Journal of Econometrics 152 (2009) 81–92

Contents lists available at ScienceDirect

Journal of Econometrics

journal homepage: www.elsevier.com/locate/jeconom

Functional-coefficient cointegration models

Zhijie Xiao *

Department of Economics, Boston College, USA
Tsinghua University, China

ARTICLE INFO

Article history:
Available online 16 January 2009

JEL classification:
C22
G1

Keywords:
Cointegration
Local polynomial
Nonparametric
Time varying
Functional coefficients

ABSTRACT

This paper studies estimation and inference of functional coefficient cointegration models. The proposed model offers a more flexible structure of cointegration where the value of cointegrating coefficients may be affected by informative covariates and thus may vary over time. The model may be viewed as a stochastic cointegration model and includes the conventional cointegration model as a special case. The proposed new model provides a useful complement to the conventional fixed coefficient cointegration models. Both kernel and local polynomial estimators are investigated. Inference procedures for instability of cointegrating parameters and a test for cointegration are proposed based on the functional-coefficient estimates. Limiting distributions of the estimates and testing statistics are derived.
© 2009 Elsevier B.V. All rights reserved.

1. Introduction

Since Granger (1981) and Engle and Granger (1987), cointegration models have attracted a huge amount of research attention in econometrics. While the concept of cointegration provides an attractive and appealing characterization for long term equilibrium relationships, less evidence of cointegration is found in empirical applications. One reason for this empirical consequence is probably due to parameter instability. In the traditional cointegration model of Engle and Granger (1987), the cointegrating parameters are constants. However, cointegration characterizes long-term equilibrium relationships, and the exact quantitative relationship among economic variables may vary over a long time horizon. Many financial and economic applications suggest that the value of cointegrating vector might be time-varying. For example, application of cointegration in investment analysis shows that frequent rebalancing is necessary to keep the portfolio in line with the index, indicating the value of cointegrating vector may be changing over time.

Many empirical studies on asset pricing consider the rational expectations model for stock prices

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

which is a first-order expectational difference equation, where P_t is the real stock price at time t, γ is the real rate of return, and D_t is the dividend. In empirical analyses, dividends (or other fundamental variables) are usually characterized as integrated processes ($I(1)$). A forward-looking solution to the above equation suggests that stock prices and market fundamentals should be cointegrated in the sense that a linear combination between P_t and D_t is stationary, leading to the following cointegrating regression model:

$$P_t = \alpha + \beta' D_t + u_t.$$

Based on such a cointegration relationship, there is a large collection of empirical study on asset pricing. [see, inter alia, Campbell and Shiller (1987), Diba and Grossman (1988), Evans (1991), Campbell et al. (1997), Cerchi and Havenner (1988), Chowdhury (1991) and Hendry (1996)] A general conclusion of empirical studies is: cointegration relationships can not be found from these time series. Although the present value model suggests that asset prices are cointegrated with market fundamentals, empirically it is well known that stock prices are much more volatile than market fundamentals. Again, a plausible source of the additional volatility may come from time-varying cointegrating vectors.

In this paper, we attempt to extend the traditional cointegration model

$$y_t = \beta' x_t + u_t. \tag{2}$$

where x_t is an integrated process ($I(1)$), to a more general class of cointegration model

$$y_t = \beta(z_t)' x_t + u_t, \tag{3}$$

in which the cointegrating coefficients β are allowed to be varying. In particular, we wish to consider cointegrating regression

* Corresponding address: Department of Economics, Boston College, Chestnut Hill, MA 02467, USA. Tel.: +1 617 552 1709.
E-mail address: xiaoz@bc.edu.

0304-4076/$ – see front matter © 2009 Elsevier B.V. All rights reserved.
doi:10.1016/j.jeconom.2009.01.008

where the value of cointegrating coefficients $\beta = \beta(z_t)$ may be affected by covariates (z_t) containing useful information of market/macroeconomic conditions. The proposed model may be viewed as a stochastic cointegration model which includes the conventional cointegration as a special case.

We focus our analysis on kernel estimators but limiting distributions of both kernel and local polynomial estimators of $\beta(z)$ are given. We show that the functional cointegrating coefficients are (nonparametric) super-consistent in the sense that it converges at rate $n\sqrt{h}$ (instead of the conventional nonparametric rate $\sqrt{nh}$) to the true values. In addition, the limiting distribution of the functional coefficient estimates is mixture of normal. The joint distribution of the estimators at different points of z is also developed. An augmented regression is proposed to accommodate the possible endogeneity problem caused by correlation between the regressors and the error term. The functional coefficient cointegration models lead to an interesting inference apparatus for time series. We study the inference for parameter instability based on the proposed model; a simple new testing procedure for cointegration (with either fixed or varying coefficients) against hypothesis of no cointegration (allowing for possible varying coefficients) is also proposed.

Related to this paper there is a large literature on cointegration models with structural instabilities (e.g. Hansen (1992), Quintos and Phillips (1993), and Qu (2005)), stochastic cointegration (e.g. Harris et al. (2002)). Both Park and Hahn (1999) and Xiao (2009) studied cointegrating regression models with varying coefficients. In particular, Park and Hahn (1999) considered a model where the cointegrating vector is a deterministic function of time t. Xiao (2009) studied cointegration models where the value of cointegrating parameters may vary over quantiles of the conditional distribution. In the unit root setting, Juhl (2005) studied functional coefficient estimation for an unit root autoregression model where the true autoregression parameter is unity (and thus not varying).

The rest of this paper is organized as follows: we study functional coefficient regressions with nonstationary $(I(1))$ regressors in Sections 2 and 3. In particular, limiting distributions of the nonparametric kernel estimators are derived in Section 2. Inference procedures on both parameter instability and the hypothesis of cointegration are proposed in Section 3. Extensions are studied in Section 4. We extend our analysis to the general model that includes an intercept term and allows for endogeneity, and local polynomial estimators are also briefly discussed in Section 4. In particular, we augment the regression by including leads and lags of the differenced regressors to remove the endogeneity. Summary of a limited Monte Carlo and conclusion is reported in Section 5, Proofs are contained in the Appendix.

In matters of notation, we use "$\Rightarrow$" to signify weak convergence of the associated probability measures, $[nr]$ to signify the integer part of nr, $:=$ to signify definitional equality, and $I(k)$ to denote integration of order k. Continuous stochastic process such as the Brownian motion $B(r)$ on $[0, 1]$ are usually written simply as B and integrals $\int$ are understood to be taken over the interval $[0, 1]$, unless otherwise specified. For a kernel function K, we define $\mu_j(K) = \int u^j K(u)\mathrm{d}u$, $\nu_j(K) = \int u^j K^2(u)\mathrm{d}u$, and matrices:

$$M(K) = \begin{bmatrix} 1 & \mu_1(K) & \cdots & \mu_P(K) \\ \mu_1(K) & \mu_2(K) & & \mu_{P+1}(K) \\ \cdots & & \ddots & \\ \mu_P(K) & \mu_{P+1}(K) & & \mu_{2P}(K) \end{bmatrix},$$

$$B(K) = \begin{bmatrix} \mu_{P+1}(K) \\ \mu_{P+2}(K) \\ \cdots \\ \mu_{2P+1}(K) \end{bmatrix},$$

$$N(K) = \begin{bmatrix} \nu_0(K) & \nu_1(K) & \cdots & \nu_P(K) \\ \nu_1(K) & \nu_2(K) & & \nu_{P+1}(K) \\ \cdots & & \ddots & \cdots \\ \nu_P(K) & \nu_{P+1}(K) & \cdots & \nu_{2P}(K) \end{bmatrix}.$$

2. Functional-coefficient regression with I(1) regressors

In this section, we study functional coefficient regression (3) where x_t is a k-dimensional I(1) process, z_t is an I(0) process, and the coefficients $\beta = \beta(z_t)$ are smooth functions of z_t. Without loss of generality and for simplicity of illustration, we assume in this paper that z_t is a univariate random variable. Extension to multivariate z involves no fundamentally new ideas. We are interested in the estimation of $\beta(z)$ in this section. Inference on model (3) will be studied in Section 3. Both kernel estimators and local polynomial estimators can be analyzed. We focus our attention on the kernel estimator in this paper, and briefly describe the results of local polynomial estimation in the end of Section 4.

For the purpose of asymptotic analysis, we make the following assumptions on dependence of the time series, the kernel function, and bandwidth.

Assumption A. Let $v_t = \Delta x_t$, v_t and z_t are a zero-mean stationary processes and are independent with the error term u_t. In addition, the process $Y_t = \{u_t, v_t, z_t\}$ is geometrically absolute regular (β-mixing) with finite fourth moment and a nonsingular long-run variance matrix.

Assumption B. z_t has Lebesgue density $f_z(\cdot)$ with bounded uniformly continuous partial derivatives up to the order q. In addition, $\beta(z)$ is uniformly continuous up to the order q.

Assumption K. The kernel K has support $[-1, 1]$ and is symmetric about zero and satisfies $\int K(u)\mathrm{d}u = 1$. In addition, $\int u^j K(u)\mathrm{d}u = 0, j = 1, \ldots, q-1$, and $\int u^q K(u)\mathrm{d}u \neq 0$.

Assumption W. As $n \to \infty$, $h \to 0$, and $nh \to \infty$.

Assumption A ensures that the partial sums of u_t and v_t satisfy the following functional central limiting theorems:

$$\frac{1}{\sqrt{n}} \sum_{t=1}^{[nr]} u_t \Rightarrow B_1(r) = BM(\omega_1^2),$$

$$\frac{1}{\sqrt{n}} \sum_{t=1}^{[nr]} v_t \Rightarrow B_2(r) = BM(\Omega_{22}),$$

and $B_1(r)$ and $B_2(r)$ are independent Brownian motions with variance and covariance matrix ω_1^2 and Ω_{22}. It also ensures that the covariance matrix Ω_{22} is nonsingular. Assumption B concerns about the smoothness of the coefficient function and density and ensures a Taylor expansion to appropriate order. Assumption K for the kernel function and Assumption W for the bandwidth expansion are quite standard in nonparametric estimation. In addition with Assumption A, the kernel weighted partial sum ($\underline{K}_t u_t$ below) converges to $B_3(r)$, a Brownian motion independent of $B_1(r)$ and $B_2(r)$.

Due to nonparametric smoothing, we need to analyze the behavior of kernel weighted time series and denote

$$\underline{K}_t = K\left(\frac{z_t - z}{h}\right) - \mathrm{E}K\left(\frac{z_t - z}{h}\right).$$

Lemma 1. *Under Assumptions A, B, K and W, we have the following multivariate invariance principle:*

$$\begin{pmatrix} \frac{1}{\sqrt{n}} \sum_{t=1}^{[nr]} u_t \\ \frac{1}{\sqrt{n}} \sum_{t=1}^{[nr]} v_t \\ \frac{1}{\sqrt{nh}} \sum_{t=1}^{[nr]} \underline{K}_t u_t \end{pmatrix} \Rightarrow \begin{pmatrix} B_1(r) \\ B_2(r) \\ B_3(r) \end{pmatrix},$$

where $(B_1(r), B_2'(r), B_3(r))'$ *is a* $(k+2)$*-dimensional Brownian motion with covariance matrix*

$$\begin{bmatrix} \omega_1^2 & 0 & 0 \\ 0 & \Omega_{22} & 0 \\ 0 & 0 & \nu_0(K) f_z(z)\sigma_u^2 \end{bmatrix}$$

where $\nu_0(K) = \int K(u)^2 \, du$.

Remark 1. Notice that $\omega_1^2 = \sigma_u^2 + 2\sum_{h=1}^{\infty} \mathrm{E}(u_t u_{t+h})$, although u_t is weakly dependent and its partial sum process converges to a Brownian Motion with variance equals to the *long-run variance* of u_t (ω_1^2), the kernel weighted sequence $\underline{K}_t u_t$ behaves asymptotically uncorrelated and thus its partial sum process converges to a Brownian Motion with variance proportional to the variance of u_t (σ_u^2).

The kernel estimator of $\beta(z)$ is determined by the following optimization problem:

$$\widehat{\beta}(z) = \arg\min_{\beta} \sum_{t=1}^{n} K\left(\frac{z_t - z}{h}\right) \{y_t - x_t'\beta\}^2$$

giving

$$\widehat{\beta}(z) = \left[\frac{1}{nh}\sum_{t=1}^{n} x_t x_t' K\left(\frac{z_t - z}{h}\right)\right]^{-1} \left[\frac{1}{nh}\sum_{t=1}^{n} x_t y_t K\left(\frac{z_t - z}{h}\right)\right].$$

The asymptotic behavior of the nonparametric estimator is summarized in the following Theorem. The proof of this result is given in the Appendix.

Theorem 1. *Under Assumptions A, B, K and W,*

$$n\sqrt{h}\left(\widehat{\beta}(z) - \beta(z) - h^q \mathcal{B}_1^K(z)\right)$$

$$\Rightarrow \left\{f_z(z)\int B_2(r)B_2(r)'\mathrm{d}r\right\}^{-1} \int B_2(r)\mathrm{d}B_3(r) = MN\left(0, \Omega(z)\right)$$

where the bias term

$$\mathcal{B}_1^K(z) = \frac{\mu_q(K)}{f_z(z)}\left[\sum_{i=1}^{q} \frac{1}{i!} f_z^{(q-i)}(z)\beta^{(i)}(z)\right],$$

and covariance matrix

$$\Omega(z) = \frac{\sigma_u^2}{f_z(z)}\nu_0(K)\left[\int B_2 B_2'\right]^{-1}. \tag{4}$$

Remark 2. Just like the conventional nonparametric regression, the bias term is determined by the order of kernel function and the smoothness of the regression function. If we use a second order kernel as most empirical applications, the bias term is

$$\mathcal{B}_1^K(z) = \frac{\mu_2(K)}{f_z(z)}\left\{f_z^{(1)}(z)\beta^{(1)}(z) + \frac{1}{2}f_z(z)\beta^{(2)}(z)\right\}$$
$$= \mu_2(K)\left\{\frac{f_z^{(1)}(z)}{f_z(z)}\beta^{(1)}(z) + \frac{1}{2}\beta^{(2)}(z)\right\}.$$

Remark 3. The nonparametric estimation of the cointegrating vector function $\widehat{\beta}(z)$ is nonparametrically super-consistent in the sense that it converges to the true cointegrating vector function $\beta(z)$ at rate $n\sqrt{h}$. This is expected: (i) Due to nonstationarity (and thus strong signal) of the regressors, we get rate $n\sqrt{h}$ which is faster than the $\sqrt{nh}$ rate comparing to the conventional functional coefficient estimators in stationary time series regression, a similar result as the super consistency of linear nonstationary time series regression. (ii) Due to local smoothing, we get rate $n\sqrt{h}$ which is slower than the rate n comparing to the conventional linear regression with integrated regressors, a similar result as the nonparametric time series regression.

Remark 4. Limiting distribution of the estimator is a mixture of normal, (thus the standardized estimator has a normal limit). This facilitates statistical inference on the cointegrating vector.

3. Inference

We consider two important inference problems related to the nonstationary time series regression models. The first inference problem tests instability of the cointegrating vector, i.e. constancy of regression coefficients $\beta(z)$. This has been the focus of the conventional functional coefficient regression in iid or stationary time series models, see, e.g. Cai et al. (2000). The second inference problem is to test for cointegration between y and x (irrespective of (in)stability of cointegrating vector). This is one of the most important inference problem in cointegration models.

3.1. Testing stability of the cointegrating vector

We first study testing stability of the cointegrating parameters, corresponding to

$$H_0: \beta(z) = \beta, \quad \text{v.s. } H_1: \text{varying coefficients } \beta(z). \tag{5}$$

In addition to the nonparametric functional coefficient estimator $\widehat{\beta}(z)$, we consider the (null) restricted regression (2). Denote, say, the OLS regression estimator of (2) as $\widetilde{\beta}$, then under the null of constant coefficients,

$$n(\widetilde{\beta} - \beta) \Rightarrow \left(\int B_2 B_2'\right)^{-1} \int B_2 \mathrm{d}B_1.$$

A natural candidate in testing the stability of $\beta(z)$ is to look at $\left(\widehat{\beta}(z) - \widetilde{\beta}\right)$ over a range of z. Notice that, if we under smooth in the nonparametric estimation by choosing $nh^{q+1/2} \to 0$, and under the null hypothesis,

$$n\sqrt{h}\left(\widehat{\beta}(z) - \widetilde{\beta}\right) = n\sqrt{h}\left(\widehat{\beta}(z) - \beta\right) + o_p(1) \Rightarrow MN\left(0, \Omega(z)\right).$$

Notice that $f_z(z)\int B_2 B_2'$ may be estimated by

$$\frac{1}{n^2 h}\sum_{t=1}^{n} K\left(\frac{z_t - z}{h}\right) x_t x_t'$$

and σ_u^2 can be estimated by

$$\frac{1}{n}\sum_{t=1}^{n} \widehat{u}_t^2, \quad \text{where } \widehat{u}_t \text{ is the nonparametric regression residual,}$$

and $\nu_0(K)$ is a constant which only depends on the kernel functions, we can consistently estimate $\Omega(z)$. Denote the consistent estimator of $\Omega(z)$ as $\widehat{\Omega}(z)$, we have

$$n\sqrt{h}\widehat{\Omega}(z)^{-1/2}\left(\widehat{\beta}(z) - \widetilde{\beta}\right) \Rightarrow N\left(0, I_k\right),$$

and

$$\left\|n\sqrt{h}\widehat{\Omega}(z)^{-1/2}\left(\widehat{\beta}(z) - \widetilde{\beta}\right)\right\|^2 \Rightarrow \chi^2(k),$$

Assumption W2. As $n \to \infty$, $h \to 0$, $nh^{q+1/2} \to 0$, and $n\sqrt{h} \to \infty$.

In order to look at $\widehat{\beta}(z) - \widetilde{\beta}$ over a range of z, and construct an asymptotically valid test, we need to find out the joint distribution of the estimated functional coefficients over a number of points. Let z_i^* $(i = 1, \ldots, m)$ be m distinct points, the joint distribution of $\widehat{\beta}(z_i^*)$ $(i = 1, \ldots, m)$ is given by the following Theorem.

Theorem 2. *Under Assumptions A, B, K and W2, the kernel estimators of the cointegrating parameters have the following limiting joint distribution*

$$n\sqrt{h}\begin{pmatrix}\widehat{\beta}(z_1^*) - \beta(z_1^*)\\ \cdots\cdots \\ \widehat{\beta}(z_m^*) - \beta(z_m^*)\end{pmatrix} \Rightarrow MN\left(0, \begin{pmatrix}\Omega(z_1^*) & & 0\\ & \ddots & \\ 0 & & \Omega(z_m^*)\end{pmatrix}\right)$$

where $\Omega(z)$ is defined by (4).

Now we can define the test statistic. For distinct $z_1^*, \ldots, z_m^*$, define

$$T_m = \max_{1\le i\le m}\left\| n\sqrt{h}\widehat{\Omega}(z_i^*)^{-1/2}\left(\widehat{\beta}(z_i^*) - \widetilde{\beta}\right)\right\|^2. \tag{6}$$

Then we obtain the following result.

Theorem 3. *Let $z_1^*, \ldots, z_m^*$ be distinct points and suppose Assumptions A, B, K and W2 hold. Under H_0 given by (5), as $n \to \infty$,*

$$T_m \xrightarrow{d} \max_{1\le i\le m} \chi_i^2(k),$$

where $\chi_1^2(k), \ldots, \chi_m^2(k)$ are independent chi-square random variates with k degrees of freedom.

Thus one rejects the null if T_m is too large. The critical value of T_m can be easily simulated and tabulated since the limiting distribution of T_m is a functional of independent chi-square random variates (with k degrees of freedom) that is free of nuisance parameters.

The testing procedure given by (6) and Theorem 4 is an asymptotic test. It has the advantage that its limiting distribution is free of nuisance parameters. As an alternative, we may consider a bootstrap based test of (6), which may give some finite sample improvement.

Another issue related to the proposed test is the choice of finite distinct points $\{z_i^*\}_{i=1}^m$. In practice, we may consider, say choosing lower quartile, median, and upper quartiles, or we may construct the test based on all deciles. In some applications, different choices of m and the points $\{z_i^*\}_{i=1}^m$ may potentially lead to different conclusions in finite samples, thus it would be desirable to consider all points z on the domain z, and treat $\widehat{\beta}(z)$ as a process in z, and Kolmogorov–Smirnov or Cramer-von-Mises type tests may be constructed.

3.2. Testing for cointegration

In many economic applications, researchers are interested in whether there is a long run equilibrium relationship between y_t and x_t, i.e. cointegration between y and x irrespective of (in)stability of cointegrating vector. Arguably the most important inference problem in these models might be testing the hypothesis of cointegration. For that reason, we focus on this inference problem in the current paper. Under the null hypothesis H_0 : y_t and x_t are cointegrated (possibly with varying cointegrating parameters), i.e., $u_t = y_t - \beta(z_t)'x_t$ is stationary. Under the alternative, H_1 : y_t and x_t are not cointegrated, i.e. $y_t - \beta(z_t)'x_t$ is I(1) under H_1. In particular, in place of Assumptions A and B (or B′), we make the following assumption for the behavior of time series y_t and x_t under the alternative hypothesis.

Assumption C. The process $Z_t = \{\Delta x_t', \Delta y_t\}$ is a zero-mean stationary processes and its partial sum process converge to a $(k+1)$ dimensional Brownian motion with a non-singular covariance matrix, i.e.

$$\frac{1}{\sqrt{n}}\sum_{t=1}^{[nr]} Z_t = \left(\frac{x_{[nr]}'}{\sqrt{n}}, \frac{y_{[nr]}}{\sqrt{n}}\right)' \Rightarrow \begin{pmatrix}B_2(r)\\ B_y(r)\end{pmatrix} = BM(\Omega)$$

where Ω is nonsingular.

We also modify the bandwidth condition from W to W′. It is easy to see that Assumption W3 implies Assumption W.

Assumption W3. As $n \to \infty$, $h^q n^{3/4} \to 0$, and $\sqrt{n}h \to \infty$ ($q = P + 1$ in the case of local polynomial estimation).

A natural way to consider this inference problem is to look at the nonparametric regression residuals

$$\widehat{u}_t = y_t - \widehat{\beta}(z_t)'x_t.$$

We could consider constructing a test statistic in a similar way as Xiao and Phillips (2002) by examining the fluctuation of $\widehat{u}_t$. However, the limiting distribution of the testing statistic will now be affected by the nonparametric estimation and the limiting distribution of the test may depend on the function, a bootstrap procedure may be needed to obtain the critical value. In this paper, we consider a different approach.

Notice that when u_t is weakly stationary, $E(u_t^2) = \sigma_u^2 =$ constant. However, when u_t is a unit root nonstationary process, say

$$u_t = u_{t-1} + \varepsilon_t, \quad t = 1, \ldots, n,$$

where $\varepsilon_t = iid(0, \sigma^2)$ and the initial value u_0 is $O_p(1)$. Then it is well-known that $E(u_t^2)$ increases over time:

$$E(u_t^2) = \alpha + \beta t, \tag{7}$$

where $\alpha = \mathrm{Var}(u_0)$ and $\beta = \sigma^2$. This suggest that, intuitively, we may consider an analogue of regression (7), i.e. a regression of $\widehat{u}_t^2$ on a linear trend.[1]

Let $\widehat{u}_t$ be the nonparametric regression residual, we consider the following regression based on:

$$\widehat{u}_t^2 = \widehat{a} + \widehat{b}t + \widehat{e}_t,$$

we construct tests for cointegration based on the t-ratio statistic of the coefficients with the time trend t. Denote

$$\mathcal{T}_n = \frac{\widehat{b}}{\widehat{s}(b)},$$

where

$$\widehat{s}(b) = \sqrt{\underline{\widehat{\omega}}^2 / \sum (t - \bar{t})^2},$$

where $\bar{t} = n^{-1}\sum_{t=1}^n t$, and $\underline{\widehat{\omega}}^2$ is a consistent nonparametric estimator (with lag truncation bandwidth parameter M) of the long-run variances of u_t^2. For example, we may use the following kernel estimator

$$\underline{\widehat{\omega}}^2 = \sum_{h=-M}^{M} k\left(\frac{h}{M}\right) C(h), \tag{8}$$

where $k(\cdot)$ is the lag window defined on $[-1, 1]$ with $k(0) = 1$, and M is the bandwidth parameter satisfying the property that $M \to \infty$ and $M/n \to 0$ as the sample size $n \to \infty$. The quantity $C(h)$ is the sample covariance defined by $C(h) = n^{-1}\sum'[\widehat{u}_t^2 - n^{-1}\sum_j \widehat{u}_j^2][\widehat{u}_{t+h}^2 - n^{-1}\sum_j \widehat{u}_j^2]$, and $\sum'$ signifies summation over $1 \le t, t+h \le n$.

[1] This intuition motivates our construction of the test $\mathcal{T}_n$. A more rigorous way of analysis can be given based on the theory of spurious regression.

Theorem 4. (i) *If the time series is characterized by model (3), under Assumptions A, B, K and W3,*

$$\mathcal{T}_n \Rightarrow N(0,1), \quad \text{as } n \to \infty.$$

(ii) *Under the alternative of no cointegration and Assumptions C, K and W3,*

$$\mathcal{T}_n \xrightarrow{P} \infty, \quad \text{as } n \to \infty.$$

In particular, as $n \to \infty$, $\Pr[\mathcal{T}_n > B_n] \to 1$ for any nonstochastic sequence $B_n = o(n^{1/2}M^{-1/2})$., where M is the bandwidth in estimating $\widehat{\omega}^2$ which satisfies the assumption $M \to \infty$, $M/n \to 0$, as $n \to \infty$.

Remark 5. In addition to the two inference problems that we consider in this paper, other applications are possible based the results of Section 2. For example, the result in Section 2 is also useful for deriving nonparametric confidence intervals for the cointegrating coefficient function. With undersmoothing bandwidth assumptions, the bias term can be ignored and a confidence interval for $\beta_i(z)$ $(i = 1, \ldots, k)$ at significance level α can be constructed as follows

$$\left[\widehat{\beta}_i(z) - \frac{\widehat{s}_i(z)}{n\sqrt{h}} z_{\alpha/2}, \widehat{\beta}_i(z) + \frac{\widehat{s}_i(z)}{n\sqrt{h}} z_{\alpha/2}\right], \tag{9}$$

where $\Phi(z_{\alpha/2}) = 1 - \alpha/2$ with $\Phi(\cdot)$ the standard normal distribution, and $\widehat{s}_i(z)^2$ is the i-th diagonal element of the covariance matrix estimator $\widehat{\Omega}(z)$.

4. Generalizations

4.1. Cointegration model with functional coefficients

In many economic applications of cointegration, the residual term u_t in regression (3) is correlated with regressors. In this case, subtle issues arise due to endogeneity of the cointegration model.

For this reason, we decompose the residual term u_t into a pure innovation component ε_t and a component related to (and thus can be represented as leads and lags of) Δx_t. In particular, we make the following assumption which modifies Assumption A and allows for correlation between u_t and Δx_t.

Assumption A′. Let $v_t = \Delta x_t$, $\{u_t, v_t, z_t\}$ is a zero-mean, stationary sequence of $(k+2)$-dimensional random vectors and u_t has the following representation

$$u_t = \sum_{j=-K}^{K} v'_{t-j}\pi_j + \varepsilon_t, \tag{10}$$

where ε_t is a stationary process such that

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

In addition, the process $Y_t = \{\varepsilon_t, v_t, z_t\}$ is geometrically absolute regular (β-mixing) with finite forth moment and a nonsingular long-run variance matrix.

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

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

where $f_{\varepsilon\varepsilon}(\lambda)$, $f_{uu}(\lambda)$, $f_{vv}(\lambda)$ are spectral densities of ε, u, v, and $f_{uv}(\lambda)$ is the cross spectral of u and v, implying that the long run variance of ε is $\omega^2_{\varepsilon\varepsilon} = \omega^2_{uu} - \Omega_{uv}\Omega^{-1}_{vv}\Omega_{vu}$.

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

$$y_t = \beta(z_t)'x_t + \sum_{j=-K}^{K} \Delta x'_{t-j}\pi_j + \varepsilon_t.$$

Many applications also introduce an intercept term in the regression. For this reason, we consider the following cointegrating regression model with stationary covariates w_t which are leads and lags of $\Delta x'_t$:

$$y_t = \alpha(z_t) + \beta(z_t)'x_t + \pi(z_t)'w_t + \varepsilon_t. \tag{11}$$

Let

$$\theta(z_t) = (\alpha(z_t), \beta(z_t)', \pi(z_t)')', \qquad X_t = (1, x'_t, w'_t)'$$

then the regression model (11) can be expressed as

$$y_t = \theta(z_t)'X_t + \varepsilon_t. \tag{12}$$

Remark 6. Similar to the traditional cointegration analysis, in addition to the method of "leads and lags" that we use here, "fully-modification" and canonical regression methods may also be used in functional-coeffcient cointegration models to deal with the endogeneity problem. In a time varying-coefficient cointegration model, Park and Hahn (1999) used the technique of canonical regression to handle endogeneity. Xiao (2009) studied fully-modification in cointegration models where the coefficients may vary over quantiles of the conditional distribution. Similar ideas may be applied to the functional-coefficient cointegration models.

The kernel estimation of (12) is given by

$$\widehat{\theta}(z) = \arg\min \sum_{t=1}^{n} K\left(\frac{z_t - z}{h}\right) \left\{y_t - X'_t\theta\right\}^2.$$

In the presence of both stationary and nonstationary regressors, the convergence rate of different component in $\widehat{\theta}(z)$ will be different. For this reason, conformable with the three components of θ $(\alpha(z), \beta(z), \pi(z))$ we introduce the following standardizing matrix

$$D_n = \operatorname{diag}\left[1, n^{-1/2}I_k, I_\ell\right].$$

The following Lemma is an extension of Lemma 1.

Lemma 2. *Under Assumptions A′, B, K and W,*

$$\begin{pmatrix} \frac{1}{\sqrt{n}}\sum_{t=1}^{[nr]}\varepsilon_t \\ \frac{1}{\sqrt{n}}\sum_{t=1}^{[nr]} v_t \\ \frac{1}{\sqrt{nh}}\sum_{t=1}^{[nr]} \underline{K}_t\varepsilon_t \\ \frac{1}{\sqrt{nh}}\sum_{t=1}^{[nr]} w_t\underline{K}_t\varepsilon_t \end{pmatrix} \Rightarrow \begin{pmatrix} B_1(r) \\ B_2(r) \\ B_3(r) \\ B_4(r) \end{pmatrix}$$

with Covariance matrix

$$\begin{bmatrix} \omega_1^2 & 0 & 0 & 0 \\ 0 & \Omega_{22} & 0 & 0 \\ 0 & 0 & \nu_0(K)f_z(z)\sigma_\varepsilon^2 & \nu_0(K)f_z(z)\mu'_w(z)\sigma_\varepsilon^2 \\ 0 & 0 & \nu_0(K)f_z(z)\sigma_\varepsilon^2\mu_w(z) & \nu_0(K)f_z(z)\sigma_\varepsilon^2 E\left(w_t w'_t|z\right) \end{bmatrix}$$

$$= \begin{bmatrix} \omega_1^2 & 0 & 0 \\ 0 & \Omega_{22} & 0 \\ 0 & 0 & V(K) \end{bmatrix}$$

where $\nu_0(K) = \int K(u)^2\,\mathrm{d}u$,

$$V(K) = \nu_0(K) f_z(z)\sigma_\varepsilon^2 \begin{bmatrix} 1 & \mu_w'(z) \\ \mu_w(z) & \Sigma_w(z) \end{bmatrix}$$

and

$$\mu_w(z) = E\left[w_t | z_t = z\right], \qquad \Sigma_w(z) = E\left[w_t w_t' | z_t = z\right].$$

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

Theorem 5. *Under Assumptions A′, B, K and W,*

$$\sqrt{nh} D_n^{-1}\left(\widehat{\theta}(z) - \theta(z) - h^q \mathcal{B}^K(z)\right) \Rightarrow MN\left(0, \frac{\nu_0(K)}{f_z(z)}\sigma_\varepsilon^2 Q^{-1}\right)$$

where

$$\mathcal{B}^K(z) = \frac{\mu_q(K)}{f_z(z)} Q^{-1} \times \left(\sum_{i=1}^q \frac{1}{i!}\begin{bmatrix} f_z^{(q-i)}(z)\int \overline{B}_2\overline{B}_2' & \int \overline{B}_2 \mu_{w,i}(z)' \\ \mu_{w,i}(z)\int \overline{B}_2' & \Sigma_{w,i}(z)\end{bmatrix}\theta^{(i)}(z)\right)$$

with

$$f_{zw(1,j)}(z, w) = \frac{\partial^j f_{zw}(z, w)}{\partial z^j},$$

$$\mu_{w,i}(z) = \int w f_{zw(1,q-i)}(z, w)\mathrm{d}w,$$

$$\Sigma_{w,i}(z) = \int w w' f_{zw(1,q-i)}(z, w)\mathrm{d}w, \qquad \overline{B}_2(r) = \begin{bmatrix} 1 \\ B_2(r)\end{bmatrix},$$

$$Q = \begin{bmatrix} \int \overline{B}_2\overline{B}_2'\mathrm{d}r & \left(\int \overline{B}_2\right)\mu_w'(z) \\ \mu_w(z)\int \overline{B}_2 & \Sigma_w(z)\end{bmatrix}.$$

Remark 7. In the special cases when z_t is independent with w_t, both Q and $V(K)$ are block diagonal:

$$Q = \begin{bmatrix} \int \overline{B}_2\overline{B}_2'\mathrm{d}r & 0 \\ 0 & \Sigma_w(z)\end{bmatrix}, \quad \text{and}$$

$$V(K) = \nu_0(K) f_z(z)\sigma_\varepsilon^2 \begin{bmatrix} 1 & 0 \\ 0 & \Sigma_w(z)\end{bmatrix},$$

thus the limiting distributions of estimators $\widehat{\beta}(z)$ and $\widehat{\pi}(z)$ are independent. In particular,

$$\begin{pmatrix} \sqrt{nh}\left[\widehat{\alpha}(z) - \alpha(z)\right] \\ n\sqrt{h}\left[\widehat{\beta}(z) - \beta(z)\right] \end{pmatrix} - h^q B^K(z) \Rightarrow MN\left(0, \frac{\nu_0(K)}{f_z(z)}\sigma_\varepsilon^2\left(\int \overline{B}_2\overline{B}_2'\mathrm{d}r\right)^{-1}\right)$$

where

$$B^K(z) = \mu_q(K)\sum_{i=1}^q \frac{1}{i!}\frac{f_z^{(q-i)}(z)}{f_z(z)}\begin{bmatrix}\alpha^{(i)}(z) \\ \beta^{(i)}(z)\end{bmatrix},$$

and the estimator of the cointegrating vector function $\widehat{\beta}(z)$ has the following limit

$$\left(n\sqrt{h}\left[\widehat{\beta}(z) - \beta(z)\right] - h^q B_1^K(z)\right) \Rightarrow MN\left(0, \frac{\nu_0(K)}{f_z(z)}\sigma_\varepsilon^2\left(\int \underline{B}_2\underline{B}_2'\mathrm{d}r\right)^{-1}\right)$$

where

$$B_1^K(z) = \mu_q(K)\sum_{i=1}^q \frac{1}{i!}\frac{f_z^{(q-i)}(z)}{f_z(z)}\beta^{(i)}(z),$$

$$\underline{B}_2(r) = B_2(r) - \int_0^1 B_2(s)\mathrm{d}s.$$

Remark 8. All inference procedures proposed in Section 3 can be extended to the general model with minor modification.

4.2. *Local polynomial regression*

In addition to the kernel estimation used in this paper, other nonparametric methods can also be applied and qualitatively similar results can be obtained. One important alternative approach is the local polynomial estimation. See Fan (1992), and Fan and Gijbels (1996) for discussions on the attractive properties of local polynomials. In this subsection, we briefly discuss asymptotic result of local polynomial estimators of our models. For simplicity and without loss of generality, we focus on model (12).

We use a nonnegative kernel $\mathcal{K}$ for the local polynomial estimation and make the following assumption about $\mathcal{K}$. In addition, we assume that $q = P + 1$ so that the following results are comparable to the previous ones.

Assumption L. The kernel $\mathcal{K}$ is nonnegative, symmetric about zero, bounded, and has compact connected support ($\mathcal{K}(u) = 0$ for $\|u\| > A_0$ some A_0).

The local polynomial nonparametric estimator of the functional coefficient model minimizes the locally smoothed mean square error

$$\sum_{t=1}^n K\left(\frac{z_t - z}{h}\right)\left\{y_t - X_t'\left[\sum_{p=0}^P \theta_p (z_t - z)^p\right]\right\}^2. \tag{13}$$

Let

$$\Theta = \Theta(z) = \left[\theta_0', \theta_1', \ldots, \theta_P'\right] = \left[\theta(z)', \frac{\partial\theta'(z)}{\partial z}, \ldots, \frac{1}{P!}\frac{\partial^P\theta(z)}{\partial z^P}\right]$$

$$H_P = \begin{bmatrix} 1 & & & \\ & h & & \\ & & \ddots & \\ & & & h^P \end{bmatrix},$$

$$H = H_P \otimes I_k = \begin{bmatrix} I_k & & & \\ & hI_k & & \\ & & \ddots & \\ & & & h^P I_k \end{bmatrix}$$

and

$$\mathbf{X}_t(z)' = \left[X_t', X_t'(z_t - z), \ldots, X_t'(z_t - z)^p, \ldots, X_t'(z_t - z)^P\right]$$

then

$$\widehat{\Theta} = \arg\min \sum_{t=1}^n K\left(\frac{z_t - z}{h}\right)\left\{y_t - \mathbf{X}_t(z)'\Theta\right\}^2.$$

The limiting distribution of the local polynomial estimators is given by Theorem 6.

Theorem 6. *Under Assumptions A′, B, L and W,*

$$\sqrt{nh}\left[H_P \otimes D_n^{-1}\left(\widehat{\Theta}-\Theta\right)-h^{P+1}\mathcal{B}(z,D_n^{-1})\right]$$
$$\Rightarrow MN\left(0,\frac{\sigma_\varepsilon^2}{f_z(z)}\left(\mathcal{M}(\mathcal{K})\otimes Q^{-1}\right)\right)$$

where

$$\mathcal{B}(z,D_n^{-1})=\frac{1}{(P+1)!}\left[M(\mathcal{K})^{-1}B\otimes D_n^{-1}\theta^{(P+1)}(z)\right],$$
$$\mathcal{M}(\mathcal{K})=M(\mathcal{K})^{-1}N(\mathcal{K})M(\mathcal{K})^{-1}.$$

In particular, if we look at θ, and denote the first element of $M(K)^{-1}B(K)$ as $\eta(K)$, then

$$\sqrt{nh}D_n^{-1}\left[\widehat{\theta}-\theta-h^{P+1}B_{\mathcal{K}}(z)\right]\Rightarrow MN\left(0,\frac{\sigma_\varepsilon^2}{f_z(z)}\varphi(\mathcal{K})Q^{-1}\right)$$

where

$$B_{\mathcal{K}}(z)=\frac{1}{(P+1)!}\eta(\mathcal{K})\theta^{(P+1)}(z),$$

and $\varphi(K)$ is the upper-left $(1,1)$ element of the matrix $M(K)$.

Remark 9. In the special case of local linear regression, $p=1$, and $X_t=x_t$, if we consider the estimation of the cointegrating vector $\beta(z)$:

$$n\sqrt{h}\left[\widehat{\beta}(z)-\beta(z)-\frac{h^2}{2}\frac{\mu_2^2-\mu_1\mu_3}{\mu_2-\mu_1^2}\beta''(z)\right]$$
$$\Rightarrow MN\left(0,\frac{\sigma_u^2}{f_z(z)}\frac{\mu_2^2\nu_0-2\mu_1\mu_2\nu_1+\mu_1^2\nu_2}{\left(\mu_2-\mu_1^2\right)^2}\left(\left(\int B_2B_2{}'\right)^{-1}\right)\right).$$

In the case $\mu_1=0$,

$$n\sqrt{h}\left[\widehat{\beta}(z)-\beta(z)-h^2\frac{\mu_2\beta''(z)}{2}\right]$$
$$\Rightarrow MN\left(0,\frac{\nu_0\sigma_u^2}{f_z(z)}\left(\left(\int B_2B_2{}'\right)^{-1}\right)\right).$$

5. Finite sample performance

A small Monte Carlo experiment was conducted to examine the finite sample performance of the proposed tests. In particular, we compared the proposed tests with the CUSUM test for cointegration of Xiao and Phillips (2002) for different sample sizes ($n = 100, 300$), and examined the empirical size and power of these tests.

Two kernel smoothing procedures are used: one for the functional coefficient estimation, another for long-run variance estimation. For the functional coefficient estimation, we use the forth order kernel $k(u) = 15(7u^4 - 10u^2 + 3)1(|u| \le 1)/32$, and a simple bandwidth choice $h = d \cdot s_z \cdot T^{-1/3}$ where s_z is the standard deviation of z and d is a constant. We choose $d = 1$. In estimating the long-run variances, we use the Bartlett window $k(x) = 1 - |x|$ so that the nonnegativity of the long-run variance estimator is guaranteed, and two truncation numbers $M_1 = 1$ and $M_2 = [4(n/100)^{1/4}]$ are investigated. The first bandwidth value, $M_1 = 1$, is small and fixed, while the other is a function of the sample sizes and increases with n, it has been used in Schwert (1989), Kwiatkowski et al. (1992), Xiao and Phillips (2002) and other simulations. In the presence of serial correlation, we need the bandwidth increase with n in estimating the long-run variance. Thus, we expect that tests with M_2 will work better for cases with high serial correlation. All experiments use 5000 replications.

We consider the following data-generating process:

$$y_t = \beta(z_t)x_t + u_t, \tag{14}$$

where

$$x_t = x_{t-1} + v_t,$$

and

$$u_t = \rho u_{t-1} + \varepsilon_t$$

where ε_t and v_t are *iid* $N(0,1)$ and are independent with each other. The initial values are all set to be zero. In model (14), y_t and x_t are cointegrated when $|\rho| < 1$.

When $\rho = 1$, there is no cointegration. The AR coefficient ρ measures the distance from the null to the alternative. The following cases are investigated in the Monte Carlo experiment: (i) $\beta(z_t)=1, \rho=0$; (ii) $\beta(z_t)=1, \rho=0.5$; (iii) $\beta(z_t)=1, \rho=0.8$; (iv) $\beta(z_t)=z_t^2, \rho=0$; (v) $\beta(z_t)=1, \rho=1$; (vi) $\beta(z_t)=z_t^2$, $\rho=1$. Empirical rejection rates in the first 4 cases ((i)–(iv)) give empirical sizes, and results for the last two cases ((v)–(vi)) deliver the power.

Tables of the empirical rejection rates of the tests for different sample sizes can be found in an earlier version of this paper. The Monte Carlo results suggest: (1) The finite sample performance of the proposed test is largely affected by the bandwidth choices, and selection of bandwidth is an important issue for future research; (2) When the true models have constant coefficients, the traditional methods has better performance over the functional coefficient estimation-based approach as we expected. (3) In the presence of varying coefficients with values depending on the covariates, the traditional methods have very high size distortion, and the functional coefficient model display better properties, corroborating the asymptotic theory.

Acknowledgements

The author thank Roger Koenker, Steve Portnoy, Dan Bernhardt, Jianqing Fan, two referees, conference participants at "Econometrics at Rio" and seminar participants at University of Illinois for their helpful comments.

Appendix. Sketch of proofs

Proof of Lemma 1. Under our assumptions, we can verify the regularity conditions of, say, Theorem 27.17 of Davidson (1994) and use the multivariate invariance principle. Notice that

$$\frac{1}{n}\sum_{t=1}^{n}u_tu_{t+j}\xrightarrow{P}\gamma_u(j)=E\left(u_tu_{t+j}\right),$$
$$\frac{1}{n}\sum_{t=1}^{n}v_tv_{t+j}\xrightarrow{P}\Gamma_v(j)=E\left(v_tv_{t+j}\right),$$

and

$$\frac{1}{nh}\sum_{t=1}^{n}\underline{K}_tu_t\underline{K}_{t+j}u_{t+j}\xrightarrow{P}\gamma_u(0)f_z(z)\nu_0(K),\quad \text{if } j=0$$
$$\frac{1}{nh}\sum_{t=1}^{n}\underline{K}_tu_t\underline{K}_{t+j}u_{t+j}=O_p(h)\xrightarrow{P}0,\quad \text{if } j\neq 0$$

the multivariate partial sum process converges to a vector Brownian motion with the covariance matrix specified in the Theorem. ■

Proof of Theorem 1. Notice that

$$\widehat{\beta}(z) = n^{-1/2}\left[\frac{1}{nh}\sum_{t=1}^{n}\frac{x_t}{\sqrt{n}}\frac{x_t'}{\sqrt{n}}K\left(\frac{z_t-z}{h}\right)\right]^{-1} \times\left[\frac{1}{nh}\sum_{t=1}^{n}\frac{x_t}{\sqrt{n}}y_tK\left(\frac{z_t-z}{h}\right)\right]$$
$$= \beta(z) + n^{-1/2}\left[\frac{1}{nh}\sum_{t=1}^{n}\frac{x_t}{\sqrt{n}}\frac{x_t'}{\sqrt{n}}K\left(\frac{z_t-z}{h}\right)\right]^{-1} \times\left[\frac{1}{nh}\sum_{t=1}^{n}\frac{x_t}{\sqrt{n}}\frac{x_t'}{\sqrt{n}}K\left(\frac{z_t-z}{h}\right)\sqrt{n}\left[\beta(z_t)-\beta(z)\right]\right]$$
$$+ n^{-1/2}\left[\frac{1}{nh}\sum_{t=1}^{n}\frac{x_t}{\sqrt{n}}\frac{x_t'}{\sqrt{n}}K\left(\frac{z_t-z}{h}\right)\right]^{-1} \times\left[\frac{1}{nh}\sum_{t=1}^{n}K\left(\frac{z_t-z}{h}\right)\frac{x_t}{\sqrt{n}}u_t\right].$$

Denote

$$\Sigma_{n,h} = \frac{1}{nh}\sum_{t=1}^{n}\frac{x_t}{\sqrt{n}}\frac{x_t'}{\sqrt{n}}K\left(\frac{z_t-z}{h}\right),$$
$$b_{n,h} = \frac{1}{nh}\sum_{t=1}^{n}\frac{x_t}{\sqrt{n}}\frac{x_t'}{\sqrt{n}}K\left(\frac{z_t-z}{h}\right)\left[\beta(z_t)-\beta(z)\right],$$
$$V_{n,h} = \frac{1}{\sqrt{nh}}\sum_{t=1}^{n}K\left(\frac{z_t-z}{h}\right)\frac{x_t}{\sqrt{n}}u_t$$

we have

$$n\sqrt{h}\left(\widehat{\beta}(z)-\beta(z)-\Sigma_{n,h}^{-1}b_{n,h}\right) = \Sigma_{n,h}^{-1}V_{n,h}.$$

Let us first look at $\Sigma_{n,h}$.

$$\frac{1}{n^2h}\sum_{t=1}^{n}K\left(\frac{z_t-z}{h}\right)x_tx_t' = \frac{1}{n^2h}\sum_{t=1}^{n}\mathrm{E}\left[K\left(\frac{z_t-z}{h}\right)\right]x_tx_t' + \frac{1}{n^2h}\sum_{t=1}^{n}\left[K\left(\frac{z_t-z}{h}\right)-\mathrm{E}\left[K\left(\frac{z_t-z}{h}\right)\right]\right]x_tx_t'$$
$$= \frac{1}{n^2h}\sum_{t=1}^{n}\mathrm{E}\left[K\left(\frac{z_t-z}{h}\right)\right]x_tx_t' + o_p(1)$$
$$\Rightarrow f_z(z)\int B_2(r)B_2(r)'\mathrm{d}r$$

since

$$\frac{1}{n^2h}\sum_{t=1}^{n}\left[K\left(\frac{z_t-z}{h}\right)-\mathrm{E}\left[K\left(\frac{z_t-z}{h}\right)\right]\right]x_tx_t' \xrightarrow{P} 0$$

and

$$\mathrm{E}\left[\frac{1}{h}K\left(\frac{z_t-z}{h}\right)\right] = \int\frac{1}{h}K\left(\frac{z_t-z}{h}\right)f_z(u)\mathrm{d}u \to f_z(z).$$

Next we consider the "variance" term $V_{n,h}$.

$$\frac{1}{\sqrt{nh}}\sum_{t=1}^{n}\left(n^{-1/2}x_tK\left(\frac{z_t-z}{h}\right)u_t\right) = \left(\frac{1}{n\sqrt{h}}\sum_{t=1}^{n}x_t\underline{K}_tu_t\right) + \left(\frac{1}{n\sqrt{h}}\sum_{t=1}^{n}x_tu_tEK\left(\frac{z_t-z}{h}\right)\right)$$
$$= \left(\frac{1}{n\sqrt{h}}\sum_{t=1}^{n}x_t\underline{K}_tu_t\right) + o_p(1)$$
$$\Rightarrow\left[\int B_2(r)\mathrm{d}B_3(r)\right].$$

Notice that the k-dimensional Brownian motion $B_2(r)$ is independent with Brownian motion $B_3(r)$, the limiting variate $\int B_2(r)\mathrm{d}B_3(r)$ is a mixture of normal random vector $MN\,(0, V_1)$, where

$$V_1 = \nu_0(K)f_z(z)\sigma_u^2\int B_2(r)B_2(r)'\mathrm{d}r.$$

Finally, for the bias effect, under Assumption B, we can obtain a Taylor expansion of $\beta(z_t)$ around z, and

$$b_{n,h} \approx \sum_{i=1}^{q}\frac{1}{nh}\sum_{t=1}^{n}\frac{x_t}{\sqrt{n}}\frac{x_t'}{\sqrt{n}}K\left(\frac{z_t-z}{h}\right)\left[\frac{h^i}{i!}\beta^{(i)}(z)\right]$$
$$\approx h^q\mu_q(K)\sum_{i=1}^{q}\frac{1}{i!}f_z^{(q-i)}(z)\left[\int B_2B_2{}'\right]\beta^{(i)}(z)$$
$$= h^q\mu_q(K)\left[\int B_2B_2{}'\right]\sum_{i=1}^{q}\frac{1}{i!}f_z^{(q-i)}(z)\beta^{(i)}(z)$$

thus, the bias term $\Sigma_{n,h}^{-1}b_{n,h}$ is asymptotically equivalent to $h^q\mathcal{B}_1^K(z)$, where:

$$\mathcal{B}_1^K(z) = \left\{f_z(z)\int B_2B_2{}'\right\}^{-1}\left[\mu_q(K)\left[\int B_2B_2{}'\right] \times\sum_{i=1}^{q}\frac{1}{i!}f_z^{(q-i)}(z)\beta^{(i)}(z)\right]$$
$$= \frac{\mu_q(K)}{f_z(z)}\left[\sum_{i=1}^{q}\frac{1}{i!}f_z^{(q-i)}(z)\beta^{(i)}(z)\right]. \quad\blacksquare$$

Proof of Theorem 2. Based on the results of Theorem 1, we only need to show that the covariance matrix is block diagonal, this is true because for $t \neq s$

$$E\left(\frac{1}{h}K\left(\frac{z_t-z}{h}\right)K\left(\frac{z_s-z}{h}\right)\right) = o(1). \quad\blacksquare$$

Proof of Theorem 4. Notice that $\widehat{u}_t = y_t - \widehat{\beta}(z_t)'x_t = u_t - \left(\widehat{\beta}(z_t)-\beta(z_t)\right)'x_t$, and denote

$$\underline{u}_t^2 = u_t^2 - E(u_t^2)$$

under the null of cointegration and Assumptions A, B, K and W′,

$$\left[\sum_t(t-\bar{t})^2\right]^{-1/2}\left[\sum_t(t-\bar{t})\widehat{u}_t^2\right]$$
$$= \left[\sum_t(t-\bar{t})^2\right]^{-1/2}\left[\sum_t(t-\bar{t})\left(u_t-\left(\widehat{\beta}(z_t)-\beta(z_t)\right)'x_t\right)^2\right]$$
$$= \left[\sum_t(t-\bar{t})^2\right]^{-1/2}\left[\sum_t(t-\bar{t})u_t^2\right] + \left[\sum_t(t-\bar{t})^2\right]^{-1/2} \times\left[\sum_t(t-\bar{t})\left(\widehat{\beta}(z_t)-\beta(z_t)\right)'x_tx_t'\left(\widehat{\beta}(z_t)-\beta(z_t)\right)'\right]$$
$$-2\left[\sum_t(t-\bar{t})^2\right]^{-1/2}\left[\sum_t(t-\bar{t})\left(\widehat{\beta}(z_t)-\beta(z_t)\right)'x_tu_t\right].$$

We show that under our assumptions, the second and third terms above are $o_p(1)$. This calculation is quite long. The general strategy is to plug-in the representation of $\widehat{\beta}(z_t) - \beta(z_t)$ and expand out the random denominator around its probability limit and then calculate the moments of the resulting statistics term by term. For example, for the last term, notice that

$$\frac{1}{n}\sum_t \left(\frac{t-\bar{t}}{n}\right)^2 = O(1)$$

and by a geometric expansion it can be shown that the leading terms in

$$\frac{1}{n^{3/2}}\sum_t (t-\bar{t})\left(\widehat{\beta}(z_t)-\beta(z_t)\right)' x_t u_t$$

are

$$\sum_t \sum_{s=1}^n \phi_{1,K,h}(t,s), \quad \text{and} \quad \sum_t \sum_{s=1}^n \phi_{2,K,h}(t,s)$$

where

$$\phi_{1,K,h}(t,s) = \left(\frac{t-\bar{t}}{n}\right)\left[\frac{1}{nh}\frac{x_s}{\sqrt{n}}\frac{x_s'}{\sqrt{n}}\frac{K\left((z_s-z_t)/h\right)}{f_z(z_t)} \times \sqrt{n}\left[\beta(z_s)-\beta(z_t)\right]\right]' \frac{x_t}{\sqrt{n}}\frac{u_t}{\sqrt{n}}$$

and

$$\phi_{2,K,h}(t,s) = \left(\frac{t-\bar{t}}{n}\right)\left[\frac{1}{nh}\frac{K\left((z_s-z_t)/h\right)}{f_z(z_t)}\frac{x_s}{\sqrt{n}}u_s\right]' \frac{x_t}{\sqrt{n}}\frac{u_t}{\sqrt{n}}$$

and by calculation of moments we can verify that under our bandwidth assumptions

$$\frac{1}{n^{3/2}}\sum_t (t-\bar{t})\left(\widehat{\beta}(z_t)-\beta(z_t)\right)' x_t u_t = o_p(1).$$

Thus,

$$\begin{aligned}
&\left[\sum_t (t-\bar{t})^2\right]^{-1/2}\left[\sum_t (t-\bar{t})\widehat{u}_t^2\right] \\
&= \left[\sum_t (t-\bar{t})^2\right]^{-1/2}\left[\sum_t (t-\bar{t})\underline{u}_t^2\right] + O_p\left(\frac{1}{\sqrt{n}}\right) \\
&\Rightarrow \underline{\omega}\left[\int_0^1 \left(r-\frac{1}{2}\right)^2 \mathrm{d}r\right]^{-1/2}\int_0^1\left(r-\frac{1}{2}\right)\mathrm{d}W(r),
\end{aligned}$$

where $W(r)$ is a standard Brownian motion. Thus the t-statistic

$$\mathcal{T}_n = \left[\widehat{\underline{\omega}}^2\sum(t-\bar{t})^2\right]^{-1/2}\left[\sum(t-\bar{t})\underline{u}_t^2\right] \Rightarrow N(0,1).$$

Under the alternative of no cointegration,

$$\widehat{\beta}(z) = \left[\frac{1}{nh}\sum_{t=1}^n \frac{x_t}{\sqrt{n}}\frac{x_t'}{\sqrt{n}}K\left(\frac{z_t-z}{h}\right)\right]^{-1} \times \left[\frac{1}{nh}\sum_{t=1}^n \frac{x_t}{\sqrt{n}}\frac{y_t}{\sqrt{n}}K\left(\frac{z_t-z}{h}\right)\right].$$

By a similar argument as the proof of Theorem 1,

$$\frac{1}{n^2h}\sum_{t=1}^n K\left(\frac{z_t-z}{h}\right)x_t x_t' \Rightarrow f_z(z)\int B_2(r)B_2(r)'\mathrm{d}r,$$

and

$$\frac{1}{n^2h}\sum_{t=1}^n K\left(\frac{z_t-z}{h}\right)x_t y_t \Rightarrow f_z(z)\int B_2(r)B_y(r)\mathrm{d}r,$$

thus

$$\widehat{\beta}(z) \Rightarrow \left[\int B_2(r)B_2(r)'\mathrm{d}r\right]^{-1}\left[\int B_2(r)B_y(r)'\mathrm{d}r\right] = \xi$$

and $\widehat{u}_t = y_t - \widehat{\beta}(z_t)'x_t$

$$\frac{1}{\sqrt{n}}\widehat{u}_{[nr]} = \frac{1}{\sqrt{n}}y_{[nr]} - \widehat{\beta}'\frac{x_{[nr]}}{\sqrt{n}} \Rightarrow B_y(r) - \xi' B_2(r) = B_\xi(r).$$

Then

$$\begin{aligned}
&\frac{1}{n^{3/2}}\left[\sum_t (t-\bar{t})^2\right]^{-1/2}\left[\sum_t (t-\bar{t})\widehat{u}_t^2\right] \\
&= \left[\frac{1}{n}\sum\left(\frac{t-\bar{t}}{n}\right)^2\right]^{-1/2}\left[\frac{1}{n}\sum\left(\frac{t-\bar{t}}{n}\right)\left(\frac{\widehat{u}_t}{\sqrt{n}}\right)^2\right] \\
&\Rightarrow \left[\int_0^1\left(r-\frac{1}{2}\right)^2\mathrm{d}r\right]^{-1/2}\int_0^1\left(r-\frac{1}{2}\right)B_\xi(r)^2\mathrm{d}r.
\end{aligned}$$

Thus, $\left[\sum_t (t-\bar{t})^2\right]^{-1/2}\left[\sum_t (t-\bar{t})z_t\right]$ is $O_p(n^{3/2})$. For the denominator,

$$\widehat{\underline{\omega}}^2 = \frac{1}{2\pi}\sum_{h=-M}^{M} k\left(\frac{h}{M}\right)C(h),$$

where $C(h) = n^{-1}\sum'[\widehat{u}_t^2 - n^{-1}\sum_j \widehat{u}_j^2][\widehat{u}_{t+h}^2 - n^{-1}\sum_j \widehat{u}_j^2]$, and $\sum'$ signifies summation over $1 \le t, t+h \le n$. By a similar argument as Xiao and Phillips (1998), it can be verified that

$$\widehat{\underline{\omega}}^2 = \frac{1}{2\pi}\sum_{h=-M}^{M} k\left(\frac{h}{M}\right)C(h) = O_p(n^2 M).$$

Thus, under the alternative hypothesis,

$\tau_n = O_p(n^{1/2}M^{-1/2})$. ■

Proof of Theorem 5. (Similar to the proof of Theorem 1.) ■

Proof of Theorem 6. By Assumption B, a Taylor expansion of $\theta(z)$ can be obtained to the P-th order ($p = q - 1$):

$$\begin{aligned}
\theta(z_t) &= \theta(z) + \theta'(z)(z_t - z) + \cdots \\
&\quad + \frac{\theta^{(P)}(z)}{P!}(z_t-z)^P + \frac{\theta^{(P+1)}(z)}{(P+1)!}(z_t-z)^{P+1} \\
&\approx \theta_0 + \theta_1(z_t-z) + \cdots + \theta_P(z_t-z)^P.
\end{aligned}$$

Denote

$$\Theta = \Theta(z) = \begin{bmatrix}\theta_0\\ \theta_1\\ \cdot\cdot\\ \theta_P\end{bmatrix} = \begin{bmatrix}\theta(z)\\ \theta'(z)\\ \cdot\cdot\\ \dfrac{\theta^{(P)}(z)}{P!}\end{bmatrix}$$

and

$$\begin{aligned}
&X_t'\theta_0 + X_t'\theta_1(z_t-z) + \cdots + X_t'\theta_p(z_t-z)^p + \cdots X_t'\theta_P(z_t-z)^P \\
&= \left[X_t', X_t'(z_t-z), \ldots, X_t'(z_t-z)^p, \ldots, X_t'(z_t-z)^P\right]\begin{bmatrix}\theta_0\\ \theta_1\\ \cdot\cdot\\ \theta_P\end{bmatrix} \\
&= \mathbf{X}_t(z)'\Theta
\end{aligned}$$

then the local polynomial regression estimator is determined by

$$\min_{\Theta}\sum_{t=1}^{n}\mathcal{K}\left(\frac{z_t-z}{h}\right)\left\{y_t-\mathbf{X}_t(z)'\Theta\right\}^2$$
$$=\min\sum_{t=1}^{n}K_{h,t}\left\{y_t-\mathbf{X}_t(z)'\Theta\right\}^2,$$

where $\mathcal{K}\left(\frac{z_t-z}{h}\right)$ by $K_{h,t}$. The first order condition of the above weighted least square problem is

$$\sum_{t=1}^{n}K_{h,t}\mathbf{X}_t(z)\left\{y_t-\mathbf{X}_t(z)'\widehat{\mathbf{\Theta}}\right\}=0,$$

giving

$$\widehat{\mathbf{\Theta}}=\left[\sum_{t=1}^{n}K_{h,t}\mathbf{X}_t(z)\mathbf{X}_t(z)'\right]^{-1}\left[\sum_{t=1}^{n}K_{h,t}\mathbf{X}_t(z)y_t\right].$$

In order to derive the asymptotic result, we need to re-standardize both $\sum_{t=1}^{n}K_{h,t}\mathbf{X}_t(z)\mathbf{X}_t(z)'$ and $\sum_{t=1}^{n}K_{h,t}\mathbf{X}_t(z)y_t$. Let

$$H_P=\begin{bmatrix}1&&&\\&h&&\\&&\ddots&\\&&&h^P\end{bmatrix},$$

$$H=H_P\otimes I_k=\begin{bmatrix}I_k&&&\\&hI_k&&\\&&\ddots&\\&&&h^PI_k\end{bmatrix}$$

then

$$\sqrt{nh}\left[H_P\otimes D_n^{-1}\left(\widehat{\mathbf{\Theta}}-\mathbf{\Theta}\right)-S_n^{-1}B_n\right]=S_n^{-1}U_n$$

where

$$S_n=\frac{1}{nh}\left(H_P\otimes D_n^{-1}\right)^{-1}\sum_{t=1}^{n}K_{h,t}\mathbf{X}_t(z)\mathbf{X}_t(z)'\left(H_P\otimes D_n^{-1}\right)^{-1}$$

$$B_n=\frac{1}{nh}H^{-1}\sum_{t=1}^{n}K_{h,t}\mathbf{X}_t(z)X_t'D_n\left[\frac{D_n^{-1}\theta^{(P+1)}(z)}{(P+1)!}(z_t-z)^{P+1}\right]$$

$$U_n=\frac{1}{\sqrt{nh}}\left(H_P\otimes D_n^{-1}\right)^{-1}\sum_{t=1}^{n}K_{h,t}\mathbf{X}_t(z)\varepsilon_t.$$

Let us first analyze S_n, which equals the equation given in Box I.

Similar to the previous analysis, we can show that

$$\frac{1}{nh}\sum_{t=1}^{n}K_{h,t}D_nX_tX_t'D_n$$
$$\Rightarrow f_z(z)\begin{bmatrix}\int\overline{B}_2\overline{B}_2{}'&\int\overline{B}_2(r)\mathrm{d}r\mu_w'(z)\\\mu_w(z)\int\overline{B}_2(r)'\mathrm{d}r&\Sigma_w(z)\end{bmatrix}$$
$$=f_z(z)Q.$$

Next, see the equation given in Box II and

$$\frac{1}{nh}\sum_{t=1}^{n}\left(\frac{z_t-z}{h}\right)^iK\left(\frac{z_t-z}{h}\right)\xrightarrow{P}f(z)\int u^iK(u)\,\mathrm{d}u=f(z)\mu_i(K).$$

$$\frac{1}{n^{3/2}h}\sum_{t=1}^{n}\left(\frac{z_t-z}{h}\right)^iK\left(\frac{z_t-z}{h}\right)x_t'$$
$$=\frac{1}{nh}\sum_{t=1}^{n}\left[\mathrm{E}\left(\frac{z_t-z}{h}\right)^iK\left(\frac{z_t-z}{h}\right)\right]\frac{x_t'}{\sqrt{n}}+o_p(1)$$
$$\Rightarrow\mu_i(K)f(z)\int B_2(r)'\mathrm{d}r,$$

$$\frac{1}{nh}\sum_{t=1}^{n}\left(\frac{z_t-z}{h}\right)^iK\left(\frac{z_t-z}{h}\right)w_t'$$
$$\xrightarrow{P}\mathrm{E}\left[\frac{1}{h}\left(\frac{z_t-z}{h}\right)^iK\left(\frac{z_t-z}{h}\right)w_t'\right]$$
$$\approx\int u^iK(u)\,\mathrm{d}uf_z(z)\mathrm{E}(w'|z)=f_z(z)\mu_i(K)\mu_w(z)'.$$

If w and z are independent, this term is zero.

Next,

$$\frac{1}{n^2h}\sum_{t=1}^{n}\left(\frac{z_t-z}{h}\right)^iK\left(\frac{z_t-z}{h}\right)x_tx_t'$$
$$\approx\left[\mathrm{E}\frac{1}{h}\left(\frac{z_t-z}{h}\right)^iK\left(\frac{z_t-z}{h}\right)\right]\frac{1}{n^2}\sum_{t=1}^{n}x_tx_t'$$
$$\to\mu_i(K)f_z(z)\int B_xB_x{}'.$$

Using a similar argument as Hansen (1992), we have

$$\frac{1}{n^{3/2}h}\sum_{t=1}^{n}\left(\frac{z_t-z}{h}\right)^iK\left(\frac{z_t-z}{h}\right)x_tw_t'$$
$$\approx\mu_i(K)f_z(z)\int B_2(r)\mathrm{d}r\mu_w(z)'.$$

Finally,

$$\frac{1}{nh}\sum_{t=1}^{n}\left(\frac{z_t-z}{h}\right)^iK\left(\frac{z_t-z}{h}\right)w_tw_t'$$
$$\to\mathrm{E}\left(\frac{1}{h}\left(\frac{z_t-z}{h}\right)^iK\left(\frac{z_t-z}{h}\right)w_tw_t'\right)$$
$$=\mu_i(K)f_z(z)\Sigma_w(z).$$

Thus

$$\frac{1}{nh}\sum_{t=1}^{n}K_{h,t}D_nX_tX_t'D_n\left(\frac{z_t-z}{h}\right)^i$$
$$\Rightarrow f_z(z)\mu_i(K)\begin{bmatrix}1&\int B_2(r)'\mathrm{d}r&\mu_w(z)'\\\int B_2(r)\mathrm{d}r&\int B_xB_x{}'&\int B_2(r)\mathrm{d}r\mu_w(z)'\\\mu_w(z)&\mu_w(z)\int B_2(r)\mathrm{d}r&\Sigma_w(z)\end{bmatrix}$$
$$=f_z(z)\mu_i(K)Q.$$

Thus

$$S_n\Rightarrow f_z(z)M(K)\otimes Q.$$

For the bias term,

$$B_n=\frac{1}{nh}H^{-1}\sum_{t=1}^{n}K_{h,t}\mathbf{X}_t(z)X_t'D_n\left[\frac{D_n^{-1}\theta^{(P+1)}(z)}{(P+1)!}(z_t-z)^{P+1}\right]$$
$$=h^{P+1}\begin{bmatrix}\frac{1}{nh}\sum_{t=1}^{n}K_{h,t}D_nX_tX_t'D_n\left(\frac{z_t-z}{h}\right)^{P+1}\frac{D_n^{-1}\theta^{(P+1)}(z)}{(P+1)!}\\\frac{1}{nh}\sum_{t=1}^{n}K_{h,t}D_nX_tX_t'D_n\left(\frac{z_t-z}{h}\right)^{P+2}\frac{D_n^{-1}\theta^{(P+1)}(z)}{(P+1)!}\\\cdots\\\frac{1}{nh}\sum_{t=1}^{n}K_{h,t}D_nX_tX_t'D_n\left(\frac{z_t-z}{h}\right)^{2P+1}\frac{D_n^{-1}\theta^{(P+1)}(z)}{(P+1)!}\end{bmatrix}$$

$$\frac{1}{nh}\sum_{t=1}^{n}\begin{bmatrix} K_{h,t}D_nX_tX_t'D_n & K_{h,t}D_nX_tX_t'D_n\left(\frac{z_t-z}{h}\right) & \cdots & K_{h,t}D_nX_tX_t'\left(\frac{z_t-z}{h}\right)^{P} \\ K_{h,t}D_nX_tX_t'D_n\left(\frac{z_t-z}{h}\right) & K_{h,t}D_nX_tX_t'D_n\left(\frac{z_t-z}{h}\right)^{2} & & K_{h,t}D_nX_tX_t'D_n\left(\frac{z_t-z}{h}\right)^{P+1} \\ \cdots & & \ddots & \\ K_{h,t}D_nX_tX_t'D_n\left(\frac{z_t-z}{h}\right)^{P} & K_{h,t}D_nX_tX_t'D_n\left(\frac{z_t-z}{h}\right)^{P+1} & & K_{h,t}D_nX_tX_t'D_n\left(\frac{z_t-z}{h}\right)^{2P} \end{bmatrix}$$

Box I.

$$\frac{1}{nh}\sum_{t=1}^{n}K_{h,t}D_nX_tX_t'D_n\left(\frac{z_t-z}{h}\right)^{i}$$
$$=\sum_{t=1}^{n}\begin{bmatrix} \frac{1}{nh}\left(\frac{z_t-z}{h}\right)^{i}K\left(\frac{z_t-z}{h}\right) & \frac{1}{n^{3/2}h}\left(\frac{z_t-z}{h}\right)^{i}K\left(\frac{z_t-z}{h}\right)x_t' & \frac{1}{nh}\left(\frac{z_t-z}{h}\right)^{i}K\left(\frac{z_t-z}{h}\right)w_t' \\ \frac{1}{n^{3/2}h}\left(\frac{z_t-z}{h}\right)^{i}K\left(\frac{z_t-z}{h}\right)x_t & \frac{1}{n^{2}h}\left(\frac{z_t-z}{h}\right)^{i}K\left(\frac{z_t-z}{h}\right)x_tx_t' & \frac{1}{n^{3/2}h}\left(\frac{z_t-z}{h}\right)^{i}K\left(\frac{z_t-z}{h}\right)x_tw_t' \\ \frac{1}{nh}\left(\frac{z_t-z}{h}\right)^{i}K\left(\frac{z_t-z}{h}\right)w_t & \frac{1}{n^{3/2}h}\left(\frac{z_t-z}{h}\right)^{i}K\left(\frac{z_t-z}{h}\right)w_tx_t' & \frac{1}{nh}\left(\frac{z_t-z}{h}\right)^{i}K\left(\frac{z_t-z}{h}\right)w_tw_t' \end{bmatrix}$$

Box II.

$$\Rightarrow h^{P+1}\begin{bmatrix} \frac{f_z(z)}{(P+1)!}\mu_{P+1}(K)QD_n^{-1}\theta^{(P+1)}(z) \\ \frac{f_z(z)}{(P+1)!}\mu_{P+2}(K)QD_n^{-1}\theta^{(P+1)}(z) \\ \frac{f_z(z)}{(P+1)!}\mu_{2P+1}(K)QD_n^{-1}\theta^{(P+1)}(z) \end{bmatrix}$$
$$= h^{P+1}\frac{f_z(z)}{(P+1)!}\left(B\otimes QD_n^{-1}\theta^{(P+1)}(z)\right)$$

where

$$B=\begin{bmatrix} \mu_{P+1}(K) \\ \mu_{P+2}(K) \\ \cdots \\ \mu_{2P+1}(K) \end{bmatrix}.$$

Finally

$$U_n = \frac{1}{\sqrt{nh}}\left(H_P\otimes D_n^{-1}\right)^{-1}\sum_{t=1}^{n}K_{h,t}\mathbf{X}_t(z)\varepsilon_t$$
$$=\begin{bmatrix} \frac{1}{\sqrt{nh}}\sum_{t=1}^{n}K_{h,t}D_nX_t\varepsilon_t \\ \frac{1}{\sqrt{nh}}\sum_{t=1}^{n}K_{h,t}D_nX_t\varepsilon_t\left(\frac{z_t-z}{h}\right) \\ \cdots \\ \frac{1}{\sqrt{nh}}\sum_{t=1}^{n}K_{h,t}D_nX_t\varepsilon_t\left(\frac{z_t-z}{h}\right)^{P} \end{bmatrix}$$
$$\Rightarrow MN\left(0, f_z(z)\sigma_\varepsilon^2 N(K)\otimes Q\right)$$

since

$$\frac{1}{\sqrt{nh}}\sum_{t=1}^{n}K_{h,t}D_nX_t\varepsilon_t \Rightarrow MN\left(0,\, \nu_0(K)f_z(z)\sigma_\varepsilon^2 Q\right)$$

$$\frac{1}{\sqrt{nh}}\sum_{t=1}^{n}D_nX_tK\left(\frac{z_t-z}{h}\right)\left(\frac{z_t-z}{h}\right)^{i}\varepsilon_t$$
$$=\frac{1}{\sqrt{nh}}\sum_{t=1}^{n}\begin{pmatrix} K\left(\frac{z_t-z}{h}\right)\left(\frac{z_t-z}{h}\right)^{i}\varepsilon_t \\ n^{-1/2}x_tK\left(\frac{z_t-z}{h}\right)\left(\frac{z_t-z}{h}\right)^{i}\varepsilon_t \\ w_tK\left(\frac{z_t-z}{h}\right)\left(\frac{z_t-z}{h}\right)^{i}\varepsilon_t \end{pmatrix}$$
$$=\begin{pmatrix} \frac{1}{\sqrt{nh}}\sum_{t=1}^{n}\underline{K}_t\varepsilon_t \\ \frac{1}{n\sqrt{h}}\sum_{t=1}^{n}x_t\underline{K}_t\varepsilon_t \\ \frac{1}{\sqrt{nh}}\sum_{t=1}^{n}w_t\underline{K}_t\varepsilon_t \end{pmatrix}+\begin{pmatrix} \frac{1}{\sqrt{nh}}\sum_{t=1}^{n}\varepsilon_t\mathrm{E}K\left(\frac{z_t-z}{h}\right) \\ \frac{1}{n\sqrt{h}}\sum_{t=1}^{n}x_t\varepsilon_t\mathrm{E}K\left(\frac{z_t-z}{h}\right) \\ \frac{1}{\sqrt{nh}}\sum_{t=1}^{n}w_t\varepsilon_t\mathrm{E}K\left(\frac{z_t-z}{h}\right) \end{pmatrix}$$
$$\Rightarrow MN\left(0,\frac{\sigma_\varepsilon^2}{f_z(z)}\left(\mathcal{M}(\mathcal{K})\otimes Q^{-1}\right)\right). \quad \blacksquare$$

References

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

Campbell, J.Y., Shiller, R.J., 1987. Cointegration and tests of present value models. Journal of Political Economy 95 (5), 1062–1088.

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

Cerchi, M., Havenner, A., 1988. Cointegration and stock prices. Journal of Economic Dynamics and Control 12, 333–346.

Chowdhury, A.R., 1991. Futures market efficiency: Evidence from cointegration tests. The Journal of Future Markets 11 (5), 577–589.

Davidson, J., 1994. Stochastic Limit Theory. Oxford University Press..

Diba, B., Grossman, H., 1988. Explosive rational bubbles in stock prices. American Economic Review 78, 520–530.

Engle, R.F., Granger, C.W.J., 1987. Cointegration and error correction: Representation, estimation, and testing. Econometrica 55, 251–276.

Fan, J., 1992. Design-adaptive nonparametric regression. Journal of the American Statistical Association 87, 998–1004.

Fan, J., Gijbels, I., 1996. Local Polynomial Modelling and Its Applications. Champman & Hall..

Evans, G., 1991. Pitfalls in testing for explosive bubbles in asset prices. American Economic Review 81 (4), 922–930.

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

Harris, D., Mccabe, B., Laybourne, S., 2002. Stochastic cointegration: Estimation and inference. Journal of Econometrics 111, 363–384.
Hendry, D., 1996. Dynamic Econometrics. Oxford University Press.
Hansen, B., 1992. Testing for parameter instability in regressions with I(10) processes. Journal of Business and Economic Statistics 10, 321–335.
Juhl, T., 2005. Functional-Coeffcient models under unit root behavior. Econometrics Journal 8, 197–213.
Kwiatkowski, D., Phillips, P.C.B., Schmidt, P., 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, 159–178.
Park, J., Hahn, S., 1999. Cointegrating regressions with time varying coefficients. Econometric Theory 15, 664–703.
Quintos, C., Phillips, P.C.B., 1993. Parameter constancy in cointegrating regressions. Empirical Economics 18, 675–706.
Qu, Z., 2005. Estimating and testing for structural changes in multivariate regressions. Ph.D. Dissertation, Boston University.
Saikkonen, P., 1991. Asymptotically efficient estimation of cointegration regression. Econometric Theory 7, 1–21.
Schwert, G.W., 1989. Tests for unit roots: A Monte Carlo investigation. Journal of Business and Economic Statistics 7, 147–159.
Xiao, Z., Phillips, P.C.B., 2002. A CUSUM test for cointegration using regression residuals. Journal of Econometrics 108, 43–61.
Xiao, Zhijie, 2009. Quantile cointegrating regression. Journal of Econometrics 150 (2), 248–260.