OPERATIONS RESEARCH AND DECISIONS

No. 3–4
DOI: 10.5277/ord1203-0403 2011

Manohar B. RAJARSHI*
Thekke V. RAMANATHAN*
Chanchala A. GHADGE*

RANK BASED TESTS FOR TESTING THE CONSTANCY OF THE REGRESSION COEFFICIENTS AGAINST RANDOM WALK ALTERNATIVES

A class of approximately locally most powerful type tests based on ranks of residuals is suggested for testing the hypothesis that the regression coefficient is constant in a standard regression model against the alternatives that a random walk process generates the successive regression coefficients. We derive the asymptotic null distribution of such a rank test. This distribution can be described as a generalization of the asymptotic distribution of the Cramer-von Mises test statistic. However, this distribution is quite complex and involves eigen values and eigen functions of a known positive definite kernel, as well as the unknown density function of the error term. It is then natural to apply bootstrap procedures. Extending a result due to Shorack in [25], we have shown that the weighted empirical process of residuals can be bootstrapped, which solves the problem of finding the null distribution of a rank test statistic. A simulation study is reported in order to judge performance of the suggested test statistic and the bootstrap procedure.

Keywords: *bootstrap, random coefficient regression models, random walk alternative models, rank tests, weighted empirical and rank processes*

1. Introduction

Consider the regression model

$$Y_t = \beta x_t + \varepsilon_t : t = 1, 2, \ldots \tag{1}$$

where $\{ \varepsilon_t : t = 1, 2, \ldots \}$ forms a sequence of independently and identically distributed (iid) random variables having a location-scale family of density $(1/\sigma_\varepsilon) f((\varepsilon - \mu_\varepsilon)/\sigma_\varepsilon)$

*Department of Statistics, University of Pune, Pune 411 007, India; e-mail addresses: m.b.rajarshi@gmail.com; ram@stats.unipune.ac.in;cghadge04@yahoo.com

with $\mu_\varepsilon = E(\varepsilon_t) = 0$ and $E(\varepsilon_t^2) = h_1(\sigma_\varepsilon) < \infty$, $h_1(.)$ being a nonnegative function of σ_ε. The assumption that the regression coefficient β remains constant over time may not be true always. Stochastic variation of the parameter β over time can be of various types. A number of varying parameter models have been proposed in the literature. [28, 18] give a brief survey of some of these models. One of the important modes of variation that has been extensively discussed in the literature is where the regression coefficients $\{\beta_t\}$ vary according to a random walk process. To be more specific, consider the model

$$Y_t = \beta_t x_t + \varepsilon_t, \qquad \beta_t = \beta_{t-1} + u_t, \qquad t = 1, 2, \ldots \tag{2}$$

where $\{u_t\}$ is a sequence of iid random variables having a location-scale family of density $(1/\sigma_u)f((u - \mu_u)/\sigma_u)$ with $\mu_u = E(u_t) = 0$, $E(u_t^2) = h_2(\sigma_u)$, $h_2(.)$ being a non-negative function of σ_u such that $0 < h_2(\sigma_u) < \infty$, and $E(u_t^2) = 0$ whenever $\sigma_u = 0$. Further, $\beta_0 = \beta$ is assumed to be nonrandom. The two sequences of random variables viz., $\{\beta_t\}$ and $\{\varepsilon_t\}$ are assumed to be independent. This model describes a situation, wherein there is a gradual and smooth change in the regression parameter from a time unit to the next. Model (2) belongs to the class of state space models which have been widely discussed in the literature. Note that when $\sigma_u = 0$, model (2) reduces to model (1).

A natural problem of interest is to test whether variation of the (random) regression coefficients $\{\beta_t\}$ is significant. In other words, we would like to test the hypothesis H_0: $\sigma_u = 0$ against the alternatives H_1: $\sigma_u > 0$. These types of tests are generally known as specification tests, whose history dates back to the works of Ramsey in [23] and Hausman in [8]. For simplifying the notation, while constructing the test statistic, we denote the variance of u_t as σ_u^2 instead of $h_2(\sigma_u)$.

Assuming that ε_t and u_t follow normal distributions, a number of tests have been suggested by various authors. It was Cooley and Prescott who first looked into this problem we refer [5, 6] for further references. LaMotte and McWhorter [13] have constructed an exact test for this hypothesis. Nyblom and Makelainen [19] have obtained a locally most powerful invariant test. The asymptotic null distribution theory of the locally most powerful invariant test is somewhat complicated: they handled a special case when all x_t's are identically equal to one. Nabeya and Tanaka [16] have shown that the limiting null distribution of the locally best invariant test statistic is closely related to that of Cramer–von Mises statistics, and it heavily depends on the values of the regressors x_t, $t = 1, 2, \ldots$ Even though, their test statistics is developed under the assumption of normality of ε_t and u_t, the limiting distribution theory does not require this assumption. Nabeya [15] has considered the limiting distribution under various sequences of alternatives that converge to the null. Also see [10, 17] for some of the related references. Another closely related reference is that of Shively [24], who develops an exact test for the same problem under the assumption of normality by specifying the value of $\sigma_u/\sigma_\varepsilon$ under the alternative. Rajarshi and Ramanathan [21] developed a test procedure for testing the constancy of a parameter of a Markov sequence against the alternatives that the parameter varies over time according in a random walk manner.

All the above mentioned tests and the corresponding distribution theories have been developed under the assumptions of normality. In this paper, we develop a rank test procedure for the above mentioned problem. We suggest a class of rank tests (which, in some cases, may be approximately locally most powerful rank tests) and derive their null distribution.

In Section 2, we derive a class of rank tests for testing H$_0$: $\sigma_u^2 = 0$ against H$_1$: $\sigma_u^2 > 0$. The asymptotic null distribution of a rank test is obtained in Section 3. Since the distribution of the test statistic is extremely complicated, we suggest that the test statistic can be bootstrapped. Discussion of validity of the suggested bootstrap procedure forms Section 4. An extensive simulation study is carried out in Section 5 to check the performance of these tests. Some concluding remarks are given in Section 6. All the proofs have been deferred to Section 7.

2. A class of rank tests

The locally most powerful invariant test for the hypothesis H$_0$: $\sigma_u^2 = 0$ against H$_1$: $\sigma_u^2 > 0$ in the context of model (2) has been discussed by Nyblom and Makelainen [19] and Nabeya and Tanaka [16].

Let $\boldsymbol{\varepsilon} = (\varepsilon_1, \varepsilon_2, ..., \varepsilon_T)$ be the error vector, $\mathbf{D}_X$ and $\mathbf{A}_T$ are $T \times T$ matrices, defined by

$$\mathbf{D}_X = \operatorname{diag}(x_1, x_2, ..., x_T) \qquad \text{and} \qquad \mathbf{A}_T = \min(i, j) \tag{3}$$

Let r_t be the least squares residual, $t = 1, 2, ..., T$. If ε_t and u_t are normally distributed, the locally most powerful invariant test is given by

$$S = \frac{\hat{\boldsymbol{\varepsilon}}'\mathbf{D}_X\mathbf{A}_T\mathbf{D}_X\hat{\boldsymbol{\varepsilon}}}{c(T)} \tag{4}$$

where $\hat{\boldsymbol{\varepsilon}} = (r_1, r_2, ..., r_T)'$ and $c(T)$ is an appropriate scaling factor.

To derive the locally most powerful rank test for H_0, for the time being, we assume that β is known. Let f be the probability density function of ε_1 and F be the corresponding distribution function. It is assumed that the distribution of ε_1 belongs to a location-scale family of distributions, with location 0 and scale σ_ε. Let $\boldsymbol{\beta}_T = (\beta_1, \beta_2, ..., \beta_T)'$. The joint probability density function of $\mathbf{y} = (y_1, y_2, ..., y_T)'$ under H_1, $L_1(\mathbf{y})$ (say), is given by

$$L_1(\mathbf{y}) = \int_{R^T} \prod_{i=1}^{T} f(y_i - \beta_i x_i)\, dG(\boldsymbol{\beta}_T)$$

where G is the joint distribution function of $\boldsymbol{\beta}_T$. Expanding $\prod_{i=1}^{T} f(y_i - \beta_i x_i)$ around $\prod_{i=1}^{T} f(y_i - \beta x_i)$ by the multivariate Taylor series expansion, we have

$$L_1(\mathbf{y}) = L_0(\mathbf{y}) \left\{ 1 + \frac{\sigma_u^2}{2} \sum_{i=1}^{T} \sum_{j=1}^{T} \min(i,j)\, x_i x_j h_i h_j + \frac{\sigma_u^2}{2} \sum_{i=1}^{T} i x_i^2 \left\{ \frac{\partial^2 \ln f(u)}{\partial u^2} \right\}_{u=\varepsilon_i} \right\} + \Delta \quad (5)$$

where $L_0(\mathbf{y})$ is the joint probability density function of $\mathbf{y}$ under H_0, $h_i(u) = \frac{\partial \ln f(u)}{\partial u}$ at $u = \varepsilon_i$ and Δ is the reminder term.

Under the usual regularity conditions on the probability density function f, along with the assumption that $E|(\beta_i - \beta)(\beta_j - \beta)(\beta_k - \beta)| < \infty$ for all i, j and k, it can be shown that $\frac{\partial \Delta}{\partial \sigma_u^2} \xrightarrow{p} 0$ as $\sigma_u^2 \xrightarrow{p} 0$ (cf. [29]). Further, assuming that the scaling factor $c(T)$ exists such that

$$\frac{1}{\{Tc(T)\}} \sum_{i=1}^{T} i x_i^2 \left\{ \frac{\partial^2 \ln f(u)}{\partial u^2} \right\}_{u=\varepsilon_i} \xrightarrow{P} \text{const} \quad (6)$$

this term can be ignored while constructing the test statistics (cf. Remark 3.1). Thus,

$$L_1(\mathbf{y}) \cong L_0(\mathbf{y}) \left\{ 1 + \frac{\sigma_u^2}{2} \sum_{i=1}^{T} \sum_{j=1}^{T} \min(i,j)\, x_i x_j h_i h_j \right\} \quad (7)$$

Let $\mathbf{R} = (R_1, R_2, ..., R_T)'$, where R_i is the rank of $\varepsilon_i = y_i - \beta x_i$ among $\varepsilon_1, \varepsilon_2, ..., \varepsilon_T$. Let $\boldsymbol{\alpha}$ be a permutation of $\{1, 2, ..., T\}$. Now we consider

$$P_1[\mathbf{R} = \boldsymbol{\alpha}] = \int_{\{\varepsilon_j = \varepsilon_{(\alpha_j)}\}} L_1(\boldsymbol{\varepsilon})\, d(\boldsymbol{\varepsilon}) \quad (8)$$

Upon using (7) and (8) and also the standard approximation

$$E\frac{f'\left(\varepsilon_{(i)}\right)}{f\left(\varepsilon_{(i)}\right)} \cong \frac{f'\left(F^{-1}\left(\frac{i}{T+1}\right)\right)}{f\left(F^{-1}\left(\frac{i}{T+1}\right)\right)}$$

(cf. [2], p. 189), we have

$$\left\{\frac{\partial \ln P_1(\mathbf{R}=\boldsymbol{\alpha})}{\partial u}\right\}_{\sigma_u^2=0} \cong \sum_{i=1}^{T}\sum_{j=1}^{T}\min(i,j)\,x_i x_j \left\{h_i(u)\Big|_{u=F^{-1}\left(\frac{R_i}{T+1}\right)}\right\}\times\left\{h_j(u)\Big|_{u=F^{-1}\left(\frac{R_j}{T+1}\right)}\right\} \quad (9)$$

Let ϕ be a score generating function which is assumed to be the Riemann integrable over [0, 1]. We do not assume that β is estimated by the least squares estimator. In general, let $\hat{\beta}$ be any estimator of β. Exact conditions on the estimation procedure of β will be specified in Section 3. Consider the vector

$$\mathbf{V} = \left(V_1, V_2, ..., V_T\right)', \qquad V_i = \left\{\phi\left(\frac{R_i}{T+1}\right) - \bar{\phi}_T\right\} \quad (10)$$

R_i being the rank of r_i among $r_1, r_2, ..., r_T$, and $\bar{\phi}_T = \frac{1}{T}\sum_{i=1}^{T}\phi\left(\frac{R_i}{T+1}\right)$. Based on (9), we propose the class of rank tests $S_T(\phi)$ defined by

$$S_T(\phi) = \sum_{i=1}^{T}\sum_{j=1}^{T}\min(i,j)\,x_i x_j \left\{\phi\left(\frac{R_i}{T+1}\right) - \bar{\phi}_T\right\}\left\{\phi\left(\frac{R_j}{T+1}\right) - \bar{\phi}_T\right\} = \mathbf{V}'\mathbf{D}_X\mathbf{A}_T\mathbf{D}_X\mathbf{V} \quad (11)$$

where ϕ has been conveniently suppressed in the matrix notation on the right hand side. It would be shown in the next section that $S_T(\phi)/\{Tc(t)\}$ converges in distribution to a random variable which justifies (6) and subsequently (7). For example, if the error distribution is logistic, $\phi(u)$ is given by $(2u - 1)$ (see [2], p. 189).

3. Asymptotic null distribution

Throughout this section, we assume that the null hypothesis holds. All probability statements would refer to the probability distribution of $\{y_t\}$, as defined in (1). Let $c_1, c_2, ..., c_T$ be constants and let $\gamma(\mathbf{c}) = \gamma(c_1, c_2, ..., c_T) = \left(\sum_{t=1}^{T} c_t^2\right)^{-1/2}$.

We assume throughout that $\max_{1\le t\le T}\left[\{\gamma(\mathbf{c})\}^2 c_t^2\right]\to 0$ as $T\to\infty$, $\sum_{t=1}^{T} c_t = 0$, and $\max_{1\le t\le T}\left[\{\gamma(\mathbf{x})\}^2 x_t^2\right]\to 0$ as $T\to\infty$. Define the weighted rank process $R_T(t)$ by

$$R_T(t) = \gamma(\mathbf{c})\sum_{i=1}^{T} c_i I_{\left[R_i\le (T+1)t\right]}, \qquad 0 < t \le 1 \tag{12}$$

(cf. [26], p. 90), where R_i is the rank of r_i among $r_1, r_2, \ldots, r_T$ with $r_i = y_i - \hat{\beta}x_i$ and $I_{[\cdot]}$ denotes the indicator function. Towards deriving the asymptotic distribution of the test statistic under H_0, we need the following:

Lemma 3.1. Let W denote a Brownian bridge. Let $\hat{\beta}$ be an estimator of β such that
(i) $\{\gamma(\mathbf{x})\}^{-1}(\hat{\beta}-\beta) = O_p(1)$,
(ii) there exists a function g such that

$$\left\{\left\{\frac{\gamma(\mathbf{c})}{\sigma_\varepsilon}\right\}\left(\sum_{i=1}^{T} c_i x_i\left(\hat{\beta}-\beta\right)\right)\right\} - \left\{\int_0^1 g\left(F^{-1}\right)dW\right\} \xrightarrow{P} 0$$

where g is a continuous function such that $g(F^{-1})$ is square integrable with respect to the Lebesgue measure,
(iii) the functions f and $yf(y)$ are absolutely continuous on R and
(iv) $f\cdot F^{-1}$ and $F^{-1}f\cdot F^{-1}$ are absolutely continuous on [0, 1].

Suppose that the density f is bounded. Then under H_0

$$\left\| R_T - \left\{W + \left(f\cdot F^{-1}\right)\int_0^1 g\left(F^{-1}\right)dW\right\}\right\| \xrightarrow{P} 0$$

where $f\cdot F^{-1}$ denotes $f(F^{-1}(u))$.

It is clear from the above lemma that the weighted rank process R_T converges weakly to a Gaussian process. Hereafter, we denote this Gaussian process by $\mathbf{X}$.

Lemma 3.2. Let $\psi(t)$ be a continuous, real valued function on [0, 1] such that

$$\tau^2 = \int_0^1 \psi^2(t)\,dt - \left\{\int_0^1 \psi(t)\,dt\right\}^2, \qquad 0 < \tau^2 < \infty$$

Further, let ϕ be a left continuous, square integrable function of bounded variation. Then under H_0 and under the assumptions of Lemma 3.1

$$L\left\{T^{-1/2}\sum_{i=1}^{T}\psi\left(\frac{i}{T}\right)V_i\right\} \to N\left(0, \sigma^2\tau^2\right)$$

where $\sigma^2 = \mathrm{Var}\left(\int_0^1 \phi(t)\, d\mathbf{X}(t)\right)$ and L denotes a probability law.

We note that σ^2 in the above lemma depends on the score function ϕ and the unknown density f.

Lemma 3.3. Let $\psi_1(t)$, $\psi_2(t)$, ..., $\psi_n(t)$ be continuous real valued functions on [0, 1] such that

$$\int_0^1 \psi_k(t)\psi_l(t)\, dt = \delta_{kl} \qquad \text{(Kronecker's delta)}$$

Then, under the assumptions of Lemma 3.2 and under H_0

$$L\left\{\left[T^{-1/2}\sum_{i=1}^{T}\psi_j\left(\frac{i}{T}\right)V_i\right]_{j=1,\,2,\,\ldots,\,n}\right\} \to N\left(0, \sigma^2\boldsymbol{\sigma}_n\right)$$

where the (j, k)-th element of $\boldsymbol{\sigma}_n$ is given by

$$\tau_{jk} = \delta_{jk} - \left\{\int_0^1 \psi_j(t)dt\right\}\left\{\int_0^1 \psi_k(t)dt\right\}$$

Lemma 3.4. Let $\mathbf{H}_T$ be a $T \times T$ real symmetric matrix such that $|\mathbf{H}_T(i, j)| \le \delta$ for $i, j = 1, 2, ..., T$. If ϕ is a bounded function, then

$$E\left(\left|\frac{1}{T}\mathbf{V}'\mathbf{H}_T\mathbf{V}\right|\right) \le K\delta$$

where K is a constant free of T.

Now let $K(s, t)$ be a continuous symmetric function (or kernel) defined on $[0, 1] \times [0, 1]$ (that is, $K(s, t) = K(t, s)$ for all t and s) and suppose that it is positive definite, in the sense that

$$\int_0^1\int_0^1 K(s,t)\psi(s)\psi(t)\,ds\,dt > 0 \tag{13}$$

for all continuous functions defined on [0, 1]. Then the kernel $K(s, t)$ can be decomposed in the form

$$K(s,t) = \sum_{j=1}^{\infty} \lambda_j \psi_j(s)\psi_j(t), \qquad 0 \le s, t \le 1 \tag{14}$$

where $0 < \lambda_1 \le \lambda_2 \le \dots$ are eigen values and $\psi_1(t)$, $\psi_2(t)$, ... are the corresponding eigen functions of the kernel $K(s, t)$, defined by the relationship

$$\int_0^1 \psi(\mathrm{s})K(s,t)\,ds = \lambda\psi(t), \qquad 0 \le t \le 1 \tag{15}$$

By Mercer's theorem ([26], p. 208), the series (14) converges uniformly and absolutely in $(s, t) \subset [0, 1] \times [0, 1]$. Also, $\sum_{i=1}^{\infty} \lambda_i < \infty$.

With this background, we prove the first of our main results.

Theorem 3.1. Let $\mathbf{B}_T = \left(c(T)^{-1}\right)\mathbf{D}_X\mathbf{A}_T\mathbf{D}_X$. Consider a kernel $K(s, t)$ defined on $[0, 1] \times [0, 1]$, which is of the type described through (13)–(15), such that

$$\lim_{T\to\infty} \max_{i,j} \left| B_T(i,j) - K\left(\frac{i}{T}, \frac{j}{T}\right)\right| = 0 \tag{16}$$

Then, under the assumptions of Lemmas 3.1–3.3 along with the additional assumption that the score generating function ϕ is bounded

$$\frac{L\left(S_T(\phi)\right)}{\{Tc(T)\}} \to L\left(\sum_{i=1}^{\infty} \lambda_i Z_i^2\right) \tag{17}$$

where λ_i's are the eigen values associated with the kernel $K(s, t)$ and $\{Z_i\}_{i=1}^{\infty}$ is a Gaussian sequence with $E(Z_i) = 0$ for each i and Σ_n as the variance-covariance matrix of $(Z_1, Z_2, \dots, Z_n)$ for each n.

Remark 3.1. The condition (16) has been verified by Nabeya and Tanaka [16] and Nabeya [15] for various choices of regressors. The condition (6) is naturally satisfied in situations discussed by them along with the additional assumptions

$$\operatorname{Var}\left\{\left.\frac{\partial^2 \ln f(u)}{\partial u^2}\right|_{u=\varepsilon_1}\right\} < \infty \qquad \text{and} \qquad \sum_{i=1}^{\infty} \lambda_i^2 < \infty$$

In view of the fact that $\{ix_i^2\} = c(T)\, B_T(i, i)$, it can be shown that (6) converges to

$$E\left\{\left.\frac{\partial^2 \ln f(u)}{\partial u^2}\right|_{u=\varepsilon_1}\right\} \int_0^1 K(t, t) dt, \qquad \text{in probability}$$

since the variance of the left hand side of (7) is approximately

$$\frac{1}{T} \operatorname{Var}\left\{\left.\frac{\partial^2 \ln f(u)}{\partial u^2}\right|_{u=\varepsilon_1}\right\} \int_0^1 \{K(t, t)\}^2 dt$$

for large values of T.

Remark 3.2. Since $E(Z_l^2) = \sigma^2 \tau_l^2 = \sigma^2 \left\{1 - \left[\int_0^1 \psi_l(t)\, dt\right]^2\right\} \le \sigma^2$, for all l =1, 2, ... and $\sum_{i=1}^{\infty} \lambda_i < \infty$, it is clear that the random variable on the right hand side of (17) is a proper random variable. Further, it is trivial to establish the existence of a Gaussian sequence $\{Z_i\}$, as described in Theorem 3.1.

Remark 3.3. In case the eigen functions $\psi(t)$ are such that $\int_0^1 \psi(t)\, dt = 0$, it follows from Lemma 3.3 that the random variables Z_1, Z_2, ..., Z_n are iid. However, since the variance of Z_i's depends on the unknown density f, even asymptotically the statistic $S_T(\phi)$ is not distribution free. It may be possible to estimate $\sigma^2 = \operatorname{Var}\left\{\int_0^1 \phi(t)\, d\mathbf{X}(t)\right\}$ by estimating the density function f. Since we adopt a bootstrap procedure to estimate the null distribution of $S_T(\phi)$, we do not pursue this further.

Remark 3.4. It has been shown [22] that the weighted empirical process of square of the residuals converges weakly to a Brownian bridge, when the distribution of ε_1 is symmetric around zero. Consequently, the weighted empirical rank process of the squares of the residuals also converges weakly to a Brownian bridge. Hence, $\sigma^2 = \left\{\operatorname{Var} \int_0^1 \phi(t)\, dW(t)\right\}$ is completely known and a statistic of the type $S_T(\phi)$ based on the ranks of the squares can be shown to be asymptotically distribution free. However, if $\int_0^1 \psi(t)\,dt \neq 0$, one needs to compute the eigen functions of the kernel $K(s, t)$.

Remark 3.5. It may be possible to explore the relationship of the suggested rank statistic to a generalized U-statistics. In such a case, there are several results available in the literature, which may be useful in proving the asymptotics (see [11, 14]). However we are not pursuing in that direction here.

In practice, applying the test statistic $S_T(\phi)$ faces some difficulties. Firstly, one needs to compute the eigen values and eigen functions of the kernel $K(s, t)$. Even if these are available, the actual cut-off point cannot be computed since the test statistic is not asymptotically distribution free. Bootstrap procedures naturally come very handy in situations of this type [4, 5]. In the next section, we describe a bootstrap procedure and establish its validity.

4. Bootstrapping rank tests

Applications of bootstrap procedures in hypothesis testing and computation of the p-value have been discussed by many authors. We refer to [1, 3, 7, 9] for more details. It is pointed out therein that for bootstrapping the test statistic, bootstrap observations should be drawn from the model specified by the null hypothesis. We describe below such a bootstrap procedure together with its validity. It is assumed throughout that the model (1) holds.

Let $\bar{r}$ be the mean of the residuals r_i's and let $r_i' = (r_i - \bar{r})$. Let $b(T) = hT^{-1/5}$, where h is an appropriate constant and k be a suitable kernel (cf. [27], Chapter 3). Let

$$\hat{f}_T(x) = \frac{1}{\{Tb(T)\}} \sum_{i=1}^{T} \frac{k(x - r_i')}{b(T)}, \qquad \hat{F}_T(x) = \int_{-\infty}^{x} \hat{f}_T(u)\,du \tag{18}$$

be the estimates of f and F, respectively. Properties of $\hat{f}_T$ and $\hat{F}_T$ are described below in the following lemma.

Lemma 4.1. Let f and k satisfy the following conditions:

C1: f is uniformly continuous on R and the kernel k satisfies the following conditions.

C2: f and $yf(y)$ are absolutely continuous on R.

C3: $f \cdot F^{-1}$ and $F^{-1} f \cdot F^{-1}$ are absolutely continuous on [0, 1].

C4: k is bounded and of bounded variation on R.

C5: k is a uniformly continuous function.

Then, we have the following.

1. $\left\|\hat{f}_T - f\right\| \to 0,$ in probability, and

2. $\left\{P\left[\hat{f}_T \text{ is absolutely continuous}\right] \cap \left[y\hat{f}_T \text{ is absolutely continuous}\right]\right.$
$\cap\left[\hat{f}_T \cdot \hat{F}_T^{-1} \text{ is absolutely continuous}\right]$
$\left.\cap\left[\hat{F}_T^{-1}\hat{f}_T \cdot \hat{F}_T^{-1} \text{ is absolutely continuous}\right]\right\} \to 1 \text{ as } T \to \infty$

To bootstrap the test statistics $\dfrac{S_T(\phi)}{\{Tc(T)\}}$, we proceed as follows. Given the data (x_1, y_1), (x_2, y_2), ..., (x_T, y_T), let $\varepsilon_1^*, \varepsilon_2^*, ..., \varepsilon_T^*$ be iid random variables with distribution function $\hat{F}_T$. Let

$$y_i^* = \hat{\beta}x_i + \varepsilon_i^*, \qquad i = 1, 2, ..., T \tag{19}$$

Let

$$\{(x_1, y_1^*), (x_2, y_2^*), ..., (x_T, y_T^*)\}$$

form a bootstrap sample. Let $\hat{\beta}^*$ be the estimator of β based on a bootstrap sample. It is understood that $\hat{\beta}^*$ is computed from a bootstrap sample by the same procedure by which $\hat{\beta}$ was computed from the data. Let

$$r_i^* = y_i^* - \hat{\beta}^* x_i = \varepsilon_i^* - (\hat{\beta}^* - \hat{\beta})x_i, \qquad i = 1, 2, ..., T \tag{20}$$

be residuals of a bootstrap sample. Let c_i's be constants satisfying the condition described in the beginning of Section 3. Consider the weighted empirical process of residuals and its bootstrap version:

$$\begin{aligned} Z_T(x) &= \gamma(\mathbf{c})\left\{\sum_{i=1}^{T} c_i \left\{I_{[r_i \le x]} - F(x)\right\}\right\} \\ Z_T^*(x) &= \gamma(\mathbf{c})\left\{\sum_{i=1}^{T} c_i \left\{I_{[r_i^* \le x]} - \hat{F}_T(x)\right\}\right\} \end{aligned} \tag{21}$$

Let P^* denote the conditional probability distribution of (x_1, y_1^*), $(x_2, y_2^*), ..., (x_T, y_T^*)$ given the data.

We are now in a position to state our second main result, which finally leads to the asymptotic validity of the above bootstrap procedure for estimation of the null distribution of rank statistics (11).

Theorem 4.1. Suppose that the conditions of Lemma 4.1 are satisfied. Then there exists a Brownian bridge W_1 such that for every $\varepsilon > 0$

$$P^*\left[\left\| Z_T^* - \left\{ W_1(F) + f\int_0^1 g(F^{-1})dW_1 \right\} \right\| > \varepsilon \right] \xrightarrow{P} 0$$

The above result implies that the distribution of Z_T can be consistently estimated by the above bootstrap procedure. This further implies that the weighted rank process (12) can also be bootstrapped. Now, one proves an analogue of Theorem 3.1 for $\left[\frac{S_T^*(\phi)}{\{Tc(T)\}}\right]$. Therefore, the cut-off point of $S_T(\phi)$ (for a given level of significance) or the p-value can be consistently estimated by the bootstrap procedure for a large T. Further improvement over the estimation of the p-value may be obtained by nested bootstrap procedure, cf. Section 4 of [9].

5. Simulation study

In this section, we report an extensive simulation study to judge the performance of the suggested rank test and the bootstrap procedure. For various choices of distributions of the error term (ε) such as logistic, normal and Laplace, we have generated samples of size 50, 100 and 250. The number of bootstrap replications and the number of simulations were fixed at 1000. The location and scale parameters were taken as 0 and 1 for all these distributions. We have simulated data under the null and alternative hypothesis from specified distributions. The critical point was found using the bootstrap procedure explained in Section 4. Normal kernel is used for the estimation of the density function in (18) with the bandwidth chosen as $h = \min\{\hat{\sigma}_\varepsilon, \text{IQR}/1.05\}$, where $\hat{\sigma}_\varepsilon$ is the usual estimator of the Var(ε) and IQR is the inter-quartile range of residuals. The predictor x is generated from uniform (0,10). A Wilcoxon type score function is used in our computations. Table 1 shows that the suggested rank test maintains its level at 1%, 5% and 10% nominal levels for all three sample sizes as well as under all the three different choices of distributions for ε. In Table 2, we provide the power computations of the suggested rank test. It may be noted that the rank test is quite powerful, when u is distributed as normal, logistic and Laplace.

Table 1. Levels of significance of the rank test

n	$\varepsilon \sim$ Logistic			$\varepsilon \sim$ Normal			$\varepsilon \sim$ Laplace		
	1%	5%	10%	1%	5%	10%	1%	5%	10%
50	0.008	0.043	0.088	0.007	0.035	0.078	0.010	0.047	0.094
100	0.008	0.051	0.102	0.010	0.045	0.082	0.011	0.046	0.096
250	0.011	0.051	0.097	0.009	0.044	0.085	0.010	0.047	0.095

Table 2. Power of the rank test

n	$\varepsilon \sim$ Logistic			$\varepsilon \sim$ Normal			$\varepsilon \sim$ Laplace		
	1%	5%	10%	1%	5%	10%	1%	5%	10%
$u \sim$ Normal(0, 1)									
50	0.754	0.879	0.939	0.759	0.887	0.970	0.772	0.892	0.937
100	0.916	0.972	0.983	0.923	0.971	0.985	0.886	0.956	0.976
250	0.987	0.994	0.999	0.990	1.000	1.000	0.980	0.994	0.998
$u \sim$ Logistic(0, 1)									
50	0.748	0.858	0.919	0.761	0.888	0.924	0.778	0.892	0.934
100	0.910	0.970	0.993	0.920	0.973	0.990	0.915	0.966	0.991
250	0.995	1.000	1.000	0.980	0.996	1.000	0.993	0.999	1.000
$u \sim$ Laplace(0, 1)									
50	0.727	0.859	0.926	0.754	0.874	0.916	0.767	0.876	0.934
100	0.918	0.972	0.989	0.918	0.968	0.982	0.908	0.964	0.983
250	0.981	0.988	1	0.990	1	1	0.992	0.999	1.000

Table 3. Comparison of empirical levels of significance

n	1%	5%	10%	1%	5%	10%	1%	5%	10%
	$\varepsilon \sim$ Logistic			$\varepsilon \sim$ Normal			$\varepsilon \sim$ Laplace		
	Rank test								
50	0.011	0.049	0.117	0.013	0.052	0.095	0.011	0.058	0.100
100	0.009	0.058	0.108	0.004	0.039	0.084	0.004	0.039	0.108
250	0.009	0.052	0.102	0.011	0.053	0.104	0.006	0.050	0.103
LBI test									
50	0.006	0.044	0.105	0.006	0.040	0.088	0.004	0.043	0.107
100	0.009	0.044	0.099	0.005	0.035	0.077	0.002	0.031	0.074
250	0.009	0.440	0.085	0.007	0.048	0.093	0.003	0.042	0.089
	$\varepsilon \sim$ Cauchy(0, 1)			$\varepsilon \sim t(2)$			$\varepsilon \sim$ Std. $\chi^2(4)$		
	Rank test								
50	0.013	0.056	0.115	0.016	0.056	0.107	0.009	0.049	0.099
100	0.008	0.046	0.099	0.009	0.051	0.099	0.018	0.054	0.104
250	0.011	0.039	0.094	0.009	0.047	0.088	0.007	0.038	0.082
LBI test									
50	0.002	0.031	0.082	0.005	0.046	0.114	0.004	0.036	0.081
100	0.006	0.034	0.103	0.007	0.038	0.089	0.008	0.045	0.087
250	0.006	0.046	0.095	0.008	0.031	0.081	0.002	0.034	0.083

A comparison between two tests, that is, the rank test suggested in this paper and the LBI test suggested by Nabeya and Tanaka [16] is provided in Tables 3 and 4. Here, we consider $x = t$, as given in model (19) by Nabeya [15]. It may be noted that the performance of the rank test is quite good as compared to the best parametric test.

In order to stress the fact that our procedure does not require second order moment assumptions, as opposed to that of Nabeya and Tanaka [16], we have included in Tables 3 and 4 the computations pertaining to the distributions such as Cauchy (heavy tailed), t-distribution with 2 degrees of freedom (moment assumption violated) and standardized chi-square with 4 degrees of freedom (skewed). From Table 3 we see that under all these distributional choices of ε, the suggested rank test maintains its nominal level much better than the LBI test. Also from Table 4, we may note that the suggested rank test is quite powerful when compared to the LBI test.

Table 4. Comparison of power when u is Normal(0, 1)

n	1%	5%	10%	1%	5%	10%	1%	5%	10%
	$\varepsilon\sim$ Logistic			$\varepsilon\sim$ Normal			$\varepsilon\sim$ Laplace		
	Rank test								
50	0.770	0.832	0.898	0.757	0.842	0.904	0.759	0.878	0.969
100	0.874	0.943	0.973	0.883	0.949	0.966	0.856	0.941	0.988
250	0.979	0.992	0.995	0.976	0.933	0.997	0.902	0.965	0.997
LBI test									
50	0.770	0.829	0.891	0.756	0.841	0.896	0.763	0.856	0.898
100	0.859	0.936	0.966	0.878	0.941	0.968	0.870	0.939	0.965
250	0.966	0.990	0.994	0.972	0.994	0.995	0.959	0.992	0.996
	$\varepsilon\sim$ Cauchy(0,1)			$\varepsilon\sim t(2)$			$\varepsilon\sim$ Std. $\chi^2(4)$		
Rank test									
50	0.718	0.801	0.868	0.760	0.839	0.897	0.757	0.855	0.905
100	0.863	0.923	0.949	0.879	0.949	0.971	0.887	0.942	0.966
250	0.977	0.995	0.997	0.977	0.997	1.000	0.978	0.996	0.998
LBI test									
50	0.455	0.564	0.628	0.747	0.843	0.896	0.761	0.851	0.908
100	0.666	0.75	0.805	0.873	0.942	0.964	0.872	0.933	0.965
250	0.935	0.962	0.973	0.968	0.992	1.000	0.968	0.991	0.998

From a practical point of view, it would be interesting to investigate the performance of the rank tests for situations in which the regression coefficient changes at a very slower rate. Therefore, we have considered similar kind power comparisons of these two tests when $\mathrm{Var}(u_t) < \mathrm{Var}(\varepsilon_t)$. These are given in Table 5. Here also, we observe that the performance of the rank test is quite good. It may be specifically noted that when ε follows the Cauchy distribution, LBI test perform very badly in terms of its power for smaller sample sizes.

Table 5. Comparison of power when *u* is Normal(0, 0.5)

n	1%	5%	10%	1%	5%	10%	1%	5%	10%
	$\varepsilon \sim$ Logistic			$\varepsilon \sim$ Normal			$\varepsilon \sim$ Laplace		
	Rank test								
0	0.489	0.623	0.699	0.743	0.824	0.889	0.745	0.829	0.880
100	0.866	0.945	0.972	0.875	0.944	0.973	0.515	0.640	0.718
250	0.489	0.991	0.999	0.973	0.990	0.997	0.971	0.993	0.999
LBI test									
50	0.441	0.593	0.687	0.740	0.823	0.882	0.735	0.829	0.876
100	0.860	0.936	0.962	0.870	0.944	0.970	0.440	0.570	0.642
250	0.959	0.987	0.994	0.957	0.989	0.994	0.967	0.989	0.998
	$\varepsilon \sim$ Cauchy(0, 1)			$\varepsilon \sim t(2)$			$\varepsilon \sim$ Std. $\chi^2(4)$		
	Rank test								
50	0.656	0.759	0.807	0.721	0.806	0.877	0.755	0.842	0.891
100	0.847	0.916	0.952	0.793	0.880	0.914	0.864	0.937	0.959
250	0.971	0.992	0.998	0.984	0.997	0.999	0.97	0.992	0.999
LBI test									
50	0.311	0.420	0.487	0.690	0.771	0.848	0.753	0.842	0.893
100	0.597	0.687	0.739	0.712	0.814	0.867	0.854	0.929	0.959
250	0.881	0.931	0.951	0.975	0.991	0.995	0.962	0.983	0.991

Now, we investigate the power of these two tests when Var(u_t) > Var(ε_t). From Table 6, we can see that under this set up also, rank test perform quite well.

Table 6. Comparison of power when *u* is Normal(0, 1.5)

n	1%	5%	10%	1%	5%	10%	1%	5%	10%
	$\varepsilon \sim$ Logistic			$\varepsilon \sim$ Normal			$\varepsilon \sim$ Laplace		
	Rank test								
50	0.754	0.835	0.901	0.762	0.852	0.917	0.668	0.775	0.830
100	0.892	0.952	0.970	0.884	0.946	0.975	0.866	0.943	0.966
250	0.961	0.993	0.997	0.973	0.989	0.997	0.975	0.995	0.998
LBI test									
50	0.754	0.838	0.899	0.760	0.843	0.907	0.649	0.766	0.809
100	0.874	0.941	0.970	0.875	0.945	0.970	0.857	0.929	0.966
250	0.956	0.987	0.996	0.956	0.989	0.994	0.964	0.991	0.998
	$\varepsilon \sim$ Cauchy(0, 1)			$\varepsilon \sim t(2)$			$\varepsilon \sim$ Std. $\chi^2(4)$		
	Rank test								
50	0.457	0.587	0.673	0.744	0.828	0.874	0.765	0.853	0.902
100	0.845	0.92	0.951	0.888	0.949	0.965	0.892	0.957	0.975
250	0.975	0.994	0.999	0.974	0.991	0.997	0.976	0.994	0.997
LBI test									
50	0.121	0.223	0.305	0.685	0.797	0.849	0.767	0.851	0.904
100	0.697	0.79	0.834	0.877	0.940	0.960	0.864	0.946	0.971
250	0.896	0.94	0.959	0.969	0.991	0.997	0.971	0.991	0.997

In most of the cases, bootstrap offers a better approximation to the sampling distribution of a pivotal/test statistics as compared to the conventional procedures (such as normal approximation). In this case, the asymptotic distribution is an infinite linear combination of independent chi-squares, both the weights and variances of normal random variables being unknown. It is thus very reassuring to see that the bootstrap works and offers a very handy tool.

6. Concluding remarks

We have approached this testing problem nonparametrically to enhance its scope of application. The major advantage of our procedure is that we do not require second order moment assumptions on the innovation density as opposed to a stringent moment requirement assumption made by Nabeya and Tanaka [16]. On the other hand, the major drawback of this procedure is that it is not distribution free. However, the suggested bootstrap turned out to be performing reasonably well for the problem at hand.

The locally most powerful rank test may be derived based on the ranks of a transformed variable ω, $\omega = H'(y - \beta x_i)$, where H is an appropriate transformation, as discussed in [16]. Perhaps, this may lead to a distribution free asymptotics and hence will emphasize the reasons for resorting to rank tests. This is being investigated currently.

7. Proofs

This section gives proofs for all lemmas and theorems stated in Sections 3 and 4.

Proof of Lemma 3.1. (Theorem 2 of p.198 [26]).

Proof of Lemma 3.2. Let $\bar{\psi}_T = \frac{1}{T}\sum_{i=1}^{T}\psi\left(\frac{i}{T}\right)$. Then

$$T^{-1/2}\sum_{i=1}^{T}\psi\left(\frac{i}{T}\right)V_i = T^{-1/2}\sum_{i=1}^{T}\left\{\psi\left(\frac{i}{T}\right)-\bar{\psi}_T\right\}\phi\left(\frac{R_i}{T+1}\right)$$

$$= T^{-1/2}\left[\gamma(\mathbf{c}(\psi))\right]^{-1}\left\{[\gamma(\mathbf{c})]\sum_{i=1}^{T}c_i(\psi)\phi\left(\frac{R_i}{T+1}\right)\right\}$$

$$= T^{-1/2}\left[\gamma(\mathbf{c}(\psi))\right]^{-1}\int_0^1\phi(t)\,dR_T(t)$$

where $R_T(t)$ is as defined in (12) with $c_i(\psi) = \left\{\psi\left(\frac{i}{T}\right) - \bar{\psi}_T\right\}$. Thus, as a consequence of Lemma 3.1 and the fact that $T^{-1/2}[\gamma(\mathbf{c}(\psi)]^{-1} \to \tau$

$$L\left\{T^{-1/2}\sum_{i=1}^{T}\psi\left(\frac{i}{T}\right)V_i\right\} \to L\left\{\tau\int_0^1 \phi(t)\,d\mathbf{X}(t)\right\}$$

Proof of Lemma 3.3. Consider $\psi(t) = a_1\psi_1(t) + a_2\psi_2(t) + \ldots + a_n\psi_n(t)$ for any real $a_1, a_2, \ldots, a_n$. Applying Lemma 3.2 to ψ, we have

$$L\left\{T^{-1/2}\sum_{j=1}^{n}a_j T^{-1/2}\sum_{i=1}^{T}\psi_j\left(\frac{i}{T}\right)V_i\right\} \to N\left[0, \sigma^2\left(\sum_{j=1}^{n}a_j^2\tau_j^2 - 2\sum_{j<}^{n}\sum_{k}^{n}a_j a_k \tau_{jk}\right)\right]$$

where $\tau_j^2 = 1 - \left(\int_0^1 \psi_j(t)\,dt\right)^2$, in view of the orthogonality of ψ_i's $i = 1, 2, \ldots, n$.

Proof of Lemma 3.4. The proof is straightforward and hence omitted.

Proof of Theorem 3.1. This proof follows techniques in Nabeya and Tanaka [16]. In view of Lemma 3.4 and the assumption (16), it is enough to consider the case $B_T(i, j) = K\left(\frac{i}{T}, \frac{j}{T}\right)$. Let $0 < \lambda_1 \leq \lambda_2 \leq \ldots$ be eigen values and $\psi_1(t)$, $\psi_2(t)$, ... be the corresponding orthonormal eigen functions associated with the kernel $K(s, t)$. Let

$$K_n(s, t) = \sum_{l=1}^{n}\lambda_l\psi_l(s)\psi_l(t), \qquad n = 1,\ 2,\ \ldots$$

and put $K_T^n = \left(\left(K_n\left(\frac{i}{T}, \frac{j}{T}\right)\right)\right)$, a $T \times T$ matrix. Then

$$\frac{1}{T}\left(\mathbf{V}'K_T^n\mathbf{V}\right) = \frac{1}{T}\sum_{i=1}^{T}\sum_{j=1}^{T}K_T^n(i, j)V_iV_j$$

$$= \frac{1}{T}\sum_{i=1}^{T}\sum_{j=1}^{T}\sum_{l=1}^{n}\lambda_l\psi_l\left(\frac{i}{T}\right)\psi_l\left(\frac{j}{T}\right)V_iV_j$$

$$= \sum_{l=1}^{n}\lambda_l\left\{T^{-1/2}\sum_{i=1}^{T}\psi_l\left(\frac{i}{T}\right)V_i\right\}^2$$

By Lemma 3.3, for each fixed n

$$L\left(\frac{1}{T}\mathbf{V}'K_T^n\mathbf{V}\right) \to L\left(\sum_{i=1}^{n}\lambda_i Z_i^2\right) \qquad \text{as } T \to \infty$$

An application of Lemma 3.4 completes the proof.

Proof of Lemma 4.1. 1. Let

$$g_T(y) = \left\{\frac{1}{b(T)}\right\}\int \frac{k(y-z)}{b(T)} f(z)\,dz$$

and

$$\hat{F}_T(t) = \left\{\frac{1}{T}\right\}\sum_{i=1}^{T} I_{[r_i \le t]}$$

From [26], Theorem 2 of p. 198, we note that $\left\|\hat{F}_T - F\right\| \to 0$ in probability. Therefore,

$$\left\|\hat{f}_T - g_T\right\| = \left\|\left\{\frac{1}{b(T)}\right\}\int k\frac{(y-z)}{b(T)}d\hat{F}_T(z) - \left\{\frac{1}{b(T)}\right\}\int k\frac{y-z)}{b(T)}dF(z)\right\| \le \left\|\hat{F}_T - F_T\right\|\int(d|k|)$$

where $d|k|$ denotes the total variation measure of the function k. By Theorem 2.1.1 of [20, p. 35], it follows that $\|g_T - f\| \to 0$ uniformly. Thus, 1 follows.

2. We prove only one of the results: We show that $\hat{f}$ is absolutely continuous with probability tending to one as $T \to \infty$. Since $\left\|\hat{f}_T - f\right\| \to 0$ in probability, for every subsequence $\{T_i\}$, there exists a further subsequence $\{\mathrm{T}_i'\}$ such that $\left\|\hat{f}_{T_i'} - f\right\| \to 0$ almost surely.

Let $\{(x_i, x_i')\}_{i=1}^{N}$ be a finite collection of non-overlapping intervals of the real line R with $\sum_{i=1}^{N}|x_i' - x_i| < \delta$. Since f is absolutely continuous, we have

$$\begin{aligned}\sum_{i=1}^{N}\left|\hat{f}_{T_i'}(x_i') - \hat{f}_{T_i'}(x_i)\right| &\le \sum_{i=1}^{N}\left|\hat{f}_{T_i'}(x_i') - f(x_i')\right| \\ &\quad + \left|\sum_{i=1}^{N}\hat{f}_{T_i'}(x_i) - f(x_i)\right| + \sum_{i=1}^{N}\left|f(x_i) - f(x_i')\right| \\ &\le 2N\left\|\hat{f}_{T_i'} - f\right\| + \frac{\varepsilon}{2}\end{aligned}$$

which can be made smaller than ε in view of part 1. Since this is true for a subsequence of every subsequence, the proof is complete.

Proof of Theorem 4.1. We follow a short proof of validity of bootstrap given in [25] (also found in [12]). Let $\{\xi_i^*\}$ be a sequence of iid uniform (0, 1) random variables. Define $\xi_i^* = \hat{F}_T^{-1}\left(\xi_i^*\right)$ and for $0 \le x \le 1$,

$$U^*(x) = \gamma(\mathbf{c})\left\{\sum_{i=1}^{T} c_i \left\{I_{\left[\xi_i^* \le x\right]} - x\right\}\right\} \tag{22}$$

It is well known that there exists a probability space and a Brownian bridge W such that $\|U^* - W\| \to 0,$ in probability. Now consider

$$\left\| Z_T^*(x) - \left\{ W_1(F(x)) + \gamma(\mathbf{c})\hat{f}_T(x)\left(\hat{\beta}^* - \hat{\beta}\right)\sum_{i=1}^{T} c_i x_i \right\} \right\|$$

$$\le \left\| \gamma(\mathbf{c}) \sum_{i=1}^{T} c_i \left\{ I_{\left[\varepsilon_i^* \le x + (\hat{\beta}^* - \hat{\beta})x_i\right]} - \hat{F}_T\left(x + \left(\hat{\beta}^* - \hat{\beta}\right)x_i\right)\right\} - W_1(F) \right\|$$

$$+ \left\| \gamma(\mathbf{c}) \sum_{i=1}^{T} c_i \left\{ \hat{F}_T\left(x + \left(\hat{\beta}^* - \hat{\beta}\right)x_i\right) - \hat{F}_T(x)\right\} - \gamma(\mathbf{c})\hat{f}_T(x) + \left(\hat{\beta}^* - \hat{\beta}\right)\left(\sum_{i=1}^{T} c_i x_i\right) \right\|$$

$$= T_1 + T_2 \text{ (say)}$$

By expanding $\hat{F}_T\left(x + \left(\hat{\beta}^* - \hat{\beta}\right)x_i\right)$ around $\hat{F}_T(x)$ and using the uniform continuity of $\hat{f}_T$, it follows that $T_2 \to 0$, in P^* – probability. Now, note that

$$\begin{aligned} T_1 \le & \left\| \gamma(\mathbf{c}) \sum_{i=1}^{T} c_i \left\{ I_{\left[\xi_i^* \le \hat{F}_T\left(x + (\hat{\beta}^* - \hat{\beta})x_i\right)\right]} - \hat{F}_T\left(x + \left(\hat{\beta}^* - \hat{\beta}\right)x_i\right)\right\} - W_i\left(\hat{F}_T\right) \right\| \\ & + \left\| W_1\left(\hat{F}_T\right) - W_1(F) \right\| \end{aligned} \tag{23}$$

The second term of (23) converges to zero in probability, in view of the fact that $\left\|\hat{F}_T - F\right\| \to 0$ in probability. For the first term, one may repeat the arguments in Section 4.5 of [26]; in particular, the process Z_T^{*b} defined by

$$Z_T^{*b} = \gamma(\mathbf{c})\sum_{i=1}^{T} c_i \left\{ I_{[\varepsilon_i^* \le x + b\gamma(\mathbf{x})x_i]} - \hat{F}_T\left(x + b\gamma(\mathbf{x})x_i\right)\right\}$$

can be shown to satisfy that $\left\|\dot{Z}_T^{*b} - W_1\left(\hat{F}_T\right)\right\| \to 0$ in P^* – probability, uniformly in $|b| \le B$, $0 \le B < \infty$ (cf. Theorem 2 of [26], p. 186).

Further, under the conditions of the theorem, $\left(\hat{\beta}^* - \hat{\beta}\right)$ satisfies $\{\gamma(\mathbf{x})\}^{-1}\left(\hat{\beta}^* - \hat{\beta}\right)$ $= O_{p^*}(1)$ (cf. Theorem 2.2 of [25]). Therefore, for every $\varepsilon > 0$ there exists a B_ε such that $P^*[|b^*| \le B_\varepsilon] > (1-\varepsilon)$, $T > T_\varepsilon$, where $b^* = \left\{\{\gamma(\mathbf{x})\}^{-1}\left(\hat{\beta}^* - \hat{\beta}\right)\right\}$.

Since $\max\limits_{1\le t\le T}\left[\{\gamma(\mathbf{x})\}^2 x_t^2\right] \to 0$ as $T \to \infty$, the result follows, in view of the assumption (2) of Lemma 4.1 and the consistency of $\hat{\sigma}$.

Acknowledgements

Authors thank the Associate Editor for some of the valuable suggestions. Chanchala Ghadge would like to acknowledge the financial support from the University Grants Commission, India, in the form of a Senior Research Fellowship.

References

[1] CHERNICK M., *Bootstrap Methods: A Practitioner's Guide*, Wiley, New York, 2007.
[2] COX D.R., HINKLEY D.V., *Theoretical Statistics,* Chapman and Hall, London, 1974.
[3] DAVISON A.C., HINKLEY D.V., *Bootstrap Methods and Their Application*, Cambridge Series in Statistical and Probabilistic Mathematics, No. 1, 1999.
[4] DELICADO F., ROMO J., G*oodness-of-fit tests in random coeffient regression models*, Annals of the Institute of Statistical Mathematics, 1999, 51, 125–148.
[5] DELICADO F., ROMO J., *Random coefficient regressions: Parametric goodness-of-fit tests*, Journal of Statistical Planning and Inference, 2004, 119 (2), 377–400.
[6] GARBADE K., *Two methods for examining the stability of regression coefficients*, Journal of American Statistical Association, 1977, 72, 54–63.
[7] HALL P., WILSON S.R., *Two guidelines for bootstrap hypothesis testing*, Biometrics, 1991, 47, 757–762.
[8] HAUSMAN J.A., *Specification tests in econometrics*, Econometrica, 1978, 46 (6), 1251–1271.
[9] HINKLEY D.V., *Bootstrap significance tests*, Bulletin of the International Statistical Institute, Proceedings of the 47th Session, 1989, 53, 65–74.
[10] JANDHYALA V.K., MACNEILL I.B., *On testing for the constancy of regression coefficients under random walk and change-point alternatives*, Econometric Theory, 1992, 8 (4), 501–517.
[11] KOROLJUK V.S., BOROVISKICH Y.V., *Theory of Statistics*, [in:] *Mathematics and its application*, Vol. 273, Kluwer Academic Publishers Group, Dordrecht, 1994.
[12] LAHIRI S.N., *Resampling methods for dependent data*, Springer, New York, 2003.

[13] LaMotte, L.R., McWhorter A., *An exact test for the presence of random walk coefficients in a linear model*, Journal of American Statistical Association, 1978, 73, 816–820.
[14] Lee A.J., *U-Statistics Theory and Practice*, Dekker, New York, 1990.
[15] Nabeya S., *Asymptotic distributions of the test statistics for the constancy of regression coefficients under a sequence of random walk alternatives*, Journal of the Japan Statistical Society, 1989, 19, 13–33.
[16] Nabeya S., Tanaka K., *Asymptotic theory of a test for the constancy of regression coefficients against the random walk alternative*, Annals of Statistics, 1988, 16, 218–235.
[17] Nabeya S., Tanaka K., *Acknowledgment of priority*, The Annals of Statistics, 1994, 22 (1), 563.
[18] Newbold P., Bos T., *Stochastic Parameter Regression Models*, Series: Quantitative Applications in Social Sciences, A Sage University Paper No. 51, 1985.
[19] Nyblom S., Maakelaainen T., *Comparison of tests for the presence of random walk coefficients in a simple linear model*, Journal of American Statistical Association, 1983, 78, 856–864.
[20] Prakasa Rao B.L.S., *Nonparametric Functional Estimation*, Academic Press, New York, 1983.
[21] Rajarshi M.B., Ramanathan T.V., *Testing constancy of a Markovian parameter against random walk alternatives*, Journal of Indian Statistical Association, 2000, 38, 23–44.
[22] Ramanathan T.V., Rajarshi M.B., *Rank tests for testing randomness of a regression coefficient in a linear regression model*, Metrika, 1992, 39, 113–124.
[23] Ramsey J.B., *Tests for specification errors in classical linear least squares regression analysis*, Journal of the Royal Statistical Society B, 1969, 31 (2), 350–371.
[24] Shively T.S., *An exact test for a stochastic coefficient in a time series regression model*, Journal of Time Series Analysis, 1988, 9, 81–88.
[25] Shorack G.R., *Bootstrapping robust regression*, Communications in Statistics: Theory and Methods, 1982, 11, 961–972.
[26] Shorack G., Wellner J.A., *Empirical Processes with Applications to Statistics*, Wiley, New York, 1986.
[27] Silverman B.W., *Density Estimation for Statistics and Data Analysis*, Chapman and Hall, London, 1996.
[28] Swamy P.A.V.B., *Statistical Inference in Random Coeffcient Regression Model*, Lecture Notes, Springer, 1971.
[29] Zelterman D., Chen C., *Homogeneity tests against central mixture alternatives*, Journal of American Statistical Association, 1988, 83, 179–182.