061	0.063	0.439	0.413	0.357	0.408	0.099	0.471	0.428	0.318
$\gamma = 0.2$	0.031	0.031	0.045	0.054	0.040	0.048	0.051	0.041	0.443	0.390	0.340	0.438	0.143	0.468	0.437	0.366
$\gamma = 0.4$	0.033	0.072	0.047	0.067	0.057	0.058	0.039	0.044	0.450	0.376	0.355	0.452	0.167	0.438	0.422	0.365
$\gamma = 0.6$	0.052	0.067	0.059	0.062	0.042	0.053	0.060	0.064	0.413	0.363	0.348	0.406	0.203	0.429	0.393	0.364
$\gamma = 0.8$	0.086	0.114	0.119	0.109	0.058	0.054	0.079	0.053	0.377	0.369	0.289	0.383	0.349	0.453	0.455	0.446
$\gamma = 1.0$	0.544	0.656	0.638	0.468	0.058	0.054	0.054	0.034	0.550	0.558	0.548	0.540	0.872	0.900	0.888	0.830

Note: The *MQCS* is our proposed test, *SupY*, $\mathcal{T}_n$ and *CUSUM* are the tests proposed by Xiao (2009b), Xiao (2009a) and Xiao and Phillips (2002), respectively. $M_2 = [\tau(1-\tau)/\phi(\Phi^{-1}(\tau))]n^{1/3}$.

statistics exhibit different degrees of distortion as γ increases, since larger value of γ implies more persistent property for u_t. When $\gamma = 0.8$, the DGP is a nearly cointegrated system and the size are inflated for most cases. Second, when the true model has constant coefficients, it is interesting to find that the *SupY* test has better size performance than $MQCS$, $\mathcal{T}_n$ and $CUSUM$ tests but worse power performance under all the considered three error distributions. Therefore, when the cointegration coefficient is constant, none of the four considered tests uniformly dominates the others, even the *SupY* and $CUSUM$ tests are designed specifically for constant cointegration test. Third, when $\beta(u_t, z_t) = sin\left(\sqrt{2}\pi z_t\right)$, we can see that the *SupY* test has no power against the false null hypothesis of cointegration. Although the powers for $\mathcal{T}_n$ and $CUSUM$ are much higher than the *SupY* test, they are lower than our proposed $MQCS$ test. Fouth, $MQCS$ test performs the best when $\beta(u_t, z_t) = [0.5 + \Phi(u_t)]cos\left(\sqrt{2}\pi z_t\right)$, while the $\mathcal{T}_n$ and $CUSUM$ tests based on mean regression display severer size distortion and misleading high power. To summarize, the proposed $MQCS$ test statistic displays reasonably size and good power performance compared to other alternative tests even in small samples.

Finally, we offer additional comment on the impact of bandwidth selection on the size performance of the proposed $MQCS$. As has been mentioned, all the tests exhibits size distortion when γ is large which implies stronger persistence of the data. Therefore when estimating the variance, it is natural to consider using larger M, that is, the selection of M should also depend on γ. However, for simplicity and by convention in the literature, we do not consider such a case as there is yet no former theoretical justice. We expect the size performance could be improved further and leave the selection of more appropriate bandwidth selection in variance estimation for quantile regression as a future research.

5.2.2 Tests on various quantiles

In this section, we investigate the finite sample performance of $MQCS$ test across different quantile levels and the DGP is set as follow:

$$
\begin{aligned}
DGP3: \quad y_t &= \beta(u_t, z_t)x_t + u_t, \quad t = 1, \cdots, n, \\
x_t &= x_{t-1} + v_t, \\
u_t &= \gamma_u u_{t-1} + \epsilon_t, \\
\gamma_u &= \begin{cases} 0, & u_{t-1} < 0, \\ 1, & u_{t-1} \geq 0, \end{cases}, \\
\begin{pmatrix} \epsilon_t \\ v_t \end{pmatrix} &\sim F\left(\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 & \rho \\ \rho & 1 \end{pmatrix}\right).
\end{aligned}
$$

where v_t and ϵ_t are generated from N(0,1) or student-t distributions. Notice that, γ is a function of u_{t-1}. If the previous exogenous shock on the dependent variable is positive, then its impact would continue to the present period. If negative, it has no persistent impact. For the symmetric error distribution above, cointegration relation between y_t and x_t only exists at quantile levels below 0.5. Therefore, the rejection rates at lower quantiles of the test correspond to size performances and the rest scenarios demonstrate the power of the proposed test. We consider $\tau = 0.25, 0.5$ and 0.75 in the following analysis and $n = 500, 1000$ since large sample could provide more accurate estimation for density function at the tail.

It is also interesting to compare the performances of the $MQSC$ test with that of the $SupY$. To this end, we allow the errors following different distributions and consider two function patterns in this section,

1. Constant-coefficient, $\beta(u_t, z_t) = \begin{cases} 1, & u_{t-1} < 0, \\ 0, & u_{t-1} \geq 0. \end{cases}$

2. Functional-coefficient, $\beta(u_t, z_t) = \begin{cases} [0.5 + \Phi(u_t)]cos\left(\sqrt{2}\pi z_t\right), & u_{t-1} < 0, \\ 0, & u_{t-1} \geq 0. \end{cases}$

where the coefficient is a constant at specific quantile in the first case. However it is allowed to be time-varying and also quantile dependent in the second case.

Table 5: Empirical size and power $(n = 500, M = M_2)$

		$MQCS(1)$			$MQCS(2)$			$SupY$		
		$N(0,1)$	$t(2)$	$t(4)$	$N(0,1)$	$t(2)$	$t(4)$	$N(0,1)$	$t(2)$	$t(4)$
Constant-coefficient										
	$\tau = 0.25$	0.076	0.048	0.064	0.090	0.142	0.108	0.088	0.050	0.070
$\rho = 0.5$	$\tau = 0.50$	0.950	0.910	0.950	0.986	0.974	0.984	0.946	0.954	0.962
	$\tau = 0.75$	0.224	0.198	0.214	0.548	0.532	0.514	0.268	0.260	0.234
	$\tau = 0.25$	0.074	0.060	0.086	0.134	0.122	0.158	0.096	0.080	0.108
$\rho = 0.8$	$\tau = 0.50$	0.943	0.938	0.940	0.990	0.918	0.974	0.957	0.942	0.954
	$\tau = 0.75$	0.260	0.276	0.254	0.508	0.438	0.492	0.264	0.226	0.302
Functional-coefficient										
	$\tau = 0.25$	0.090	0.106	0.062	0.124	0.112	0.162	0.112	0.112	0.102
$\rho = 0.5$	$\tau = 0.50$	0.964	0.960	0.938	0.980	0.999	0.986	0.934	0.970	0.958
	$\tau = 0.75$	0.276	0.258	0.276	0.514	0.218	0.490	0.243	0.320	0.298
	$\tau = 0.25$	0.074	0.060	0.090	0.180	0.168	0.164	0.092	0.138	0.128
$\rho = 0.8$	$\tau = 0.50$	0.964	0.924	0.942	0.986	0.982	0.982	0.978	0.964	0.960
	$\tau = 0.75$	0.252	0.244	0.250	0.532	0.444	0.468	0.271	0.290	0.250

Note: The $MQCS$ is our proposed test, where the value in parentheses indicates the order of Taylor expansion. $SupY$ is the test proposed by Xiao (2009b). $M_2 = [\tau(1-\tau)/\phi(\Phi^{-1}(\tau))]n^{1/3}$

Table 6: Empirical size and power ($n = 1000, M = M_2$)

		$MQCS(1)$			$MQCS(2)$			$SupY$		
		$N(0,1)$	$t(2)$	$t(4)$	$N(0,1)$	$t(2)$	$t(4)$	$N(0,1)$	$t(2)$	$t(4)$
					Constant-coefficient					
$\rho = 0.5$	$\tau = 0.25$	0.060	0.062	0.066	0.162	0.120	0.116	0.072	0.094	0.072
	$\tau = 0.50$	0.982	0.990	0.974	0.998	1.000	1.000	0.982	0.982	0.974
	$\tau = 0.75$	0.186	0.256	0.194	0.464	0.450	0.472	0.264	0.342	0.232
$\rho = 0.8$	$\tau = 0.25$	0.058	0.042	0.058	0.168	0.092	0.126	0.086	0.058	0.066
	$\tau = 0.50$	0.976	0.970	0.978	0.998	0.980	1.000	0.994	0.970	0.970
	$\tau = 0.75$	0.192	0.214	0.198	0.450	0.456	0.474	0.244	0.250	0.236
					Functional-coefficient					
$\rho = 0.5$	$\tau = 0.25$	0.096	0.094	0.064	0.110	0.104	0.164	0.108	0.096	0.070
	$\tau = 0.50$	0.984	0.982	0.980	1.000	0.998	1.000	0.994	0.994	0.982
	$\tau = 0.75$	0.272	0.302	0.212	0.462	0.460	0.512	0.214	0.304	0.234
$\rho = 0.8$	$\tau = 0.25$	0.084	0.072	0.090	0.170	0.160	0.156	0.106	0.110	0.118
	$\tau = 0.50$	0.986	0.984	0.982	1.000	0.998	1.000	0.988	0.990	0.990
	$\tau = 0.75$	0.240	0.240	0.226	0.470	0.480	0.436	0.228	0.194	0.278

Note: The $MQCS$ is our proposed test, where the value in parentheses indicates the order of Taylor expansion. $SupY$ is the test proposed by Xiao (2009b). $M_2 = [\tau(1-\tau)/\phi(\Phi^{-1}(\tau))]n^{1/3}$

The results are shown in Table 5 and 6. Both the proposed $MQCS$ test and the traditional $SupY$ test are able to distinguish the cointegration relation at low quantile levels (e.g. $\tau = 0.25$) and non cointegration relation at high quantile levels (e.g. $\tau = 0.5$, 0.75).

When the coefficient of the true model is constant, $MQCS(1)$ test where the coefficient is estimated by the first order Taylor expansion has similar size performances to that of $SupY$ test for all three distributions, but the size of $MQCS(2)$ exhibits size distortion. Meanwhile, $MQCS(2)$ test has the best power performances, followed by $SupY$ test and $MQCS(1)$ test.

Turning to the functional coefficients model, the sizes of $SupY$ become significantly inflated. When the degree of endogeneity is high ($\rho = 0.8$), it reaches the highest level of around 13.8% in the low quantile ($\tau = 0.25$). $MQCS(1)$ test displays size improvement by reducing 56.5% ((0.138-0.06)/0.138) distortion and still has comparable power. While $MQCS(2)$ test suffers the same size distortion as $SupY$ test does, it exhibits the highest power for all cases. As the sample size becomes larger ($n = 1000$), $MQCS$ tests show slightly better size performance in functional coefficient model while $SupY$ test displays considerable size improvement in constant coefficient model, which is consistent with the theory that the convergence rate of nonparametric estimation is slower than that of parametric estimation.

6 An empirical application to PPP

6.1 Background

The PPP theory is among the most influential theories about exchange rate determination, which also serves as the basis for the formation of other exchange rate determination theories. Because of its implication for policy-making, whether it could be applied in interpreting real economy has drawn much attention from both the government and researchers. In the literature, most research believe that the PPP theory correctly explains the determination of the long-term trend of the exchange rate, that is, in the long run, the trend of the exchange rate is parallel to that of the purchasing power parity. However, in the short term, the real exchange rate fluctuates sharply, and the adjustment speed towards long-term equilibrium is very slow. This phenomenon is usually termed as the "purchasing power parity puzzle" (Rogoff, 1996). Moreover, based on employing different research methods, the empirical findings in the literature have not yet reached a consensus on the applicability of PPP theory. The failure might be caused by the complex relationships between economics variables, which is more likely to be asymmetric, nonlinear and time-varying.[6] As have been well documented in our analysis above, our proposed method is flexible enough, incorporating most of the existing models as special cases. In the following analysises, we apply our proposed method and re-examines the PPP theory for three major Asian countries: China, South Korea and Japan against US.

The PPP theory suggests that one unit domestic currency should have the same purchasing power when converted into foreign currency. In a perfect competitive market environment, if the prices for the same goods are different in different regions, then arbitrageurs can make profits through the price difference until the prices converge. In the literature, the PPP theory is usually expressed as the following form:

$$s_t = \alpha + \beta(p_t - p_t^*) + u_t, \tag{16}$$

where s_t, p_t, p_t^* are the logarithms of the nominal exchange rate, domestic price and foreign price at time t, respectively, and the slope coefficient, β is expected to be positive. In this paper, we re-estimate the above model (16) by our proposed method.

6.2 Data and Preliminary Statistical Analysis

In this article, we calculate the nominal exchange rate, s_t by converting the US dollar to the domestic currency directly, and use Producer Price Index (PPI) as the measure for the price

[6] While linear cointegration analysises usually fail in supporting the PPP theory, studies using nonlinear cointegration analysis tend to find evidence in favour of it (Cai et al., 2009; Hong and Phillips, 2010).

level.[7] We choose the interest rate spread as the smoothing variable since first, it could affect the flow of short-term capital across the world, resulting in exchange rate changes, and second, the interest rate is also closely related to the price level. Following Li et al. (2015), we construct the smoothing variable as follows:

$$z_t = (R_{t,l} - R_{t,s}) - (R^*_{t,l} - R^*_{t,s}), \tag{17}$$

where $R_{t,l}, R_{t,s}$ are the domestic long-run and short-run interest rate, respectively, and $R^*_{t,l}$ and $R^*_{t,s}$ are their foreigner counterparts. Apart for South Korea, the 10-year treasury bond yield and the three-month inter-bank lending rate are used as proxy indicators for long-term and short-term interest rates. Regarding for the South Korea, since the three-month interbank lending rate is not available, we use the three-month certificate of deposit yield rate as an alternative. The sample period for Chinese YUAN (CNY) against the US dollars (USD) is from August 2005, the time at which China conducted its reform moving away from a fixed exchange rate, to December 2020. Due to data availability, the time span of Japaneses YEN (JPY) against USD is from January 1989 to December 2020, and the data for South Korean won (KRW) against USD is from October 2000 to December 2020. The PPI index is calculated on the basis of 2010. All the data are collected from the International Financial Statistics (IFS) and Wind databases.

Figure 1 plots the time series of exchange rate and PPI difference for each of the three country pairs. From the left plot, we could see that the CNY exchange rate and the nominal exchange rate determined by PPP differ greatly at different stages. During the 10-year time period of the so called "721 exchange rate reform" from 2005 to the end of 2014, the CNY exchange rate against the US dollar exhibits a unilateral appreciation trend, and the divergence between the CNY exchange rate and the exchange rate suggested by PPP is also significant. The CNY enters a depreciation period after 2015 due to, on the one hand, the central bank adjust the CNY central parity pricing mechanism; on the other hand, multiple macro shocks happen such as domestic stock market crash, and the Federal Reserve's interest rate hike. At the same time, the price levels characterized by PPI difference also demonstrates an upward trend. This depreciation ends after 2017 when the fourth exchange rate reform is carried out to eliminate the impact of unilateral market irrational expectations. During this time period, it is interesting to see that the CNY exchange rate determined by the PPP theory and the nominal exchange rate display an opposite trend. The middle plot shows that the fluctuation of the nominal exchange rate of Japan against the US dollar is cyclical, while the exchange rate

[7] Compared with Consumer Prices Index (CPI), PPI also considers intermediate products, and therefore involves a wider range of goods portfolios. Moreover, the PPI is more sensitive since prices changes usually initial in the production sectors.

suggested by by PPP has a downward trend. From the the plot on the right we find that the KRW experienced the largest depreciation rate of 25.7% against the US dollar at around 2008 due to the financial crisis. During this period, the PPI difference between the two countries also rise sharply. Prior to this time period, the exchanges rate of KRW and the prices difference display quite similar patters and their relationship becomes less unclear afterwards.

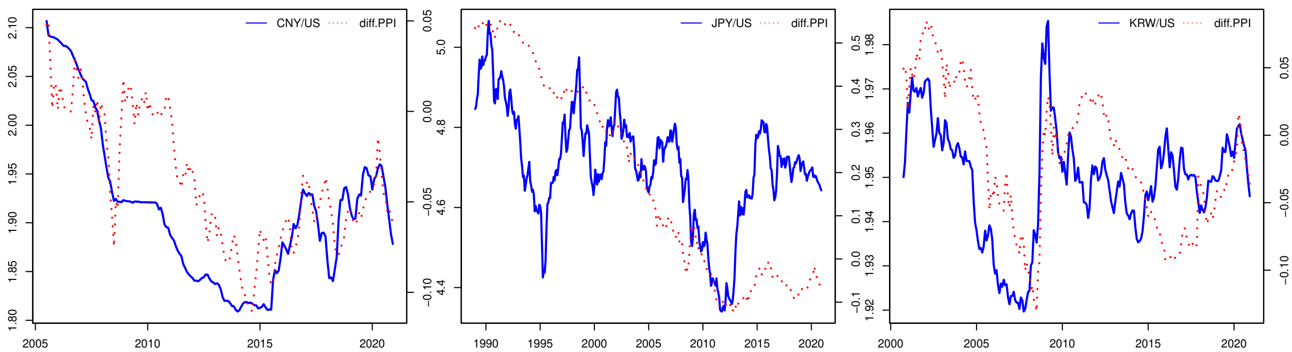

Figure 1: Plots of the logarithms of the nominal exchange rate (i.e. s_t, see the left axes) and the price difference (i.e. $p_t - p_t^*$, see the right axes).

Table 7 shows descriptive statistics of each variable. The values of skewness and kurtosis suggest that the distributions of all series are asymmetrical with heavy tails . The *Jarque-Bera* test shows further evidence that all the series are non-normal with one exception: South Korea. All these results suggest that quantile regression method is preferable in analysing the relationships between the variables, which could provide more comprehensive and reliable results.

Table 7: Descriptive statistics

		N	Mean	SD	Skewness	Kurtosis	JB-statistic
$p_t - p_t^*$	China	186	-0.039	0.036	0.199	-0.964	7.982**
	Japan	384	0.188	0.231	0.188	-1.516	38.650***
	South Korea	243	-0.016	0.051	-0.006	-0.944	8.694**
s_t	China	186	1.913	0.079	0.802	-0.090	20.311***
	Japan	384	4.703	0.144	-0.407	0.265	11.969***
	South Korea	243	7.025	0.092	-0.196	0.092	1.720
z_t	China	186	-0.597	1.152	0.494	-0.657	10.782***
	Japan	384	-0.396	1.424	0.255	-1.073	22.261***
	South Korea	243	-0.526	1.020	0.310	-0.890	11.457***

Note: JB denotes the Jarque-Bera normality test. ***, ** and * indicate significance at 1%, 5% and 10%level, respectively.

We also explore the stationarity of the series s_t, $p_t - p_t^*$ and z_t by using four conventional tests: the GLS modified ADF (DF-GLS) test by Elliott et al. (1996), nonparametric adjusted z_α test (or PP) by Phillips and Perron (1988), structural ZA test by Zivot and K. (1992) and stationary KPSS test by Kwiatkowski et al. (1992). The results are summarized in Table 8. We could see that in most cases, the tests suggest the series s_t, $p_t - p_t^*$ to be non-stationary while that z_t is stationary.

Table 8: Unit roots test results

		DF-GLS	PP	ZA	KPSS
$p_t - p_t^*$	China	-0.561	-2.684^*	-4.505	2.270***
	Japan	1.582	-0.975	-3.769	6.298***
	South Korea	-1.232	-1.946	-3.392	1.757***
s_t	China	-0.068	-2.504	-3.692	1.590***
	Japan	-1.292	-2.232	-3.212	1.770**
	South Korea	-3.036^{***}	-2.570	-3.766	0.315
z_t	China	-3.766^{***}	-10.740^{***}	-7.470^{***}	0.091
	Japan	-2.045^{**}	-16.654^{***}	-6.566^{***}	0.062
	South Korea	-2.392^{**}	-13.069^{***}	-6.984^{***}	0.246

Note: KPSS test is a test statistics for the null hypothesis that the series is stationary while other tests assume that the series are nonstationary. ***, ** and * indicate significance at 1%, 5% and 10%level, respectively.

6.3 Estimation Results

In Figure 2, we plot the estimated results for β by our proposed method at three quantile levels: 0.2, 0.5 and 0.8, representing economy status of recession, normal and prosperity. For comparison, we also report OLS estimation results. In general, we can find that the estimated coefficients fluctuate greatly with the change of interest rate spread, and there are significant differences for the estimated coefficients under different quantile levels.

Specifically, for China, the PPP exchange rate is generally positively correlated with the nominal exchange rate, consistent with the theoretical prediction of the PPP hypothesis. When the domestic price index rises relative to the foreign price index, it will put domestic export at an disadvantage, leading to the decrease of foreign exchange supply and the rise of exchange rate. We could also find that the variation of cointegration coefficient's change with z_t increases with quantile levels. When the domestic interest rate spread is significantly higher than foreign interest rate spread at the high quantile of CNY, the sign of function coefficient is negative because the capital movements caused by spreads may lead to a fall in the stock market and property prices, or even a burst of asset bubbles. For JPY, the correlation between the PPP exchange rate and the nominal exchange rate shows an obvious trend with the change of interest rate spread. The liberalization degree of the commodity market and the financial market in open economies, represented by Japan, is much higher than that of China, so the relationship between interest rate change and exchange rate is further closer. Different from CNY and JPY, the relationships between the PPP exchange rate and the nominal exchange rate of the KRW at different quantile levels share a high similarity in the evolution path, and fluctuate around a fixed level. And in the case of extreme interest rate changes, the positive correlation between the variables becomes more significant.

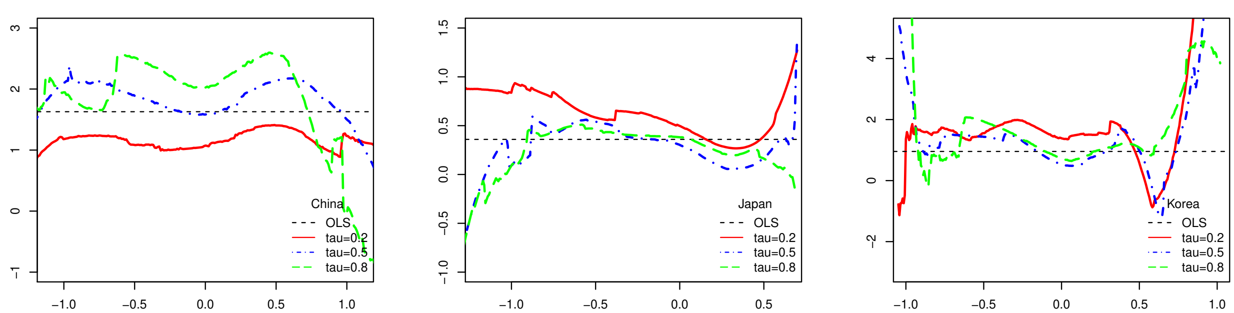

Figure 2: Plots of $\hat{\beta}$ against interest rate spread (z_t) on the x-axis: OLS v.s. QR estimation.

We report in the results of cointegration test in Table 9, and for comparison, we also report the results of $SupY$ test by Xiao (2009b). The results obtained form this two tests are in general different at different quantiles.

First, we analyse the results for China. we find that $SupY$ test tends to reject the null hypothesis at upper tail quantile, whereas it cannot reject the null hypothesis at the low quantile levels. This provides evidence in support of PPP theory only in the the process of exchange rate appreciation for all the three pairs. By contrats, the results of the MQSC-statistic for CNY/US pair at different quantiles are not significant at 10% significance level, thus fail to reject the null hypothesis of cointegration. From the simulation results in Section 5, MQCS test has better/comparable finite sample performance than/to the $SupY$ test. Moreover, considering that China has conducted reforms gradually, the results for $SupY$ test may be not reliable as it assumes a constant cointegration relationship. This suggests that PPP theory is valid for exchange rate policy-making between China and the United States. The similar pattern can also be found for KWR/US, except the case when $\tau = 0.1$. In the extreme quantile, the sign of function coefficient is also contrary to the theory, thus PPP theory is not applicable. Finally, the MQSC test for Japan and the United States is significant at 1% significance level when $\tau \in [0.2, 0.4]$, while it cannot reject the cointegrarion hypothesis in the high quantile levels. This asymmetric pattern is consistent with the findings of Protopapadakis and Stoll (1986), which show that the PPP theory is more likely to be held in countries with high inflation rate. On the one hand, there may be arbitrage opportunities in these countries, but trade barriers and other factors make speculation difficult to achieve Lucio et al. (2004); On the other hand, structural factors will also lead to the continuous deviation of nominal exchange rate from PPP, such as changes in technology and productivity in non trade sectors.

Table 9: Results for the quantile cointegration test

	$\tau = 0.1$	$\tau = 0.2$	$\tau = 0.3$	$\tau = 0.4$	$\tau = 0.5$	$\tau = 0.6$	$\tau = 0.7$	$\tau = 0.8$	$\tau = 0.9$
					$SupY$ test				
China	0.180	0.445	0.812	1.281	1.895**	2.524***	3.740***	4.559***	4.590***
Japan	0.086	0.303	0.721	1.051	2.165**	3.663***	4.458***	4.712***	4.616***
South Korea	-0.028	0.020	0.278	0.526	1.152	2.508***	3.638***	4.199***	4.332***
					MQCS test				
China	1.395	3.179	2.936	4.184	2.114	1.985	2.040	2.274	2.373
Japan	1.576	2.087***	2.956***	5.863***	2.302	7.559	2.236	3.240	4.974
South Korea	4.768*	3.374	3.747	4.464	3.996	5.411	5.336	6.085	6.499

Note: Critical values for the constant coefficient quantile cointegration test are 1.616, 1.842 and 2.326 from Xiao and Phillips (2002) .The p-values for functional coefficient quantile cointegration test statistic,$MQCS(\tau)$, are computed by 1000 bootstrap replications. ***, ** and * indicate significance at 1%, 5% and 10%level, respectively.

7 Conclusion

This paper proposes a generalized quantile cointegrating regressive model for nonstationary time series, allowing coefficients to be unknown functions of informative covariates at each quantile level. A local polynomial nonparametric estimator are proposed which is shown to be $n\sqrt{h}$ consistent. Due to endogeneity issue, the proposed estimator suffers from asymptotic bias and we further propose a "fully modified" estimators, and establish its mixed asymptotic normality. We also propose two test statistics and study their asymptotic properties for two important inference problems: parameter (in)stability and null hypothesis of cointegration. Monte Carlo Simulation results confirm that our proposed tests have good finite sample performance. Finally, we apply our method in re-examining the PPP hypothesis for three major Asian countries China, Japan and South Korea against US, and the results are more in favor of the PPP hypothesis in comparison to other existing method assuming constant relationship. However, at low quantile levels, the test supports the alternative of no cointegration for Japan and the United States. Thus the regulatory authority might pay more attention to the downward movement of the nominal exchange rate when it is in the low level.

References

Andrews, D. W. K. (1991). Heteroskedasticity and autocorrelation consistent covariance matrix estimation. *Econometrica*, 59(3):817–858.

Balke, N. S., Doan, T., and Fomby, T. B. (1997). Threshold cointegration. *International Economic Review*, 38(3):627–645.

Bickel, P. J. and Freedman, D. A. (1981). Some asymptotic theory for the bootstrap. *The annals of statistics*, pages 1196–1217.

Bierens, H. J. (1997). Nonparametric cointegration analysis. *Journal of Econometrics*, 77(2):379–404.

Bierens, H. J. and Martins, L. F. (2010). Time-varying cointegration. *Econometric Theory*, 26(5):1453–1490.

Burdekin, R. C. K. and Siklos, P. L. (2012). Enter the dragon: Interactions between chinese, us and asia-pacific equity markets, 1995–2010. *Pacific-Basin Finance Journal*, 20(3):521–541.

Cai, Z., Fan, J., and Yao, Q. (2000). Functional-coefficient regression models for nonlinear time series. *Journal of the American Statistical Association*, 95(451):941–956.

Cai, Z., Qi, L., and Park, J. Y. (2009). Functional-coefficient models for nonstationary time series data. *Journal of Econometrics*, 148(2):101–113.

Cai, Z. and Xu, X. (2009). Nonparametric quantile estimations for dynamic smooth coefficient models. *Journal of the American Statistical Association*, 104(485):371–383.

Caner, M. and Hansen, B. E. (2001). Threshold autoregression with a unit root. *Econometrica*, 69(6):1555–1596.

Caner, M. and Hansen, B. E. (2004). Instrumental variable estimation of a threshold model. *Econometric Theory*, 20(05):813–843.

Chen, B. and Hong, Y. (2012). Testing for smooth structural changes in time series models via nonparametric regression. *Econometrica*, 80(3):1157–1183.

Cho, J. S., Kim, T. H., and Shin, Y. (2015). Quantile cointegration in the autoregressive distributed-lag modeling framework. *Journal of Econometrics*, 188(1):281–300.

Culver, S. E. and Papell, D. H. (1999). Long-run purchasing power parity with short-run data: evidence with a null hypothesis of stationarity. *Journal of International Money and Finance*, 18(5):751–768.

Elliott, G., Rothenberg, T. J., and Stock, J. H. (1996). Efficient tests for an autoregressive unit root. *Econometrica*, 64(4):813–836.

Engle, R. F. and Granger, C. W. J. (1987). Co-integration and error correction: representation, estimation, and testing. *Econometrica*, 55(2):251–276.

Fan, J. and Gijbels, I. (1996). *Local polynomial modelling and its applications*. Chapman and Hall, London.

Giacomini, R., Politis, D. N., and White, H. (2013). A warp-speed method for conducting monte carlo experiments involving bootstrap estimators. *Econometric Theory*, 29(3):567–589.

Granger, C. W. J. (1981). Some properties of time series data and their use in econometric model specification. *Journal of Econometrics*, 16(1):121–130.

Granger, C. W. J. (2008). Non-linear models: Where do we go next-time varying parameter models? *Studies in Nonlinear Dynamics and Econometrics*, 12(3):1639–1639.

Granger, C. W. J. (2010). Some thoughts on the development of cointegration. *Journal of Econometrics*, 158(1):3–6.

Granger, C. W. J. and Newbold, P. (1974). Spurious regressions in econometrics. *Journal of Econometrics*, 2(2):111–120.

Granger, C. W. J. and Terasvirta, T. (1993). *Modelling non-linear economic relationships*. Oxford University Press, Oxford.

Gregory, R. P. and Shelley, G. (2011). Purchasing power parity and the chinese yuan. *Economics Bulletin*, 31(2):1247–1255.

Gu, J. and Liang, Z. (2014). Testing cointegration relationship in a semiparametric varying coefficient model. *Journal of Econometrics*, 178:57–70.

Hall, A. D., Anderson, H. M., and Granger, C. W. J. (1992). A cointegration analysis of treasury bill yields. *Review of Economics and Statistics*, 74(1):116–126.

Hall, P. and Sheather, S. J. (1988). On the distribution of a studentized quantile. *Journal of the Royal Statistical Society: Series B (Methodological)*, 50(3):381–391.

Hong, S. H. and Phillips, P. C. B. (2010). Testing linearity in cointegrating relations with an application to purchasing power parity. *Journal of Business and Economic Statistics*, 28(1):96–114.

Juhl, T. and Xiao, Z. (2005). Partially linear models with unit roots. *Econometric Theory*, 21(5):877–906.

Knight, K. (1989). Limit theory for autoregressive-parameter estimates in an infinite-variance random walk. *Canadian Journal of Statistics*, 17(3):261–278.

Koenker, R. and Xiao, Z. (2004). Unit root quantile autoregression inference. *Journal of the American Statistical Association*, 99(467):775–787.

Koenker, R. and Xiao, Z. (2006). Quantile autoregression. *Journal of the American Statistical Association*, 101(475):980–990.

Kwiatkowski, D., Phillips, P. C. B., Schmidt, P., and Shin, Y. (1992). Testing the null hypothesis of stationarity against the alternative of a unit root: How sure are we that economic time series have a unit root? *Journal of Econometrics*, 54(1-3):159–178.

Lee, C. C. and Zeng, J. H. (2011). Revisiting the relationship between spot and futures oil prices: evidence from quantile cointegrating regression. *Energy Economics*, 33(5):924–935.

Lettau, M. and Ludvigson, S. (2001). Consumption, aggregate wealth, and expected stock returns. *The Journal of Finance*, 56(3):815–849.

Li, D., Phillips, P. C. B., and Gao, J. (2020). Kernel-based inference in time-varying coefficient cointegrating regression. *Journal of Econometrics*, 215(2):607–632.

Li, H., Lin, Z., and Hsiao, C. (2015). Testing purchasing power parity hypothesis: a semiparametric varying coefficient approach. *Empirical Economics*, 48(1):427–438.

Li, H., Zheng, C., and Guo, Y. (2016). Estimation and test for quantile nonlinear cointegrating regression. *Economics Letters*, 148:27–32.

Liang, H. Y., S., Y., and Wang, Q. (2019). Functional-coefficient quantile regression with nonstationary time series. *Working paper*.

Lin, Y., Tu, Y., and Yao, Q. (2020). Estimation for double-nonlinear cointegration. *Journal of Econometrics*, 216(1):175–191.

Lucio, S. M., Taylor, P., and Ibrahim, C. (2004). Nonlinear dynamics in deviations from the law of one price: a broad-based empirical study. *Journal of International Money and Finance*, 23(1):1–25.

Nelson, C. R. and Plosser, C. R. (1982). Trends and random walks in macroeconmic time series: some evidence and implications. *Journal of Monetary Economics*, 10(2):139–162.

Park, J. Y. and Hahn, S. B. (1999). Cointegrating regressions with time varying coefficients. *Econometric Theory*, pages 664–703.

Park, J. Y. and Phillips, P. C. B. (2001). Nonlinear regressions with integrated time series. *Econometrica*, 69(1):117–161.

Perron, P. (1989). The great crash, the oil price shock, and the unit root hypothesis. *Econometrica*, 57(6):1361–1401.

Pesaran, M. H. and Shin, Y. (1999). *An autoregressive distributed lag modelling approach to cointegration analysis*. Cambridge University Press, Cambridge.

Phillips, P. C. (2010). Bootstrapping i (1) data. *Journal of Econometrics*, 158(2):280–284.

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

Phillips, P. C. B., Li, D., and Gao, J. (2017). Estimating smooth structural change in cointegration models. *Journal of Econometrics*, 196(1):180–195.

Phillips, P. C. B. and Perron, P. (1988). Testing for a unit root in time series regression. *Biometrika*, 75(2):335–346.

Pollard, D. (1991). Asymptotics for least absolute deviation regression estimators. *Econometric Theory*, 7(2):186–199.

Protopapadakis, A. A. and Stoll, H. R. (1986). The law of one price in international commodity markets: a reformulation and some formal tests. *Journal of International Money and Finance*, 5(3):335–360.

Rogoff, K. (1996). The purchasing power parity puzzle. *Journal of Economic Literature*, 34(2):647–668.

Saikkonen, P. (1991). Asymptotically efficient estimation of cointegration regressions. *Econometric Theory*, 7(1):1–21.

Saikkonen, P. and Choi, I. (2004). Cointegrating smooth transition regressions. *Econometric Theory*, 20(2):301–340.

Serletis, A. and Banack, D. (1990). Market efficiency and cointegration: an application to petroleum markets. *Review of Futures Markets*, 9(2):372–385.

Shin, Y. (1994). A residual-based test of the null of cointegration against the alternative of no cointegration. *Econometric Theory*, 10(1):91–115.

Silverman, B. W. (1986). *Density estimation for statistics and data analysis*. Chapman and Hall, London.

Spokoiny, V., Wang, W., and Härdle, W. K. (2013). Local quantile regression. *Journal of Statistical Planning and Inference*, 143(7):1109–1129.

Tu, Y., Liang, H. Y., and Wang, Q. (2021). Nonparametric inference for quantile cointegrations with stationary covariates. *Journal of Econometrics,* Forthcoming.

Tu, Y. and Wang, Y. (2019). Functional coefficient cointegration models subject to time–varying volatility with an application to the purchasing power parity. *Oxford Bulletin of Economics & Statistics*, 81(6):1401–1423.

Uematsu, Y. (2019). Nonstationary nonlinear quantile regression. *Econometric Reviews*, 38(4):386–416.

Xiao, Z. (2009a). Functional-coefficient cointegration models. *Journal of Econometrics*, 152(2):81–92.

Xiao, Z. (2009b). Quantile cointegrating regression. *Journal of Econometrics*, 150(2):248–260.

Xiao, Z. and Phillips, P. C. B. (2002). A CUSUM test for cointegration using regression residuals. *Journal of Econometrics*, 108(1):43–61.

Yu, K. and Jones, M. C. (1998). Local linear quantile regression. *Journal of the American Statistical Association*, 93(441):228–237.

Zivot, E. and K., A. D. W. (1992). Further evidence on the great crash, the oil price shock and the unit root hypothesis. *Journal of Business and Economic Statistics*, 10(3):251–270.

Appendix A Proofs of Theorems

Proof of Lemma 1

By the defining of $\psi_\tau(u_{t\tau}) = \tau - I(u_{t\tau} < 0)$, it is easy to verify that

$$P(u_{t\tau} < 0|\mathcal{F}_t) = P(u_t < F^{-1}(\tau)|\mathcal{F}_t) = \tau,$$

$$E[E(\psi_\tau(u_{t\tau})|\mathcal{F}_t)] = \tau - P(u_{t\tau} < 0|\mathcal{F}_t) = 0,$$

$$E[E(\psi_\tau(u_{t\tau})^2|\mathcal{F}_t)] = E[\tau^2 - 2\tau I(u_{t\tau} < 0) + I(u_{t\tau} < 0)|\mathcal{F}_t] = \tau^2 - 2\tau^2 + \tau = \tau(1-\tau).$$

Then, by the Law of Iterated Expectation,

$$E(\psi_\tau(u_{t\tau})) = 0, \qquad E(\psi_\tau(u_{t\tau})^2) = \tau(1-\tau).$$

Further, notice that,

$$\begin{aligned}
\frac{1}{nh}\sum_{t=1}^{n} \underline{K}_t\psi_\tau(u_{t\tau})\underline{K}_{t+j}\psi_\tau(u_{t+j\tau}) =& \frac{1}{nh}\sum_{t=1}^{n} \psi_\tau(u_{t\tau})^2[K(z_t - z) - EK(z_t - z)]^2 \\
=& E\left(E\left(\frac{1}{nh}\sum_{t=1}^{n} \psi_\tau(u_{t\tau})^2[K(z_t - z) - EK(z_t - z)]^2|\mathcal{F}_t\right)\right) \\
=& \tau(1-\tau)\frac{1}{h}[EK(z_t - z)^2 - (EK(z_t - z))^2] \\
=& \tau(1-\tau)\frac{1}{h}\left[\int k^2\left(\frac{y-z}{h}\right)f(y)dy - \left[\int k\left(\frac{y-z}{h}\right)f(y)dy\right]^2\right] \\
=& \tau(1-\tau)f_z(z)\int k^2(u)du + O_p(h), \quad \text{if } j = 0, \\
\frac{1}{nh}\sum_{t=1}^{n} \underline{K}_t\psi_\tau(u_{t\tau})\underline{K}_{t+j}\psi_\tau(u_{t+j\tau}) =& O_p(h) \xrightarrow{p} 0, \quad \text{if } j \neq 0,
\end{aligned}$$

and

$$\frac{1}{n}\sum_{t=1}^{n} v_t v_{t+j} \xrightarrow{p} \Gamma_v(j).$$

The desired results in Lemma 1 then follows directly.

Proof of Theorem 1

For simplicity, consider the case for $q = 1$ and denote $\beta(\tau, z_t)$ as $\beta(z_t)$ in a given quantile level, we then have the following first order approximation for $\beta(z_t)$ at a neighbourhood around the point z,

$$\beta(z_t) \approx \beta(z) + \beta^{(1)}(z)(z_t - z).$$

Let $\tilde{\beta} = \beta(z_t) - \beta(z) - \beta^{(1)}(z)(z_t - z),\ u_t^* = u_t + x_t'\tilde{\beta}$, then model (2) could be rewritten as

$$y_t = x_t'\beta(z) + x_t'\beta^{(1)}(z)(z_t - z) + u_t^*.$$

Define $X_{tz} = (x_t', x_t'(z_t - z)/h)',\ \boldsymbol{\beta}(z) = \left(\beta(z), \beta^{(1)}(z)\right)'$, the above model then could be simplified as

$$y_t = X_{tz}'\boldsymbol{\beta}(z) + u_t^*,$$

and the local polynomial estimator of the functional coefficient can be obtained by solving the problem

$$\min_{\beta} \sum_{t=1}^{n} \rho_\tau \left(y_t - X_{tz}'\boldsymbol{\beta}(z)\right) K(z_t - z), \tag{18}$$

where $K(z_t - z) = k\left(\frac{z_t - z}{h}\right)$.

Denote $v_n = n\sqrt{h},\ \pi_0 = v_n\left(\beta(z) - \beta(z)\right),\ \pi_1 = v_n h\left(\beta^{(1)}(z) - \beta^{(1)}(z)\right),\ \pi = (\pi_0, \pi_1)'$ and the corresponding estimators $\hat{\pi}_0 = v_n\left(\hat{\beta}(z) - \beta(z)\right),\ \hat{\pi}_1 = v_n h\left(\hat{\beta}^{(1)}(z) - \beta^{(1)}(z)\right),\ \hat{\pi} = (\hat{\pi}_0, \hat{\pi}_1)'$, then

$$\rho_\tau\left(y_t - X_{tz}'\hat{\boldsymbol{\beta}}(z)\right) = \rho_\tau\left\{u_{t\tau}^* - X_{tz}'\left(\hat{\boldsymbol{\beta}}(z) - \boldsymbol{\beta}(z)\right)\right\} = \rho_\tau\left(u_{t\tau}^* - X_{tz}'(v_n^{-1}\hat{\pi})\right),$$

where $u_{t\tau}^* = y_t - X_{tz}'\hat{\boldsymbol{\beta}}(z)$, and (18) can be further written as

$$\begin{aligned}
\hat{\pi} &= arg\,\min_{\pi} \sum_{t=1}^{n} \rho_\tau\left(u_{t\tau}^* - X_{tz}'(v_n^{-1}\pi)\right) K(z_t - z) \\
&= arg\,\min_{\pi} \sum_{t=1}^{n} \left\{\rho_\tau\left(u_{t\tau}^* - X_{tz}'(v_n^{-1}\pi)\right) - \rho_\tau(u_{t\tau}^*)\right\} K(z_t - z) \\
&:= arg\,\min_{\pi} H_n(\pi).
\end{aligned}$$

Since $H_n(\pi)$ is a convex stochastic function, Knight (1989) (see alsoPollard (1991)) shows that if the finite-dimensional distributions of $H_n(\pi)$ converge weakly to those of $H(\pi)$,and $H(\pi)$ has a unique minimum, then the minimizer of $H(\pi)$ converges in distribution to that of $H_n(\pi)$.

Denote $\psi_\tau(u) = \tau - I(u < 0)$, if $u \neq 0$, then by the Knight identity,

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

Thus $H_n(\pi)$ could be expanded as

$$\begin{aligned}
H_n(\pi) &= \sum_{t=1}^{n}\left\{-X_{tz}'(v_n^{-1}\pi)\psi_\tau(u_{t\tau}^*) + \int_0^{X_{tz}'(v_n^{-1}\pi)} \left(I(u_{t\tau}^* \le u) - I(u_{t\tau}^* \le 0)\right) du\right\} K(z_t - z) \\
&= -v_n^{-1}\pi' \sum_{t=1}^{n} X_{tz}K(z_t - z)\psi_\tau(u_{t\tau}^*) + \sum_{t=1}^{n} K(z_t - z)\int_0^{X_{tz}'(v_n^{-1}\pi)} \left(I(u_{t\tau}^* \le u) - I(u_{t\tau}^* \le 0)\right) du \\
&:= -\pi' H_{1n} + H_{2n},
\end{aligned}$$

where

$$H_{1n} = v_n^{-1}\sum_{t=1}^{n} X_{tz}K(z_t - z)\psi_\tau(u_{t\tau}^*),$$

$$H_{2n} = \sum_{t=1}^{n}\eta_t(\pi), \quad \eta_t(\pi) = K(z_t - z)\int_0^{X_{tz}'(v_n^{-1}\pi)} \left(I(u_{t\tau}^* \le u) - I(u_{t\tau}^* \le 0)\right) du.$$

We first consider the asymptotic property for H_{2n}, Let $Y_t = K(z_t - z)$, then according to Assumption 2(iv), we have

$$0 \ \le\ \eta_t(\pi) \ \le\ Cv_n^{-1}|X_{tz}'\pi|K(z_t - z),$$

and the conditional expectation of $\eta_t(\pi)$

$$\begin{aligned}
E\left(\eta_t(\pi)|\mathcal{F}_t\right) &= K(z_t - z)\int_0^{X_{tz}'(v_n^{-1}\pi)}\int_0^r f_t\left(s - x_t'\tilde{\beta} + F^{-1}(\tau)\right) dsdr \\
&= Y_t\int_0^{X_{tz}'(v_n^{-1}\pi)}\int_0^r \left\{f_t\left(s - x_t'\tilde{\beta} + F^{-1}(\tau)\right) - f_t(F^{-1}(\tau))\right\} dsdr \\
&\quad + K(z_t - z)\int_0^{X_{tz}'(v_n^{-1}\pi)}\int_0^r f_t(F^{-1}(\tau)) dsdr \\
&= Y_t\int_0^{X_{tz}'(v_n^{-1}\pi)}\int_0^r \left\{f_t\left(s - x_t'\tilde{\beta} + F^{-1}(\tau)\right) - f_t(F^{-1}(\tau))\right\} dsdr \\
&\quad + \frac{f_t(F^{-1}(\tau))}{2}K(z_t - z)v_n^{-2}(X_{tz}'\pi)^2 \\
&:= R_{1t} + R_{2t}.
\end{aligned}$$

We now consider the property of $\sum_{t=1}^n R_{1t}$ as $n \to \infty$. First, by assumption $2(iv)$, we have

$$
\begin{aligned}
\sum_{t=1}^n R_{1t} &= \sum_{t=1}^n Y_t \int_0^{X'_{tz}(v_n^{-1}\pi)} \int_0^r \left\{ f_t\left(s - x_t'\tilde{\beta} + F^{-1}(\tau)\right) - f_t(F^{-1}(\tau)) \right\} dsdr \\
&\leq C \sum_{t=1}^n Y_t \int_0^{X'_{tz}(v_n^{-1}\pi)} \int_0^r |(s - x_t'\tilde{\beta})|^\lambda dsdr \\
&\leq C v_n^{-(\lambda+2)} \sum_{t=1}^n K(z_t - z)|X'_{tz}\pi|^{\lambda+2} \\
&\quad + C v_n^{-2} \sum_{t=1}^n K(z_t - z)(X'_{tz}\pi)^2 |x_t'\tilde{\beta}|^\lambda \\
&:= R_{11t} + R_{12t}.
\end{aligned}
$$

Since $x_t/\sqrt{n}$ is assumed to be uniformally bounded in Assumption $3(i)$, $|\mu_{2+\lambda}(k)| \leq \infty$ and $nh \to \infty$, we then have

$$
R_{11t} = C(nh)^{-\lambda/2} \frac{1}{n} \sum_{t=1}^n \frac{K(z_t - z)}{h} |\frac{X'_{tz}}{\sqrt{n}}\pi|^{\lambda+2} = O_p((nh)^{-\lambda/2}) \xrightarrow{p} 0.
$$

In addition, under Assumptions $1(i)$-(ii) and $2(i)$, we have

$$
\begin{aligned}
R_{12t} &= C \frac{1}{n} \sum_{t=1}^n \frac{K(z_t - z)}{h} (\frac{X'_{tz}}{\sqrt{n}}\pi)^2 |\frac{x_t'}{\sqrt{n}}\tilde{\beta}|^\lambda \\
&\leq C \frac{1}{n} \sum_{t=1}^n \frac{K(z_t - z)}{h} (\frac{X'_{tz}}{\sqrt{n}}\pi)^2 |\frac{x_t'}{\sqrt{n}}|^\lambda |\frac{z_t - z}{h}|^{2+\lambda} (\sqrt{n}h^{2+\lambda}) \\
&= O_p((\sqrt{n}h^{2+\lambda})) \xrightarrow{p} 0.
\end{aligned}
$$

We thus have shown $\sum_{t=1}^{n} R_{1t} = o_p(1)$, and as a result,

$$\begin{aligned}
\sum_{t=1}^{n} E\left(\eta_t(\pi)|\mathcal{F}_t\right) &= \frac{f_t(F^{-1}(\tau))}{2}\frac{1}{n^2 h}\sum_{t=1}^{n} K(z_t - z)\left(x_t'\pi_0 + x_t'\pi_1(\frac{z_t - z}{h})\right)^2 + o_p(1) \\
&= \frac{f_t(F^{-1}(\tau))}{2}\frac{1}{n^2 h}\sum_{t=1}^{n} \pi_0' x_t x_t' \pi_0 K(z_t - z) \\
&\quad + \frac{f_t(F^{-1}(\tau))}{2}\frac{1}{n^2 h}\sum_{t=1}^{n} 2\pi_0' x_t x_t' \pi_1(\frac{z_t - z}{h})K(z_t - z) \\
&\quad + \frac{f_t(F^{-1}(\tau))}{2}\frac{1}{n^2 h}\sum_{t=1}^{n} \pi_1' x_t x_t' \pi_1(\frac{z_t - z}{h})^2 K(z_t - z) + o_p(1) \\
&\Rightarrow \frac{f_t(F^{-1}(\tau))}{2}\pi_0' \int B_v B_v' \pi_0 \cdot \frac{1}{nh}\sum_{t=1}^{n} K(z_t - z) \\
&\quad + \frac{f_t(F^{-1}(\tau))}{2}\pi_1' \int B_v B_v' \pi_1 \cdot \frac{1}{nh}\sum_{t=1}^{n} (\frac{z_t - z}{h})^2 K(z_t - z) + o_p(1) \\
&\Rightarrow \frac{f_t(F^{-1}(\tau))}{2}\pi' diag\left\{f_z(z)\int B_v B_v', \mu_2 \int B_v B_v'\right\}\pi + o_p(1),
\end{aligned}$$

where $\mu_2 = \int u^2 k(u)du$. Since

$$0 \ \leq \ E\left(\eta_t(\pi)|\mathcal{F}_t\right) = Y_t \int_0^{X_{tz}'(v_n^{-1}\pi)} \int_0^r f_t\left(s - x_t'\tilde{\beta} + F^{-1}(\tau)\right) ds dr \ \leq \ C v_n^{-2} Y_t (X_{tz}'\pi)^2,$$

and according to Xiao (2009b), $\eta_t(\pi) \xrightarrow{p} 0$ holds uniformally for all t,

$$E(\eta_t^2(\pi)|\mathcal{F}_t) \leq max\ \eta_t(\pi) E(\eta_t(\pi)|\mathcal{F}_t) \xrightarrow{p} 0,$$

we then have

$$H_{2n} = \sum_{t=1}^{n} \eta_t(\pi) \xrightarrow{p} \sum_{t=1}^{n} E(\eta_t(\pi)|\mathcal{F}_t) \Rightarrow \frac{f_t(F^{-1}(\tau))}{2}\pi' diag\left\{f_z(z)\int B_v B_v', \mu_2 \int B_v B_v'\right\}\pi + o_p(1).$$

We now consider the asymptotic property for H_{1n}. Let $\xi_t = X_{tz}K(z_t - z)\left(\psi_\tau(u_{t\tau}^*) - \psi_\tau(u_{t\tau})\right)$, then

$$H_{1n} = v_n^{-1}\sum_{t=1}^{n} X_{tz}K(z_t - z)\psi_\tau(u_{t\tau}) + v_n^{-1}\sum_{t=1}^{n} \xi_t.$$

Similarly, let $\lambda_t = -x_t'\tilde{\beta}$, and the conditional expectation for ξ_t would be

$$
\begin{aligned}
E(\xi_t|\mathcal{F}_t) &= X_{tz}K(z_t - z)E\left(I(u_{t\tau} \le 0) - I(u^*_{t\tau} \le 0)|\mathcal{F}_t\right) \\
&= X_{tz}K(z_t - z)E\left(I(u_{t\tau} \le 0) - I(u_{t\tau} \le -x_t'\tilde{\beta})|\mathcal{F}_t\right) \\
&= X_{tz}K(z_t - z)\left(F_t(F^{-1}(\tau)) - F_t(F^{-1}(\tau) + \lambda_t)\right) \\
&= -X_{tz}K(z_t - z)f_t(F^{-1}(\tau))\lambda_t \\
&= f_t(F^{-1}(\tau))X_{tz}x_t'(\beta(z_t) - \beta(z) - \beta^{(1)}(z)(z_t - z))K(z_t - z) \\
&= f_t(F^{-1}(\tau))X_{tz}x_t'(\frac{h^2}{2})\beta^{(2)}(z)(\frac{z_t - z}{h})^2K(z_t - z) + o_p(1) \\
&= f_t(F^{-1}(\tau))X_{tz}x_t'(\frac{h^2}{2})\beta^{(2)}(z)(\frac{z_t - z}{h})^2K(z_t - z) \\
&\quad + f_t(F^{-1}(\tau))(-X_{tz}x_t'(\frac{h^2}{2})\beta^{(2)}(z)(\frac{z_t - z}{h})^2K(z_t - z) + o_p(1) \\
&= f_t(F^{-1}(\tau))X_{tz}x_t'(\frac{h^2}{2})\beta^{(2)}(z)(\frac{z_t - z}{h})^2K(z_t - z) + o_p(1).
\end{aligned}
$$

Further, according to Liang et al. (2019)

$$
v_n^{-1}\sum_{t=1}^{n}[\xi_t - E(\xi_t|\mathcal{F}_t)] = o_p(1),
$$

we then have

$$
\begin{aligned}
H_{1n} &= v_n^{-1}\sum_{t=1}^{n} x_t K(z_t - z)\psi_\tau(u^*_{t\tau}) \\
&= \frac{1}{n\sqrt{h}}\sum_{t=1}^{n} x_t K(z_t - z)\psi_\tau(u_{t\tau}) \\
&\quad + (\frac{h^2}{2})\beta^{(2)}(z)f_t(F^{-1}(\tau)) \cdot \frac{1}{n\sqrt{h}}\sum_{t=1}^{n} x_t x_t'(\frac{z_t - z}{h})^2 K(z_t - z) + o_p(1).
\end{aligned}
$$

Under Assumption 2(iii), 3(ii) and by the Functional Central Limit Theorem,

$$
\begin{bmatrix} \frac{1}{\sqrt{nh}}\sum_{t=1}^{nr}\underline{K}_t\psi_\tau(u_{t\tau}) \\ \frac{1}{\sqrt{n}}\sum_{t=1}^{nr} v_t \end{bmatrix} \Rightarrow \begin{bmatrix} B^k_\psi(r) \\ B_v(r) \end{bmatrix} \Rightarrow BM(0,\Omega), \qquad \Omega = \begin{bmatrix} \omega^2_\psi & \Omega_{\psi v} \\ \Omega_{\psi v} & \Omega_{vv} \end{bmatrix},
$$

and

$$
\frac{1}{n\sqrt{h}}\sum_{t=1}^{n} x_t K(z_t - z)\psi_\tau(u_{t\tau}) \;\Rightarrow\; \int B_v dB^k_\psi + \lambda_{v\psi},
$$

where $\lambda_{v\psi}$ is the one-side long-run variance-covariance matrix between v_t and $K(z_t - z)\psi_\tau(u_{t\tau})$.

Denote $\mu_j = \int u^j K(u) du$ and based on all above results, we have

$$\begin{aligned}
\hat{\pi}_0 &= v_n \left(\hat{\beta}(z) - \beta(z)\right) = \left(\frac{f_t(F^{-1}(\tau))}{} f_z(z) \int B_v B_v'\right)^{-1} \cdot H_{11n} + o_P(1) \\
&\Rightarrow \frac{1}{f_t(F^{-1}(\tau))} \left(f_z(z) \int B_v B_v'\right)^{-1} \left(\int B_v dB_\psi^k + \lambda_{v\psi}\right) + n\sqrt{h}(\frac{h^2}{2})\beta^{(2)}(z)\mu_2,
\end{aligned}$$

which implies

$$n\sqrt{h}\left[\hat{\beta}(z) - \beta(z) - \mathcal{B}\right] \Rightarrow \frac{1}{f_t(F^{-1}(\tau))} \left(f_z(z) \int B_v B_v'\right)^{-1} \left(\int B_v dB_\psi^k + \lambda_{v\psi}\right),$$

where $\mathcal{B} = (\frac{h^2}{2})\beta^{(2)}(z)\mu_2$.

When $q > 1$, the proof is very similar and the corresponding bias formula is now

$$\mathcal{B} = \left(\frac{h^{q+1}}{(q+1)!}\right)\beta^{(q+1)}(z)\mu_{q+1}.$$

The proof of Theorem 1 is then completed.

Proof of Theorem 2

First, we decompose $B_\psi^k(r)$ into,

$$B_{\psi\cdot v}^k(r) = B_\psi^k(r) - \Omega_{\psi v}\Omega_{vv}^{-1} B_v(r),$$

where $B_{\psi\cdot v}^k(r)$ is a Brownian motion uncorrelated with $B_v(r)$ and its variance is $\omega_{\psi\cdot v}^2 = \omega_\psi^2 - \Omega_{\psi v}\Omega_{vv}^{-1}\Omega_{v\psi}$. The asymptotic distribution in Theorem 1 then could be rewritten as

$$\begin{aligned}
n\sqrt{h}\left[\hat{\beta}(z) - \beta(z) - \mathcal{B}\right] \quad \Rightarrow \quad & \frac{1}{f_t(F^{-1}(\tau))} \left(f_z(z) \int B_v B_v'\right)^{-1} \int B_v dB_{\psi\cdot v}^k + \\
& \frac{1}{f_t(F^{-1}(\tau))} \left(f_z(z) \int B_v B_v'\right)^{-1} \times \\
& \left[\int B_v dB_v \Omega_{vv}^{-1}\Omega_{\psi v} + \lambda_{v\psi}\right].
\end{aligned}$$

Since $B^k_{\psi\cdot v}(r)$ is uncorrelated with $B_v(r)$, $\left(f_z(z)\int B_vB_v'\right)^{-1}\int B_vdB^k_{\psi\cdot v}$ would be a mixed normal distribution and

$$
\begin{aligned}
n\sqrt{h}\left[\hat{\beta}^m(z)-\beta(z)-\mathcal{B}\right] &= n\sqrt{h}\left[\hat{\beta}(z)-\beta(z)-\mathcal{B}\right]- \\
&\quad \frac{\widehat{\ }}{f_t(\widehat{F^{-1}(\tau)})}\left(\frac{1}{n^2h}\sum_{t=1}^{n}x_tx_t'K(z_t-z)\right)^{-1}\times \\
&\quad \left[\frac{1}{n}\sum_{t=1}^{n}x_tv_t'\hat{\Omega}_{vv}^{-1}\hat{\Omega}_{\psi v}+\hat{\lambda}^m_{v\psi}\right] \\
&\Rightarrow \frac{1}{f_t(F^{-1}(\tau))}\left(f_z(z)\int B_vB_v'\right)^{-1}\int B_vdB^k_{\psi\cdot v} \\
&\equiv MN\left[0,\frac{\omega^2_{\psi\cdot v}}{f_t(F^{-1}(\tau))^2}\left(f_z(z)\int B_vB_v'\right)^{-1}\right],
\end{aligned}
$$

where $\hat{\lambda}^m_{v\psi}=\hat{\lambda}_{v\psi}-\hat{\lambda}_{vv}\hat{\Omega}_{vv}^{-1}\hat{\Omega}_{\psi v}$ and the proof for Theorem 2 is then complete.

Proof of Theorem 3

Based on the results of Theorem 2, for any z_i

$$
\begin{aligned}
\hat{V}(z_i,\tau) &:= n\sqrt{h}[\hat{\beta}^m(\tau_0,z_i)-\hat{\beta}(\tau_0)] = n\sqrt{h}[\hat{\beta}^m(\tau_0,z_i)-\beta(\tau_0)]+o_p(\sqrt{h}) \\
&= \frac{1}{f_t(F^{-1}(\tau))}\left(\frac{1}{n^2h}\sum_{t=1}^{n}x_tx_t'K(z_t-z_i)\right)^{-1}\left(\frac{1}{n\sqrt{h}}\sum_{t=1}^{n}x_tK(z_t-z_i)\psi_{\tau\cdot v}\right)+o_p(\sqrt{h}) \\
&\Rightarrow \frac{1}{f_t(F^{-1}(\tau))}\left(f_z(z_i)\int B_vB_v'\right)^{-1}\int B_vdB^k_{\psi\cdot v} \\
&\Rightarrow MN(0,\Omega(\tau_0,z_i)).
\end{aligned}
$$

Since

$$
\begin{aligned}
E\left[\frac{1}{h}K\left(\frac{z_i-z}{h}\right)K\left(\frac{z_j-z}{h}\right)\right] &= \frac{1}{h}\int K\left(\frac{z_i-z}{h}\right)K\left(\frac{z_j-z}{h}\right)f_z(z)dz_idz_j \\
&\Rightarrow hf_z(z)\int K(u_i)K(u_j)du_idu_j \\
&= O(h)\xrightarrow{p}0,
\end{aligned}
$$

then the covariance of $\hat{V}(z_i,\tau)$ and $\hat{V}(z_j,\tau)$ is zero and the joint distribution of the kernal estimator of the distinct points z, i.e.,$z_1, z_2, \cdots, z_m$

$$n\sqrt{h}\begin{pmatrix} \hat{V}(\tau,z_1,) \\ \hat{V}(\tau,z_2) \\ \cdots \\ \hat{V}(\tau,z_m) \end{pmatrix} \Rightarrow MN\left(0, \begin{pmatrix} \Omega(\tau,z_1) & 0 & \cdots & 0 \\ 0 & \Omega(\tau,z_2) & \cdots & 0 \\ \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & \cdots & \Omega(\tau,z_m) \end{pmatrix}\right).$$

giving the required results of Theorem 3.

Proof of Theorem 4

For convenience of asymptotic analysis, we define $u_{t\tau} = y_t - Q_{y_t}(\tau|\mathcal{F}_t)$. $QCS(\tau)$ could be decomposed into two parts:

$$\begin{aligned} QCS(\tau) &= \sqrt{\frac{h}{n}}\sum_{t=1}^{\lfloor nr \rfloor}\psi_\tau(\hat{u}_{t\tau}) \\ &= \sqrt{\frac{h}{n}}\sum_{t=1}^{\lfloor nr \rfloor}\psi_\tau(u_{t\tau}) - \sqrt{\frac{h}{n}}\sum_{t=1}^{\lfloor nr \rfloor}[\psi_\tau(u_{t\tau}) - \psi_\tau(\hat{u}_{t\tau})] \\ &:= \mathcal{N}_1 - \mathcal{N}_2. \end{aligned}$$

For the first term, by the Functional Central Limit theorem and use the assumption $\sqrt{h} \to 0$, we have $\mathcal{N}_1 \to o_p(1)$.

Next we consider the limit of $\mathcal{N}_2$. We can show that

$$
\begin{aligned}
E(\mathcal{N}_2|\mathcal{F}_t) &= \frac{\sqrt{h}}{\sqrt{n}}\sum_{t=1}^{\lfloor nr\rfloor} E\left[\psi_\tau(u_{t\tau})-\psi_\tau(\hat{u}_{t\tau})\right] \\
&= \frac{\sqrt{h}}{\sqrt{n}}\sum_{t=1}^{\lfloor nr\rfloor} E\left\{I\left[u_t \le F^{-1}(\tau)+x_t'\left(\hat{\beta}(z)-\beta(z)\right)\right]-I\left[u_t\le F^{-1}(\tau)\right]\right\} \\
&= \frac{\sqrt{h}}{\sqrt{n}}\sum_{t=1}^{\lfloor nr\rfloor} F_t\left[F^{-1}(\tau)+x_t'\left(\hat{\beta}(z)-\beta(z)\right)\right]-F_t\left[F^{-1}(\tau)\right] \\
&= \frac{\sqrt{h}}{\sqrt{n}}\sum_{t=1}^{\lfloor nr\rfloor} f_t\left(F^{-1}(\tau)\right)x_t'\left(\hat{\beta}(z)-\beta(z)\right)+ \\
&\quad \frac{\sqrt{h}}{\sqrt{n}}\sum_{t=1}^{\lfloor nr\rfloor} f_t\left(F^{-1}(\tau)\right)[-]\,x_t'\left(\hat{\beta}(z)-\beta(z)\right)+o_p(1) \\
&= F\left(F^{-1}(\tau)\right)\frac{1}{n\sqrt{n}}\sum_{t=1}^{\lfloor nr\rfloor} x_t' n\sqrt{h}\left(\hat{\beta}(z)-\beta(z)\right)+o_p(1) \\
&\Rightarrow \left(\int B_v'(s)ds\right)\left(\int B_v(s)B_v'(s)ds\right)^{-1}\left(\int B_v(s)B_\psi^k(s)ds+\lambda_{v\psi}\right).
\end{aligned}
$$

We can also show $Var(\mathcal{N}_2)=o_p(1)$.

Note that the variance of $\mathcal{N}_2$ mainly depends on $\frac{\sqrt{h}}{\sqrt{n}}\sum_{t=1}^{\lfloor nr\rfloor}\left[\psi_\tau(u_{t\tau})-\psi_\tau(\hat{u}_{t\tau})\right]$. Define $M_{1t}= min\left\{F^{-1}(\tau)+x_t'\left(\hat{\beta}(z)-\beta(z)\right), F^{-1}\right\}$, $M_{2t}=max\left\{F^{-1}(\tau)+x_t'\left(\hat{\beta}(z)-\beta(z)\right), F^{-1}\right\}$, and $H_t = I\left[u_t\le F^{-1}(\tau)+x_t'\left(\hat{\beta}(z)-\beta(z)\right)\right]-I[u_t\le F^{-1}(\tau)]$, thus for any positive integer m, $|H_t|^m = I(M_{1,t}<u_t<M_{2,t})$. Therefore

$$
\begin{aligned}
Var(\mathcal{N}_2) &= Var\left(\frac{1}{\sqrt{n}}\sum_{t=1}^{\lfloor nr\rfloor}H_t\right) \le Var\left(\frac{1}{\sqrt{n}}\sum_{t=1}^{n}H_t\right) \\
&= Var(H_1)+2\sum_{j=1}^{n-1}\left(1-\frac{j}{n}\right)Cov(H_1,H_{j+1}),
\end{aligned}
$$

where

$$
Var(H_1)\le E(H_1)^2 = E\left[I(M_{1,1}<u_1<M_{2,1})\right]=O_p(v_n^{-1})=o_p(1),
$$

$$\begin{aligned}
\sum_{j=1}^{n-1}\left(1-\frac{j}{n}\right)Cov(H_1,H_{j+1}) &\leq \sum_{j=1}^{n-1}\left(1-\frac{j}{n}\right)|Cov(H_1,H_{j+1})| \\
&\leq \sum_{j=1}^{\lambda_n-1}|Cov(H_1,H_{j+1})|+\sum_{j=\lambda_n}^{n-1}|Cov(H_1,H_{j+1})| \\
&:= \mathcal{J}_1+\mathcal{J}_2.
\end{aligned}$$

For $\mathcal{J}_1$, apply the Cauchy-Schwarz inequality,

$$\begin{aligned}
|Cov(H_1,H_{j+1})| &= |E(H_1H_{j+1})| \\
&\leq \sqrt{E|H_1|^2}\sqrt{E|H_{j+1}|^2} \\
&= \sqrt{E\left[I(M_{1,1} \ < \ u_t \ < \ M_{2,1})\right]}\sqrt{E\left[I(M_{1,j+1} \ < \ u_{j+1} \ < \ M_{2,j+1})\right]} \\
&= O_p(v_n^{-1}),
\end{aligned}$$

where in the first equality, we use $E(H_1)=E\left(F\left(F^{-1}(\tau)\right)x_1'\left(\hat{\beta}(z)-\beta(z)\right)\right)=o_p(1)$. Thus $\mathcal{J}_1=O_p(\lambda_n v_n^{-1})$.

For $\mathcal{J}_2$, apply the assumption $3(ii)$ and covariance inequalities for mixing sequences,

$$\begin{aligned}
\mathcal{J}_2 &= \sum_{j=\lambda_n}^{n-1}|Cov(H_1,H_{j+1})| \\
&\leq \sum_{j=\lambda_n}^{n-1}C\alpha^{1-\delta/2}(j)\{E|H_1|^\delta\}^{1/\delta}\{E|H_{j+1}|^\delta\}^{1/\delta} \\
&= Cv_n^{-2/\delta}\sum_{j=\lambda_n}^{n-1}\alpha^{1-\delta/2}(j) \\
&\leq Cv_n^{-1}(v_n^{1-2/\delta}\lambda_n^{-\gamma})\sum_{j=\lambda_n}^{n-1}j^{-\gamma}\alpha^{1-\delta/2}(j),
\end{aligned}$$

where $\delta\geq 2,\gamma\geq 1-\delta/2$. Through choosing an appropriate λ_n such that $v_n^{1-2/\delta}\lambda_n^{-\gamma}=C,C$ is a positive constant, then $\lambda_n v_n^{-1}\to 0$.

As a result, we have

$$\mathcal{J}_1=O_p(\lambda_n v_n^{-1})=o_p(1),$$
$$\mathcal{J}_2=O_p(v_n^{-1})=o_p(1),$$
$$Var(\mathcal{N}_2)=o_p(1)+\mathcal{J}_1+\mathcal{J}_2=o_p(1).$$

Therefore,

$$\begin{aligned}
\mathcal{N}_2 &\xrightarrow{p} E(\mathcal{N}_2|\mathcal{F}_t) \\
&\Rightarrow \left(\int B_v'(s)ds\right)\left(\int B_v(s)B_v'(s)ds\right)^{-1}\left(\int B_v(s)B_\psi^k(s)ds+\lambda_{v\psi}\right),
\end{aligned}$$

and then

$$QCS(\tau) \Rightarrow -\left(\int B_v'(s)ds\right)\left(\int B_v(s)B_v'(s)ds\right)^{-1}\left(\int B_v(s)B_\psi^k(s)ds + \lambda_{v\psi}\right).$$

Define the modified quantile residuals $\hat{u}_{t\tau}^m = y_t - \beta^m(z)x_t$, then the modified statistics can be written as

$$\begin{aligned}
MQCS(\tau) &= \sqrt{\frac{h}{n}}\sum_{t=1}^{\lfloor nr\rfloor}\psi_\tau(\hat{u}_{t\tau}^m) \\
&= \sqrt{\frac{h}{n}}\sum_{t=1}^{\lfloor nr\rfloor}\psi_\tau(u_{t\tau}) - \sqrt{\frac{h}{n}}\sum_{t=1}^{\lfloor nr\rfloor}[\psi_\tau(u_{t\tau}) - \psi_\tau(\hat{u}_{t\tau}^m)] \\
&:= \mathcal{M}_1 - \mathcal{M}_2.
\end{aligned}$$

Using the same argument as proving $\mathcal{N}_1 \to o_p(1).$, we have $\mathcal{M}_1 \to o_p(1)$.

Next we consider the asymptotic behavior of $\mathcal{M}_2$. Similar to the proof of $\mathcal{N}_2$,

$$\begin{aligned}
\mathcal{M}_2 &\xrightarrow{p} E(\mathcal{M}_2|\mathcal{F}_t) \\
&= \frac{\sqrt{h}}{\sqrt{n}}\sum_{t=1}^{\lfloor nr\rfloor}E\left[\psi_\tau(u_{t\tau}) - \psi_\tau(\hat{u}_{t\tau}^m)\right] \\
&= \frac{\sqrt{h}}{\sqrt{n}}\sum_{t=1}^{\lfloor nr\rfloor}F_t\left[F^{-1}(\tau) + x_t'\left(\hat{\beta}^m(z) - \beta(z)\right)\right] - F_t\left[F^{-1}(\tau)\right] \\
&= F\left(F^{-1}(\tau)\right)\frac{1}{n\sqrt{n}}\sum_{t=1}^{\lfloor nr\rfloor}x_t'n\sqrt{h}\left(\hat{\beta}^m(z) - \beta(z)\right) + o_p(1) \\
&\Rightarrow \left(\int B_v'(s)ds\right)\left(\int B_v(s)B_v'(s)ds\right)^{-1}\left(\int B_v(s)B_{\psi\cdot v}^k(s)ds\right)
\end{aligned}$$

Finally, re-standardize both $B_{\psi\cdot v}^k(r)$ and $B_v(r)$. Let $W_{\psi\cdot v}^k(r) = \hat{\omega}_{\psi\cdot v}^{-1}B_{\psi\cdot v}^k(r), \hat{\Omega}_{vv}^{-1/2}W_v(r) = B_v(r)$, then

$$\sqrt{\frac{h}{n}}\sum_{t=1}^{\lfloor nr\rfloor}\psi_\tau(\hat{u}_{t\tau}^m) \Rightarrow -\hat{\omega}_{\psi\cdot v}^{-1}\left(\int W_v'(s)ds\right)\left(\int W_v(s)W_v'(s)ds\right)^{-1}\left(\int W_v(s)W_{\psi\cdot v}^k(s)ds\right)$$

This concludes the proof.

Appendix B Supplementary Simulation Results

Table 10: Empirical size and power ($n = 200, \rho = 0.5, M = M_1.$)

	MQCS				*SupY*				$\mathcal{T}_n$				*CUSUM*			
	$N(0,1)$	$t(2)$	$t(4)$	$N(0,\sigma_t)$	$N(0,1)$	$t(2)$	$t(4)$	$N(0,\sigma_t)$	$N(0,1)$	$t(2)$	$t(4)$	$N(0,\sigma_t)$	$N(0,1)$	$t(2)$	$t(4)$	$N(0,\sigma_t)$
								$\beta(u_t, z_t) = 1$								
$\gamma = 0.0$	0.041	0.092	0.009	0.024	0.030	0.028	0.022	0.035	0.059	0.050	0.036	0.035	0.037	0.030	0.028	0.033
$\gamma = 0.2$	0.075	0.111	0.027	0.073	0.031	0.033	0.037	0.045	0.064	0.074	0.035	0.062	0.031	0.035	0.039	0.033
$\gamma = 0.4$	0.122	0.153	0.046	0.091	0.049	0.040	0.040	0.055	0.078	0.068	0.088	0.062	0.049	0.030	0.041	0.040
$\gamma = 0.6$	0.144	0.219	0.077	0.109	0.091	0.074	0.106	0.083	0.125	0.105	0.159	0.123	0.054	0.061	0.059	0.056
$\gamma = 0.8$	0.129	0.169	0.091	0.107	0.107	0.133	0.119	0.129	0.213	0.204	0.307	0.282	0.158	0.166	0.179	0.146
$\gamma = 1.0$	0.954	0.954	0.926	0.938	0.120	0.152	0.146	0.134	0.654	0.652	0.658	0.588	0.854	0.786	0.834	0.848
								$\beta(u_t, z_t) = sin\left(\sqrt{2}\pi z_t\right)$								
$\gamma = 0.0$	0.054	0.059	0.081	0.060	0.038	0.041	0.042	0.036	0.043	0.044	0.039	0.031	0.030	0.031	0.037	0.023
$\gamma = 0.2$	0.058	0.092	0.096	0.089	0.037	0.035	0.035	0.041	0.043	0.046	0.031	0.063	0.031	0.030	0.032	0.042
$\gamma = 0.4$	0.098	0.093	0.123	0.081	0.036	0.041	0.051	0.058	0.078	0.066	0.067	0.083	0.041	0.062	0.047	0.058
$\gamma = 0.6$	0.204	0.179	0.190	0.211	0.054	0.034	0.050	0.058	0.110	0.106	0.163	0.080	0.076	0.062	0.075	0.080
$\gamma = 0.8$	0.343	0.377	0.315	0.368	0.038	0.038	0.061	0.032	0.251	0.214	0.275	0.251	0.1891	0.181	0.183	0.150
$\gamma = 1.0$	0.958	0.964	0.954	0.930	0.072	0.062	0.060	0.102	0.614	0.672	0.638	0.666	0.884	0.814	0.860	0.852
								$\beta(u_t, z_t) = [0.5 + \Phi(u_t)]cos\left(\sqrt{2}\pi z_t\right)$								
$\gamma = 0.0$	0.040	0.061	0.033	0.039	0.057	0.043	0.047	0.038	0.451	0.402	0.466	0.448	0.119	0.443	0.173	0.124
$\gamma = 0.2$	0.042	0.040	0.042	0.065	0.057	0.048	0.043	0.050	0.430	0.455	0.493	0.467	0.206	0.558	0.251	0.227
$\gamma = 0.4$	0.060	0.060	0.064	0.073	0.054	0.063	0.057	0.053	0.453	0.449	0.506	0.432	0.360	0.646	0.395	0.348
$\gamma = 0.6$	0.100	0.115	0.093	0.079	0.060	0.042	0.065	0.060	0.413	0.449	0.510	0.451	0.563	0.703	0.563	0.575
$\gamma = 0.8$	0.173	0.167	0.168	0.185	0.059	0.051	0.072	0.077	0.446	0.438	0.535	0.470	0.817	0.879	0.845	0.815
$\gamma = 1.0$	0.786	0.748	0.808	0.842	0.082	0.072	0.076	0.066	0.700	0.698	0.700	0.620	0.988	0.996	0.996	0.996

Note: The *MQCS* is our proposed test, *SupY*, $\mathcal{T}_n$ and *CUSUM* are the tests proposed by Xiao (2009b), Xiao (2009a) and Xiao and Phillips (2002), respectively.

Table 11: Empirical size and power ($n = 200, \rho = 0.8, M = M_1$.)

	MQCS				*SupY*				$\mathcal{T}_n$				*CUSUM*			
	$N(0,1)$	$t(2)$	$t(4)$	$N(0,\sigma_t)$	$N(0,1)$	$t(2)$	$t(4)$	$N(0,\sigma_t)$	$N(0,1)$	$t(2)$	$t(4)$	$N(0,\sigma_t)$	$N(0,1)$	$t(2)$	$t(4)$	$N(0,\sigma_t)$
$\beta(u_t, z_t) = 1$																
$\gamma = 0.0$	0.047	0.010	0.064	0.077	0.038	0.030	0.035	0.035	0.072	0.038	0.033	0.073	0.026	0.037	0.029	0.028
$\gamma = 0.2$	0.060	0.0025	0.078	0.100	0.058	0.047	0.045	0.034	0.079	0.040	0.058	0.070	0.037	0.040	0.048	0.039
$\gamma = 0.4$	0.112	0.045	0.133	0.140	0.059	0.047	0.054	0.049	0.089	0.074	0.066	0.099	0.033	0.048	0.048	0.033
$\gamma = 0.6$	0.180	0.067	0.137	0.177	0.106	0.090	0.086	0.101	0.142	0.153	0.127	0.109	0.033	0.072	0.067	0.052
$\gamma = 0.8$	0.181	0.094	0.151	0.239	0.183	0.185	0.156	0.123	0.227	0.291	0.242	0.172	0.070	0.167	0.163	0.066
$\gamma = 1.0$	0.912	0.966	0.964	0.856	0.128	0.090	0.106	0.078	0.602	0.660	0.644	0.650	0.638	0.710	0.764	0.540
$\beta(u_t, z_t) = sin\left(\sqrt{2}\pi z_t\right)$																
$\gamma = 0.0$	0.071	0.110	0.097	0.074	0.041	0.045	0.035	0.044	0.051	0.032	0.040	0.061	0.027	0.034	0.032	0.029
$\gamma = 0.2$	0.060	0.129	0.093	0.133	0.043	0.030	0.044	0.042	0.076	0.037	0.072	0.055	0.023	0.039	0.043	0.024
$\gamma = 0.4$	0.118	0.176	0.128	0.111	0.039	0.023	0.031	0.049	0.091	0.076	0.071	0.083	0.032	0.035	0.035	0.032
$\gamma = 0.6$	0.159	0.185	0.184	0.188	0.055	0.023	0.057	0.047	0.111	0.157	0.132	0.102	0.042	0.074	0.063	0.050
$\gamma = 0.8$	0.320	0.421	0.354	0.305	0.049	0.053	0.038	0.060	0.229	0.298	0.220	0.171	0.079	0.168	0.153	0.078
$\gamma = 1.0$	0.894	0.964	0.974	0.826	0.066	0.062	0.074	0.076	0.588	0.704	0.646	0.666	0.688	0.754	0.790	0.578
$\beta(u_t, z_t) = [0.5 + \Phi(u_t)]cos\left(\sqrt{2}\pi z_t\right)$																
$\gamma = 0.0$	0.050	0.045	0.050	0.056	0.046	0.048	0.049	0.043	0.498	0.503	0.470	0.530	0.100	0.483	0.458	0.333
$\gamma = 0.2$	0.066	0.050	0.062	0.071	0.050	0.041	0.053	0.037	0.540	0.517	0.459	0.543	0.190	0.561	0.530	0.445
$\gamma = 0.4$	0.085	0.063	0.077	0.066	0.041	0.066	0.040	0.050	0.561	0.494	0.435	0.564	0.323	0.620	0.605	0.511
$\gamma = 0.6$	0.088	0.084	0.100	0.093	0.060	0.066	0.051	0.063	0.542	0.527	0.482	0.515	0.500	0.717	0.729	0.648
$\gamma = 0.8$	0.161	0.160	0.179	0.172	0.057	0.054	0.066	0.063	0.512	0.520	0.454	0.526	0.764	0.900	0.872	0.839
$\gamma = 1.0$	0.6668	0.722	0.774	0.558	0.052	0.044	0.060	0.060	0.698	0.724	0.724	0.714	0.978	1.000	1.000	0.922

Note: The *MQCS* is our proposed test, *SupY*, $\mathcal{T}_n$ and *CUSUM* are the tests proposed by Xiao (2009b), Xiao (2009a) and Xiao and Phillips (2002), respectively.

Table 12: Empirical size and power ($n = 200, \rho = 0.5, M = M_3$.)

	MQCS				*SupY*				$\mathcal{T}_n$				*CUSUM*			
	$N(0,1)$	$t(2)$	$t(4)$	$N(0,\sigma_t)$	$N(0,1)$	$t(2)$	$t(4)$	$N(0,\sigma_t)$	$N(0,1)$	$t(2)$	$t(4)$	$N(0,\sigma_t)$	$N(0,1)$	$t(2)$	$t(4)$	$N(0,\sigma_t)$
$\beta(u_t, z_t) = 1$																
$\gamma = 0.0$	0.042	0.100	0.012	0.042	0.042	0.020	0.026	0.028	0.005	0.002	0.009	0.006	0.037	0.032	0.027	0.034
$\gamma = 0.2$	0.057	0.128	0.021	0.048	0.045	0.027	0.025	0.051	0.006	0.002	0.011	0.003	0.031	0.035	0.039	0.033
$\gamma = 0.4$	0.066	0.118	0.039	0.049	0.032	0.053	0.034	0.051	0.002	0.001	0.020	0.010	0.030	0.038	0.048	0.048
$\gamma = 0.6$	0.098	0.129	0.031	0.073	0.056	0.091	0.051	0.070	0.006	0.004	0.021	0.011	0.056	0.064	0.057	0.042
$\gamma = 0.8$	0.080	0.101	0.063	0.080	0.070	0.063	0.061	0.069	0.016	0.014	0.054	0.036	0.097	0.112	0.081	0.105
$\gamma = 1.0$	0.764	0.740	0.772	0.738	0.064	0.056	0.076	0.082	0.278	0.356	0.386	0.328	0.672	0.646	0.718	0.696
$\beta(u_t, z_t) = sin\left(\sqrt{2}\pi z_t\right)$																
$\gamma = 0.0$	0.029	0.038	0.038	0.044	0.037	0.038	0.032	0.047	0.009	0.005	0.029	0.010	0.038	0.032	0.039	0.039
$\gamma = 0.2$	0.028	0.035	0.058	0.047	0.044	0.043	0.034	0.044	0.006	0.012	0.020	0.006	0.030	0.040	0.033	0.036
$\gamma = 0.4$	0.056	0.044	0.064	0.053	0.061	0.050	0.030	0.041	0.006	0.005	0.016	0.016	0.036	0.040	0.023	0.036
$\gamma = 0.6$	0.075	0.059	0.080	0.089	0.042	0.032	0.046	0.056	0.012	0.005	0.037	0.013	0.047	0.045	0.062	0.049
$\gamma = 0.8$	0.156	0.127	0.193	0.157	0.057	0.040	0.039	0.062	0.022	0.014	0.062	0.032	0.114	0.084	0.106	0.102
$\gamma = 1.0$	0.800	0.820	0.796	0.808	0.076	0.088	0.072	0.042	0.358	0.344	0.348	0.312	0.776	0.668	0.730	0.738
$\beta(u_t, z_t) = [0.5 + \Phi(u_t)]cos\left(\sqrt{2}\pi z_t\right)$																
$\gamma = 0.0$	0.033	0.039	0.019	0.042	0.049	0.049	0.043	0.046	0.309	0.287	0.359	0.333	0.118	0.395	0.180	0.124
$\gamma = 0.2$	0.038	0.050	0.030	0.037	0.054	0.065	0.038	0.051	0.279	0.320	0.366	0.314	0.143	0.420	0.191	0.149
$\gamma = 0.4$	0.043	0.074	0.029	0.035	0.053	0.044	0.075	0.066	0.315	0.279	0.360	0.318	0.157	0.393	0.189	0.149
$\gamma = 0.6$	0.062	0.044	0.062	0.057	0.064	0.064	0.048	0.046	0.259	0.270	0.333	0.294	0.195	0.386	0.203	0.176
$\gamma = 0.8$	0.079	0.086	0.088	0.106	0.074	0.071	0.056	0.070	0.219	0.220	0.310	0.250	0.291	0.397	0.323	0.325
$\gamma = 1.0$	0.642	0.592	0.626	0.646	0.078	0.044	0.082	0.066	0.426	0.458	0.462	0.434	0.818	0.816	0.862	0.830

Note: The *MQCS* is our proposed test, *SupY*, $\mathcal{T}_n$ and *CUSUM* are the tests proposed by Xiao (2009b), Xiao (2009a) and Xiao and Phillips (2002), respectively.

Table 13: Empirical size and power ($n = 200, \rho = 0.8, M = M_3.$)

	MQCS				*SupY*				$\mathcal{T}_n$				*CUSUM*			
	$N(0,1)$	$t(2)$	$t(4)$	$N(0,\sigma_t)$	$N(0,1)$	$t(2)$	$t(4)$	$N(0,\sigma_t)$	$N(0,1)$	$t(2)$	$t(4)$	$N(0,\sigma_t)$	$N(0,1)$	$t(2)$	$t(4)$	$N(0,\sigma_t)$
								$\beta(u_t, z_t) = 1$								
$\gamma = 0.0$	0.044	0.015	0.051	0.072	0.054	0.032	0.028	0.049	0.011	0.016	0.003	0.005	0.029	0.031	0.036	0.034
$\gamma = 0.2$	0.061	0.017	0.117	0.109	0.071	0.036	0.042	0.022	0.005	0.013	0.006	0.009	0.027	0.039	0.033	0.042
$\gamma = 0.4$	0.068	0.029	0.133	0.172	0.045	0.035	0.042	0.029	0.012	0.021	0.011	0.004	0.029	0.027	0.044	0.035
$\gamma = 0.6$	0.083	0.050	0.077	0.134	0.050	0.062	0.070	0.059	0.010	0.024	0.008	0.004	0.053	0.030	0.064	0.038
$\gamma = 0.8$	0.102	0.072	0.103	0.129	0.068	0.078	0.073	0.083	0.028	0.059	0.023	0.019	0.049	0.097	0.086	0.055
$\gamma = 1.0$	0.700	0.704	0.103	0.129	0.034	0.106	0.088	0.098	0.328	0.358	0.332	0.380	0.560	0.690	0.668	0.055
								$\beta(u_t, z_t) = sin\left(\sqrt{2}\pi z_t\right)$								
$\gamma = 0.0$	0.061	0.046	0.043	0.044	0.037	0.048	0.038	0.040	0.031	0.024	0.012	0.025	0.033	0.040	0.031	0.029
$\gamma = 0.2$	0.065	0.066	0.061	0.069	0.040	0.063	0.040	0.021	0.033	0.022	0.014	0.026	0.037	0.038	0.025	0.038
$\gamma = 0.4$	0.046	0.068	0.041	0.072	0.049	0.050	0.042	0.039	0.021	0.018	0.013	0.017	0.027	0.043	0.040	0.030
$\gamma = 0.6$	0.078	0.108	0.087	0.072	0.041	0.037	0.040	0.056	0.024	0.033	0.012	0.017	0.042	0.057	0.045	0.036
$\gamma = 0.8$	0.162	0.175	0.141	0.137	0.067	0.054	0.056	0.053	0.029	0.066	0.032	0.025	0.057	0.089	0.107	0.047
$\gamma = 1.0$	0.784	0.834	0.814	0.670	0.068	0.064	0.080	0.062	0.340	0.346	0.356	0.384	0.630	0.660	0.666	0.522
								$\beta(u_t, z_t) = [0.5 + \Phi(u_t)]cos\left(\sqrt{2}\pi z_t\right)$								
$\gamma = 0.0$	0.037	0.040	0.033	0.038	0.056	0.046	0.063	0.071	0.390	0.369	0.338	0.380	0.103	0.456	0.444	0.325
$\gamma = 0.2$	0.024	0.048	0.040	0.041	0.060	0.058	0.060	0.049	0.405	0.356	0.329	0.412	0.144	0.472	0.424	0.388
$\gamma = 0.4$	0.057	0.051	0.050	0.0656	0.052	0.0552	0.059	0.073	0.370	0.374	0.302	0.379	0.137	0.421	0.355	0.391
$\gamma = 0.6$	0.053	0.078	0.067	0.057	0.053	0.058	0.047	0.078	0.373	0.355	0.299	0.391	0.177	0.402	0.355	0.346
$\gamma = 0.8$	0.104	0.115	0.107	0.095	0.056	0.050	0.046	0.046	0.334	0.322	0.259	0.331	0.281	0.414	0.384	0.388
$\gamma = 1.0$	0.510	0.612	0.598	0.386	0.090	0.064	0.046	0.028	0.474	0.518	0.492	0.528	0.820	0.844	0.822	0.768

Note: The *MQCS* is our proposed test, *SupY*, $\mathcal{T}_n$ and *CUSUM* are the tests proposed by Xiao (2009b), Xiao (2009a) and Xiao and Phillips (2002), respectively.

Table 14: Empirical size and power of MQCS, ($n = 200, M = M_2, p = 2.$).

	$\rho = 0.5$				$\rho = 0.8$			
	$N(0,1)$	$t(2)$	$t(4)$	$N(0,\sigma_t)$	$N(0,1)$	$t(2)$	$t(4)$	$N(0,\sigma_t)$
				$\beta(u_t, z_t) = 1$				
$\gamma = 0.0$	0.036	0.020	0.064	0.042	0.060	0.028	0.068	0.056
$\gamma = 0.2$	0.048	0.042	0.070	0.062	0.060	0.030	0.074	0.106
$\gamma = 0.4$	0.074	0.048	0.074	0.080	0.088	0.034	0.108	0.070
$\gamma = 0.6$	0.120	0.046	0.092	0.076	0.114	0.038	0.100	0.108
$\gamma = 0.8$	0.084	0.044	0.122	0.176	0.104	0.064	0.092	0.122
$\gamma = 1.0$	0.892	0.790	0.854	0.830	0.872	0.832	0.860	0.694
				$\beta(u_t, z_t) = sin\left(\sqrt{2}\pi z_t\right)$				
$\gamma = 0.0$	0.040	0.042	0.046	0.032	0.028	0.050	0.018	0.048
$\gamma = 0.2$	0.036	0.046	0.030	0.058	0.032	0.046	0.034	0.036
$\gamma = 0.4$	0.066	0.076	0.046	0.044	0.060	0.084	0.068	0.060
$\gamma = 0.6$	0.086	0.070	0.086	0.078	0.060	0.066	0.100	0.090
$\gamma = 0.8$	0.202	0.178	0.198	0.220	0.154	0.200	0.188	0.210
$\gamma = 1.0$	0.848	0.886	0.874	0.784	0.898	0.882	0.870	0.722
				$\beta(u_t, z_t) = [0.5 + \Phi(u_t)]cos\left(\sqrt{2}\pi z_t\right)$				
$\gamma = 0.0$	0.066	0.062	0.072	0.072	0.070	0.068	0.076	0.056
$\gamma = 0.2$	0.072	0.048	0.054	0.048	0.066	0.066	0.064	0.060
$\gamma = 0.4$	0.064	0.060	0.068	0.080	0.094	0.060	0.064	0.072
$\gamma = 0.6$	0.082	0.074	0.086	0.072	0.092	0.064	0.116	0.090
$\gamma = 0.8$	0.106	0.116	0.082	0.102	0.132	0.100	0.116	0.186
$\gamma = 1.0$	0.828	0.822	0.844	0.716	0.832	0.852	0.856	0.672

Note: The $MQCS$ is our proposed test. We use nonparametric local polynomial estimator here (i.e. p=2).