KYBERNETIKA — VOLUME 37 (2001), NUMBER 6, PAGES 669-684

SOME INVARIANT TEST PROCEDURES FOR DETECTION OF STRUCTURAL CHANGES; BEHAVIOR UNDER ALTERNATIVES[1]

MARIE HUŠKOVÁ

Regression- and scale-invariant M-test procedures for detection of structural changes in linear regression model was developed and their limit behavior under the null hypothesis was studied in Hušková [9]. In the present paper the limit behavior under local alternatives is studied. More precisely, it is shown that under local alternatives the considered test statistics have asymptotically normal distribution.

1. INTRODUCTION

The present paper is a continuation of the paper Hušková [9], where a class of M-type regression- and scale- invariant test statistics for detection of a change in regression models are developed and their limit behavior under the null hypothesis (no change) is studied. Here we focus on the limit behavior under alternatives.

We consider the regression model with a change after an unknown time point m_n:

$$Y_{in} = \boldsymbol{x}_i^T\boldsymbol{\beta} + \boldsymbol{x}_i^T\boldsymbol{\delta}_n I\{i > m_n\} + e_i, \quad i = 1\dots,n, \tag{1.1}$$

where $m_n(\leq n)$, $\boldsymbol{\beta} = (\beta_1,\dots,\beta_p)^T$, $\boldsymbol{\delta}_n = (\delta_{n1},\dots,\delta_{np})^T \neq \mathbf{0}$ are unknown parameters, $\boldsymbol{x}_i = (x_{i1},\dots,x_{ip})^T$, $x_{i1} = 1$, $i = 1,\dots,n$, are known design points, and $e_1,\dots,e_n$ are iid random variables with common distribution F that fulfills regularity conditions specified below. Here $I\{A\}$ denotes the indicator of the set A.

We write the index n with the observations Y_{in} and the parameters m_n and $\boldsymbol{\delta}_n$ because we study the limit properties as $n \to \infty$ and we assume that both m_n and $\boldsymbol{\delta}_n$ are changing together with n. More precisely, we assume that, as $n \to \infty$

$$m_n = [\lambda n] \quad \text{for some } \lambda \in (0,1] \tag{1.2}$$

and

$$||\boldsymbol{\delta}_n|| \to 0. \tag{1.3}$$

[1]Partially supported by Grant 161/1999/BMAT/MFF of the Charles University and by the research project "Mathematical Methods in Stochastics" - MSM 113200008.

at a certain rate. Here $\|.\|$ denotes the Euclidean norm and $[a]$ denotes the integer part of a.

Model (1.1) describes the situation where the first m_n observations follow the linear model with the parameter $\boldsymbol{\beta}$ and the remaining $n - m_n$ ones follow the linear regression model with the parameter $\boldsymbol{\beta} + \boldsymbol{\delta}_n$. Such models are called two phase regression models. The parameter m_n is usually called *the change point*.

Hušková [9] developed regression- and scale-invariant M-type test for

$$H_0 : \lambda = 1 \quad \text{against} \quad H_1 : \lambda \in (0,1) \tag{1.4}$$

and derived their limit distribution under H_0. The null hypothesis is saying that "no change has occurred" and the alternative states "a change has occurred".

In applications one meets often the testing problem (1.4) sometimes called detection of structural changes. Typically, one observes a sequence of variables and might be interested to know whether the possible statistical model remains the same during the whole observational period or whether the model changes at some unknown time point. Such problems occur in various situations, e. g. changes in hydrological or meteorological or econometric time series. For recent references, see, e. g. Csörgő and Horváth [3].

There are not too many results on the behavior of the test statistics for detection of changes in regression models under alternatives. More information about recent development can be found, e. g. in Horváth [4] and Csörgő and Horváth [3].

Here, we show that under a class of local alternatives the limit distribution of properly standardized test statistics is normal even if the score function depends on observations.

Set

$$\boldsymbol{C}_k = \sum_{i=1}^{k} \boldsymbol{x}_i \boldsymbol{x}_i^T, \quad \boldsymbol{C}_k^0 = \boldsymbol{C}_n - \boldsymbol{C}_k, \quad k = 1, \dots, n. \tag{1.5}$$

The M-test procedures generated by a score function ψ are defined as

$$T_n(\psi) = \max_{p<k<n-p} \left\{ \boldsymbol{S}_k(\psi)^T (\boldsymbol{C}_k^{-1} \boldsymbol{C}_n (\boldsymbol{C}_k^0)^{-1}) \boldsymbol{S}_k(\psi) \right\} / \widehat{\sigma}_n^2(\psi) \tag{1.6}$$

and

$$T_n(\psi, q) = \sup_{0<t<1} \left\{ \frac{\boldsymbol{S}_{[(n+1)t]}(\psi)^T \boldsymbol{C}_n^{-1} \boldsymbol{S}_{[(n+1)t]}(\psi)}{q^2(t) \widehat{\sigma}_n^2(\psi)} \right\}, {}^* \tag{1.7}$$

where q is a weight function and

$$\boldsymbol{S}_k(\psi) = \sum_{i=1}^{k} \boldsymbol{x}_i \psi \left(Y_i - \boldsymbol{x}_i^T \widehat{\boldsymbol{\beta}}_n(\psi) \right), \quad k = 1, \dots, n \tag{1.8}$$

and $S_0(\psi) = 0$. Here $\widehat{\boldsymbol{\beta}}_n(\psi)$ is the M-estimator with the score function ψ based on $Y_1, \dots, Y_n$ and $\widehat{\sigma}_n^2(\psi)$ is an estimator of $\sigma^2(\psi) = \int \psi(x)^2 \, dF(x)$ with the property

$$\widehat{\sigma}_n^2(\psi) - \sigma^2(\psi) = o_p(1), \quad n \to \infty. \tag{1.9}$$

*the subscript $[(n+1)t]$ should replace $[nt]$ also in analogous definitions in Hušková [9].

It is known that

$$\widehat{\sigma}_n^2(\psi) = \frac{1}{n}\sum_{i=1}^{n}\psi^2\left(Y_i - \boldsymbol{x}_i^T\widehat{\boldsymbol{\beta}}_n(\psi)\right) \tag{1.10}$$

has the desired property (1.9) for a quite broad spectrum of ψ and local alternatives. However, another possible estimator is

$$\tilde{\sigma}_n^2(\psi) = \widehat{\sigma}_n^2(\psi) - \max_{1\le k<n}\frac{1}{k(n-k)}\left(\sum_{i=1}^{k}\psi\left(Y_i - \boldsymbol{x}_i^T\widehat{\boldsymbol{\beta}}_n(\psi)\right)\right)^2, \tag{1.11}$$

that has the desired property (1.9) even under more general alternatives and works well even for moderate sample sizes.

However, the test statistics $T_n(\psi)$ and $T_n(\psi, q)$ are regression-invariant but generally not scale-invariant. Our main aim is to study the regression- and scale-invariant M-tests defined in (1.6) – (1.9) with the score function ψ replaced by

$$\widehat{\psi}_n(\cdot) = \psi(\cdot;\ K_n(1-\alpha)), \tag{1.12}$$

with

$$\psi(x;K) = \begin{cases} x & |x| \le K \\ K \operatorname{sign} x & |x| \ge K \end{cases} \tag{1.13}$$

and $K_n(1-\alpha)$ is an estimator of the $(1-\alpha)$-quantile $F^{-1}(1-\alpha)$ of the distribution function F. Namely, we consider the estimator

$$K_n(1-\alpha) = K_n(1-\alpha, \boldsymbol{Y}_n) = \frac{1}{2}\left(\tilde{\beta}_{n1}(1-\alpha) - \tilde{\beta}_{n1}(\alpha)\right), \tag{1.14}$$

where $\tilde{\beta}_{n1}(1-\alpha)$ and $\tilde{\beta}_{n1}(\alpha)$ are the first components of the $(1-\alpha)$th and αth regression quantiles based on $\boldsymbol{Y}_n = (Y_1, \dots, Y_n)^T$. The score function $\widehat{\psi}_n$ was developed by Jurečková and Sen [11] and is called the *adaptive Huber score function.* We should remark that $\widehat{\psi}_n$ is an estimator of the score function $\psi(\cdot; F^{-1}(1-\alpha))$, for detail see Jurečková and Sen [11].

We remind the definition of the α-regression quantiles. Towards this we denote

$$\begin{aligned} \phi_\alpha(x) &= \alpha - I\{x \le 0\}, \quad x \in R^1, & (1.15)\\ \rho_\alpha(x) &= x\phi_\alpha(x), \qquad\quad x \in R^1. & (1.16) \end{aligned}$$

The α-regression quantile $\tilde{\boldsymbol{\beta}}_n(\alpha)$ is defined as a solution $\boldsymbol{v}$ of the following minimization problem:

$$\min_{\boldsymbol{t}\in R^p}\sum_{i=1}^{n}\rho_\alpha(Y_i - \boldsymbol{t}^T\boldsymbol{x}_i). \tag{1.17}$$

If the solution is not unique we may set a rule how to choose it.

Large values of $T_n(\widehat{\psi}_n)$ and $T_n(\widehat{\psi}_n, q)$ indicate that the null hypothesis is violated. Approximations to the critical values can be found through their limit distribution under the null hypothesis. For more details confer to Hušková [9].

In the next section the main results are formulated. The proofs are postponed to Section 3.

2. MAIN RESULTS

In this section we formulate and discuss the main assertions on the limit distribution of $T_n(\psi)$ and $T_n(\psi, q)$ under a class of local alternative hypotheses. The assumptions on the distribution function F of the error terms are identical with those considered by Jurečková and Sen [11] while the assumptions on the design points coincide with those on design points for L_2 procedures for detection of a change in linear models.

We assume that the design points $\boldsymbol{x}_i = (x_{i1}, \dots, x_{ip})^T, i = 1, \dots, n$, satisfy:

(A.1) $x_{i1} = 1$, $i = 1, \dots, n$.

(A.2) There exists a positive definite $p \times p$ matrix $\boldsymbol{C}$ such that $\lim_{n\to\infty} \frac{1}{n}\boldsymbol{C}_{[nt]} = t\boldsymbol{C}$, $t \in (0, 1]$, where $\boldsymbol{C}$ is a positive definite matrix.

(A.3) As, $n \to \infty$,

$$\max_{1\le k\le n} \left\{ \frac{1}{k} \sum_{i=1}^{k} \|\boldsymbol{x}_i\|^4 + \frac{1}{n-k} \sum_{i=k+1}^{n} \|\boldsymbol{x}_i\|^4 \right\} = O(1).$$

The distribution function F of the error terms e_i's satisfies the following set of assumptions:

(B.1) F has absolutely continuous density f and finite nonzero Fisher's information $0 < I(f) = \int_{-\infty}^{\infty} (f'(x)/f(x))^2 \,\mathrm{d}F(x) < \infty$, $f'(x) = \mathrm{d}f(x)/\mathrm{d}x$.

(B.2) $f(-x) = f(x)$, $x \in R^1$.

(B.3) $0 < f(x) < \infty$ and $f'(x)$ is bounded in a neighborhood of $K > 0$ (which will be specified later).

We consider the following class of weight functions q_η

(C.1) $q_\eta(t) = (t(1-t))^\eta$, $t \in (0, 1)$ with $\eta \in [0, 1/2)$.

We still need assumptions on the amount of the change $\boldsymbol{\delta}_n$

(D.1) as $n \to \infty$,

$$\sqrt{\log n}\|\boldsymbol{\delta}_n\| \to 0, \quad \|\boldsymbol{\delta}_n\|\sqrt{n}(\log n)^{-1/2} \to \infty.$$

The assumption (D.1) corresponds to local alternatives but contiguous ones are not covered. In the following we denote

$$\psi_{\alpha,F}(x) = \psi(x; F^{-1}(1-\alpha)), \quad x \in R^1.$$

Now, we formulate the main results.

Theorem 2.1. Let $Y_{1n}, \dots, Y_{nn}$ follow the model (1.1), let assumption (1.2) with $\lambda \in (0,1)$ be satisfied. Moreover, let assumptions (D.1), (A.1) - (A.3), (B.1) - (B.2) and (B.3) with $K = F^{-1}(1-\alpha)$ for $\alpha \in (0, 1/2)$ be satisfied. Then, the limit behavior of $T_n(\widehat{\psi}_n)$ is the same as that of

$$\{\boldsymbol{S}_{m_n}^T(\psi_{\alpha,F})(\boldsymbol{C}_{m_n}^{-1}\boldsymbol{C}_n(\boldsymbol{C}_{m_n}^0)^{-1})\boldsymbol{S}_{m_n}(\psi_{\alpha,F})\} / \sigma^2(\psi_{\alpha,F}).$$

If additionally

$$n^{1/2}||\boldsymbol{\delta}_n||^3 \to 0 \tag{2.1}$$

is satisfied, then , as $n \to \infty$,

$$\left(4\xi_{m_n}\kappa^2(\psi_{\alpha,F}, F)\right)^{-1/2}\left(T_n(\widehat{\psi}_n) - \xi_{m_n}\kappa^2(\psi_{\alpha,F}, F)\right) \to^d N_p(\boldsymbol{0}, \boldsymbol{I}_p), \tag{2.2}$$

where

$$\xi_{m_n} = \boldsymbol{\delta}_n^T \boldsymbol{C}_{m_n}\boldsymbol{C}_n^{-1}\boldsymbol{C}_{m_n}^0\boldsymbol{\delta}_n, \tag{2.3}$$

$$\kappa^2(\psi, F) = \frac{\left(\int \psi'(x)\,\mathrm{d}F(x)\right)^2}{\sigma^2(\psi)}. \tag{2.4}$$

For the case of $T_n(\widehat{\psi}_n, q_\eta)$ we have

Theorem 2.2. Let the assumptions of Theorem 2.1 be satisfied. Moreover, let (C.1) be satisfied, then the limit distribution of $T_n(\widehat{\psi}_n, q_\eta)$ is the same as that of

$$\frac{\boldsymbol{S}_{m_n}^T(\psi_{\alpha,F})\boldsymbol{C}_n^{-1}\boldsymbol{S}_{m_n}(\psi_{\alpha,F})}{\sigma^2(\psi_{\alpha,F})q_\eta(\lambda)}.$$

If additionally (2.1) is satisfied, then, as $n \to \infty$,

$$\left(4\kappa^2(\psi_{\alpha,F}, F)\zeta_{n_m}\right)^{-1/2}\left(T_n(\widehat{\psi}_n, q_\eta)(\lambda(1-\lambda))^{2\eta} - \kappa^2(\psi_{\alpha,F}, F)\zeta_{n_m}\right) \to^d N_p(\boldsymbol{0}, \boldsymbol{I}_p), \tag{2.5}$$

where

$$\zeta_{m_n} = \boldsymbol{\delta}_n^T\boldsymbol{C}_{m_n}^0\boldsymbol{C}_n^{-1}\boldsymbol{C}_{m_n}\boldsymbol{C}_n^{-1}\boldsymbol{C}_{m_n}\boldsymbol{C}_n^{-1}\boldsymbol{C}_{m_n}^0\boldsymbol{\delta}_n.$$

The proofs are postponed to the next section.

Remark 2.1. The assertions of both theorems remain valid if $\widehat{\psi}_n$ is replaced by a score function $\psi(\cdot; K)$ with arbitrary $K > 0$. The assertions hold true even for unbounded score function ψ satisfying some smoothness assumptions, however the proofs become still more cumbersome. It is in fact quite interesting that the limit behavior is not effected by the estimated score function $\widehat{\psi}_n$.

Remark 2.2. Notice that the limit behavior under alternatives is completely different from that under the null one given in Hušková [9].

Remark 2.3. It can be seen from the proof that under the alternatives the single terms in $T_n(\psi)$ can be decomposed into a random and nonrandom part. The nonrandom part dominates the random one. Moreover, the nonrandom part as a function of k attains its maximum just for $k = m$ and hence the limit behavior is determined by the term for $k = m$. Checking the proofs one can find that the index $\widehat{m}_n$ for which a max is reached is a consistent estimator of the change point m_n, which means that together with testing one can also estimate the location of the change.

Remark 2.4. By Theorem 2.1 and Theorem 2.2 in Hušková [9] the critical regions with asymptotic level α based on the limit distribution of $T_n(\widehat{\psi}_n)$ and $T_n(\widehat{\psi}_n; q_\eta)$ under the null hypothesis are

$$\begin{aligned} &\sqrt{2\log\log n}(T_n(\widehat{\psi}_n))^{1/2} \qquad (2.6)\\ > \quad &-\log\log\left((1-\alpha)^{-1/2}\right) + 2\log\log n + \frac{p}{2}\log\log\log n - \log(2\Gamma(p/2)) \end{aligned}$$

and

$$(T_n(\widehat{\psi}_n, q_\eta))^{1/2} > b_p(1-\alpha, q_\eta), \qquad (2.7)$$

respectively, where $\Gamma(p) = \int_0^\infty t^{p-1}\exp\{-t\}\,dt$ and $b_p(1-\alpha, q_\eta)$ is the quantile of $\sup_{0<t<1}\left\{\frac{\left(\sum_{i=1}^p B_i^2(t)\right)^{1/2}}{q_\eta(t)}\right\}$ with $\{B_j(t); t \in (0,1)\}$, $j = 1, \dots, p$ being independent Brownian bridges. While the critical region (2.6) is easy to calculate, the quantile $b_p(1-\alpha, q_\eta)$ is explicitly known only for some particular cases, e. g. for $\eta = 0$. In most cases $b_p(1-\alpha, q_\eta)$ has to be simulated; for more information see Csörgő and Horváth [3]. By Theorem 2.1. we find that under the considered alternatives

$$T_n(\widehat{\psi}_n) = n\boldsymbol{\delta}_n^T \boldsymbol{C}\boldsymbol{\delta}_n \kappa^2(\psi_{\alpha,F}, F)\lambda(1-\lambda)(1 + o_P(1))$$

and by the assumption (D.1) the right hand side tends to ∞ faster then $\log\log n$ and therefore the test with critical region (2.6) is consistent. Similarly, by Theorem 2.2

$$T_n(\widehat{\psi}_n, q_\eta) = n\boldsymbol{\delta}_n^T \boldsymbol{C}\boldsymbol{\delta}_n \kappa^2(\psi_{\alpha,F}, F)(\lambda(1-\lambda))^{2(1-\eta)}(1 + o_P(1))$$

and hence the test with critical region (2.7) is also consistent.

3. PROOFS

To prove Theorems 2.1 − 2.2 we have to use a number of results proved elsewhere and also we have to derive a number of generalization of results connected mostly with the so called asymptotic linearity. These results are interesting of its own.

For simplicity we will write ψ instead of $\psi(\cdot; K)$ whenever possible.

We start with various asymptotic linearity results. We give the proof of the first one and skip the proofs of the others for they are quite close. The proofs are quite technical and we try to give their essence and to avoid too many technicalities. We set

$$\Delta_{ni}(\boldsymbol{t}) = \boldsymbol{\delta}_n I\{i > m\} - \boldsymbol{C}_n^{-1}\boldsymbol{C}_m^0\delta_n - q_n\boldsymbol{t}, \quad \boldsymbol{t} \in R^p. \tag{3.1}$$

Lemma 3.1. Let assumptions of Theorem 2.1 be satisfied and let $q_n \to 0$. Then for any $c \in (0,1)$, as $n \to \infty$

$$\begin{aligned} &\max_{1\le k<cn}\sup_{\|\boldsymbol{t}\|\le D}\frac{1}{\sqrt{k}}\Big\|\sum_{i=1}^{k}\boldsymbol{x}_i\left(\psi(e_i+\boldsymbol{x}_i^T\Delta_{ni}(\boldsymbol{t}))-\psi(e_i)\right.\\ &\qquad \left.-E\psi(e_i+\boldsymbol{x}_i^T\Delta_{ni}(\boldsymbol{t}))\right)\Big\| \\ =\ & O_P((\|\boldsymbol{\delta}_n\|+|q_n|)\sqrt{\log n}), \end{aligned} \tag{3.2}$$

$$\begin{aligned} &\max_{nc\le k<n}\sup_{\|\boldsymbol{t}\|\le D}\frac{1}{\sqrt{n-k}}\Big\|\sum_{i=k+1}^{n}\boldsymbol{x}_i\left(\psi(e_i+\boldsymbol{x}_i^T\Delta_{ni}(\boldsymbol{t}))-\psi(e_i)\right.\\ &\qquad \left.-E\psi(e_i+\boldsymbol{x}_i^T\Delta_{ni}(\boldsymbol{t}))\right)\Big\| \\ =\ & O_P((\|\boldsymbol{\delta}_n\|+|q_n|)\sqrt{\log n}), \end{aligned} \tag{3.3}$$

$$\begin{aligned} &\sup_{\|\boldsymbol{t}\|\le D}\frac{1}{\sqrt{n}}\Big\|\sum_{i=1}^{n}\boldsymbol{x}_i\left(\psi(e_i+\boldsymbol{x}_i^T\Delta_{ni}(\boldsymbol{t}))-\psi(e_i)-E\psi(e_i+\boldsymbol{x}_i^T\Delta_{ni}(\boldsymbol{t}))\right)\Big\| \\ =\ & O_P(\|\boldsymbol{\delta}_n\|+|q_n|), \end{aligned} \tag{3.4}$$

$$\begin{aligned} &\max_{1\le k<n}\sup_{\|\boldsymbol{t}\|\le D}\frac{1}{k}\Big\|E\sum_{i=1}^{k}\boldsymbol{x}_i(\psi(e_i+\boldsymbol{x}_i^T\Delta_{ni}(\boldsymbol{t}))-\psi(e_i)) \\ &\qquad -\left((\boldsymbol{C}_k-\boldsymbol{C}_m)\boldsymbol{\delta}_n I\{k>m\}-\boldsymbol{C}_k\boldsymbol{C}_n^{-1}\boldsymbol{C}_m^0\boldsymbol{\delta}_n-q_n\boldsymbol{C}_k\boldsymbol{t}\right)\int\psi'(x)\,\mathrm{d}F(x)\Big\| \\ =\ & O_P\left(\|\boldsymbol{\delta}_n\|^3+|q_n|^3\right) \end{aligned} \tag{3.5}$$

and

$$\begin{aligned} &\max_{1\le k<n}\sup_{\|\boldsymbol{t}\|\le D}\frac{1}{n-k}\Big\|E\sum_{i=k+1}^{n}\boldsymbol{x}_i(\psi(e_i+\boldsymbol{x}_i^T\Delta_{ni}(\boldsymbol{t}))-\psi(e_i)) \\ &\qquad -\left(\boldsymbol{C}^0_{\max(k,m)}\boldsymbol{\delta}_n-\boldsymbol{C}_k^0\boldsymbol{C}_n^{-1}\boldsymbol{C}_m^0\boldsymbol{\delta}_n-q_n\boldsymbol{C}_k^0\boldsymbol{t}\right)\int\psi'(x)\,\mathrm{d}F(x)\Big\| \\ =\ & O_P\left(\|\boldsymbol{\delta}_n\|^3+|q_n|^3\right) \end{aligned} \tag{3.6}$$

for any $D > 0$.

Proof. It is a modification of the proof of Theorem 2.1 in Hušková [6], therefore we give only a sketch of the proof. For fix $\boldsymbol{t}$ denote

$$Z_i(\boldsymbol{t}) = \psi(e_i - \boldsymbol{x}_i^T \boldsymbol{\Delta}_{ni}(\boldsymbol{t})) - \psi(e_i) - E\psi(e_i - \boldsymbol{x}_i^T \boldsymbol{\Delta}_{ni}(\boldsymbol{t})), \quad i = 1, \dots, n,$$

where $\boldsymbol{\Delta}_{ni}(\boldsymbol{t})$, $\boldsymbol{t} \in R^p$ is defined by (3.6). By the Markov inequality for each $\boldsymbol{t}$, $z > 0$ and $A > 0$

$$\begin{aligned} &P\left(\left|\sum_{i=1}^{k} x_{ij} Z_i(\boldsymbol{t})\right| \geq A\right) \\ \leq\ &\exp\{-zA\} \left(E\exp\left\{-z\sum_{i=1}^{k} x_{ij} Z_i(\boldsymbol{t})\right\} + E\exp\left\{z\sum_{i=1}^{k} x_{ij} Z_i(\boldsymbol{t})\right\}\right). \end{aligned}$$

Since $Z_i(\boldsymbol{t})$, $i = 1, \dots, n$, are independent with zero mean and

$$EZ_i^2(\boldsymbol{t}) \leq \min((\boldsymbol{x}_i^T \boldsymbol{\Delta}_n(\boldsymbol{t}))^2 D_1, K)$$

with some $D_1 > 0$ we obtain after few standard steps for $0 < z$ and $z = O\left(\frac{\log n}{k(\|\boldsymbol{\delta}_n\| + |q_n|)}\right)$

$$P\left(\left|\sum_{i=1}^{k} \boldsymbol{x}_{ij} Z_i(\boldsymbol{t})\right| \geq A\right) \leq 2\exp\left\{-zA + z^2 D_2 k(\|\boldsymbol{\delta}_n\|^2 + q_n^2)\right\}$$

with some $D_2 > 0$. We want the right hand side smaller than $n^{-\eta}$ for an arbitrary but fixed $\eta > 0$. This will be obtained for

$$z = \left(\frac{\log n}{k(\|\boldsymbol{\delta}_n\|^2 + q_n^2\|\boldsymbol{t}\|^2)}\right)^{1/2}$$

and $A > \sqrt{k}(\eta + D_2^*(\|\boldsymbol{\delta}_n\|^2 + q_n^2)$, where $D_2^* > 0$ is large enough (it depends neither on n nor on k).

Hence for any $\eta > 0$ and $D > 0$ there exist $A_\eta > 0$ and n_η such that for all $n \geq n_\eta$

$$\begin{aligned} P\Bigg(\frac{1}{\sqrt{k}} &\left\|\sum_{i=1}^{k} \boldsymbol{x}_i \left(\psi(e_i - \boldsymbol{x}_i^T \boldsymbol{\Delta}_{ni}(\boldsymbol{t})) - \psi(e_i)\right) - E\psi(e_i - \boldsymbol{x}_i^T \boldsymbol{\Delta}_{ni}(\boldsymbol{t}))\right)\right\| \\ &\geq A_\eta \left(\|\boldsymbol{\delta}_n\| + |q_n|\right)\sqrt{\log n}\Bigg) < n^{-\eta} \end{aligned}$$

for $1 \leq k \leq n$ and fixed $\|\boldsymbol{t}\| \leq D$. Similarly we get

$$P\left(\left|\sum_{i=1}^{k} \boldsymbol{x}_{ij}(Z_i(\boldsymbol{t}_1) - Z_i(\boldsymbol{t}_2))\right| \geq A\right) \leq 2\exp\left\{-zA + z^2\|\boldsymbol{t}_1 - \boldsymbol{t}_2\|^2 D_3 k((\|\boldsymbol{\delta}_n\|^2 + q_n^2)\right\}$$

with some $D_3 > 0$. To finish the proof of (3.2) we apply Theorem 12.1 of Billingsley [1]. The proofs of (3.3) and (3.4) can be obtained in the same way hence it is omitted.

The relations (3.5) – (3.6) follow easily applying Taylor expansion and by direct calculations. □

Lemma 3.2. Let the assumptions of Theorem 2.1 be satisfied.Then for any $D > 0$

$$\sup_{\|\boldsymbol{t}\| \le D} \Big\| \sum_{i=1}^{n} \big(\rho_\alpha(e_i - F^{-1}(\alpha) + \boldsymbol{x}_i^T \boldsymbol{\Delta}_{in}(\boldsymbol{t})) - \rho_\alpha(e_i - F^{-1}(\alpha)) \qquad (3.7)$$
$$-\boldsymbol{x}_i^T \boldsymbol{\Delta}_{in}(\boldsymbol{t}) \psi_\alpha(e_i - F^{-1}(\alpha) + \boldsymbol{\Delta}_{in}(\boldsymbol{t}))\big)$$
$$+\frac{1}{2} f(F^{-1}(\alpha)) \left(\boldsymbol{\delta}_n^T \boldsymbol{C}_m^0 \boldsymbol{C}_n^{-1} \boldsymbol{C}_m \boldsymbol{\delta}_n + q_n^2 \boldsymbol{t}^T \boldsymbol{C}_n \boldsymbol{t}\right) \Big\| = O_P(\|\boldsymbol{\delta}_n\|^3 n + \|\boldsymbol{\delta}_n\| q_n^2 n)$$

Proof. The lemma is a generalization of Theorem 2.1 in Hušková [6], where we have particularly $\boldsymbol{\Delta}_{in}(\boldsymbol{t}) = n^{-1/2}\boldsymbol{t}$. In our situation the proof can be done in the same way as that of the mentioned theorem and therefore is omitted. □

Lemma 3.3. Let $Y_1, \dots, Y_n$ follow the model (1.1) and let assumptions (A.1)–(A.3), (B.1)–(B.2) and (B.3(K))) for a $K > 0$ be satisfied. Then, as $n \to \infty$,

$$\tilde{\boldsymbol{\beta}}_n(1-\alpha) - \boldsymbol{\beta}(\alpha) = \boldsymbol{C}_n^{-1} \boldsymbol{C}_m^0 \boldsymbol{\delta}_n + O_P(\|\boldsymbol{\delta}_n\|^2 + n^{-1/2}), \quad \alpha \in (0,1) \qquad (3.8)$$

and

$$\widehat{\boldsymbol{\beta}}_n(\psi) - \boldsymbol{\beta} = \boldsymbol{C}_n^{-1} \boldsymbol{C}_m^0 \boldsymbol{\delta}_n - \boldsymbol{C}_n^{-1} \frac{1}{\int \psi'(x)\,\mathrm{d}F(x)} \sum_{i=1}^{n} \boldsymbol{x}_i \psi(e_i) + O_P(\|\boldsymbol{\delta}_n\|^3 + n^{-1/2}\|\boldsymbol{\delta}_n\|), \qquad (3.9)$$

where $\boldsymbol{\beta}(\alpha) = (\beta_1 + F^{-1}(\alpha), \dots, \beta_p)^T$.

Proof. The relation (3.8) follows from Lemma 3.2 in the same way as Theorem 2.4 from Theorem 2.1 in Hušková [6] therefore it is omitted. Applying (3.3) and (3.4) together with a standard arguments one receives that (3.9) holds true. □

Lemma 3.4. Let the assumptions of Lemma 3.1 be satisfied. Then, as $n \to \infty$,

$$\sum_{i=1}^{k} \boldsymbol{x}_i \psi(Y_i - \boldsymbol{x}_i^T \widehat{\beta}_n(\psi)) = \sum_{i=1}^{k} \boldsymbol{x}_i \psi(e_i) - \boldsymbol{C}_k \boldsymbol{C}_n^{-1} \sum_{i=1}^{n} \boldsymbol{x}_i \psi(e_i) \qquad (3.10)$$
$$-\boldsymbol{H}_k \boldsymbol{\delta}_n \int \psi'(x)\,\mathrm{d}F(x) + O_P\left(\sqrt{k \log n}\|\boldsymbol{\delta}_n\| + k\|\boldsymbol{\delta}_n\|^3\right),$$

uniformly for $1 \le k \le m$,

$$\sum_{i=k+1}^{n} \boldsymbol{x}_i \psi(Y_i - \boldsymbol{x}_i^T \widehat{\beta}_n(\psi)) = \sum_{i=k+1}^{n} \boldsymbol{x}_i \psi(e_i) - \boldsymbol{C}_k^0 \boldsymbol{C}_n^{-1} \sum_{i=1}^{n} \boldsymbol{x}_i \psi(e_i) \qquad (3.11)$$
$$-\boldsymbol{H}_k \boldsymbol{\delta}_n \int \psi'(x)\,\mathrm{d}F(x) g + O_P\left(\sqrt{(n-k) \log n}\|\boldsymbol{\delta}_n\| + (n-k)\|\boldsymbol{\delta}_n\|^3\right),$$

uniformly for $m < k < n$,

where

$$\boldsymbol{H}_k = \begin{cases} \boldsymbol{C}_k \boldsymbol{C}_n^{-1} \boldsymbol{C}_m^0 & 1 \le k \le m \\ \boldsymbol{C}_k^0 \boldsymbol{C}_n^{-1} \boldsymbol{C}_m & m < k \le n. \end{cases}$$

Moreover,

$$\widehat{\sigma}_n^2(\psi(\cdot;K)) - \sigma^2(\psi(\cdot;K)) = o_p(1) \tag{3.12}$$

$$\tilde{\sigma}_n^2(\psi(\cdot;K)) - \sigma^2(\psi(\cdot;K)) = O_p(\|\boldsymbol{\delta}_n\|), \tag{3.13}$$

where $\widehat{\sigma}_n^2(\psi)$ and $\tilde{\sigma}_n^2(\psi)$ are defined in (1.10) and (1.11), respectively, and

$$\sigma^2(\psi) = \int \psi^2(x)\,\mathrm{d}F(x).$$

Proof. Since (3.9) one can apply (3.1) and (3.4) with $\Delta_{ni}(\boldsymbol{t})$ replaced by $(\widehat{\boldsymbol{\beta}}_n(\psi)-\boldsymbol{\beta})$ and we observe that

$$\begin{aligned}
\frac{1}{\sqrt{k}}\sum_{i=1}^{k}\boldsymbol{x}_i\psi(Y_i-\boldsymbol{x}_i^T\widehat{\beta}_n(\psi)) &= \sum_{i=1}^{k}\boldsymbol{x}_i\psi(e_i)+\sum_{i=1}^{k}\boldsymbol{x}_i\int\psi(e-\boldsymbol{x}_i^T(\widehat{\boldsymbol{\beta}}_n-\boldsymbol{\beta}))\,\mathrm{d}F(e)\\
= \ &\sum_{i=1}^{k}\boldsymbol{x}_i\psi(e_i)-\boldsymbol{C}_k\boldsymbol{C}_n^{-1}\sum_{i=1}^{n}\boldsymbol{x}_i\psi(e_i)-\boldsymbol{H}_k\boldsymbol{\delta}_n\int\psi'(x)\,\mathrm{d}F(x)\\
&+O_P\left(\sqrt{k\log n}\|\boldsymbol{\delta}_n\|+k\|\boldsymbol{\delta}_n\|^3\right)
\end{aligned}$$

uniformly for $1\le k\le m$, and (3.10) follows. Noticing that

$$\sum_{i=k+1}^{n}\boldsymbol{x}_i\psi(Y_i-\boldsymbol{x}_i^T\widehat{\beta}_n(\psi)) = -\sum_{i=1}^{k}\boldsymbol{x}_i\psi(Y_i-\boldsymbol{x}_i^T\widehat{\beta}_n(\psi))$$

and applying (3.2) and (3.5) we get (3.11) quite similarly as (3.10).

Now we turn to the proof of (3.13). By Lemma 3.3 in Hušková [9], as $n\to\infty$,

$$\begin{aligned}
\max_{1\le k<n}&\left(\sum_{i=1}^{k}\boldsymbol{x}_i\psi(e_i)-\boldsymbol{C}_k\boldsymbol{C}_n^{-1}\sum_{i=1}^{n}\boldsymbol{x}_i\psi(e_i)\right)^T\boldsymbol{C}_k^{-1}\boldsymbol{C}_n(\boldsymbol{C}_k^0)^{-1}\\
&\left(\sum_{i=1}^{k}\boldsymbol{x}_i\psi(e_i)-\boldsymbol{C}_k\boldsymbol{C}_n^{-1}\sum_{i=1}^{n}\boldsymbol{x}_i\psi(e_i)\right)=O_P(\log\log n).
\end{aligned}\tag{3.14}$$

Then using (3.10) – (3.11) and (3.14) and noticing that the first component of $\boldsymbol{x}_i$ equals 1 we obtain

$$\begin{aligned}
&\max_{1\le k<n}\left(\sqrt{\frac{n}{k(n-k)}}\sum_{i=1}^{k}\boldsymbol{x}_i\psi(Y_i-\boldsymbol{x}_i^T\widehat{\beta}_n(\psi))\right)^2\\
= \ &\max_{1\le k<n}\left(-\sqrt{\frac{n}{k(n-k)}}(\boldsymbol{H}_k\boldsymbol{\delta}_n)_1\int\psi'(x)\,\mathrm{d}F(x)\right.\\
&\left.+O_P\left(\sqrt{\log\log n}+\sqrt{\frac{k(n-k)}{n}}\left(\sqrt{\log n}\|\boldsymbol{\delta}_n\|+\|\boldsymbol{\delta}_n\|^3\right)\right)\right)^2\\
= \ &O_P(n\|\boldsymbol{\delta}_n\|^2\log n),
\end{aligned}$$

where $(\boldsymbol{y})_1$ denotes the first component of the vector $\boldsymbol{y}$. Moreover,

$$\begin{aligned}
&\sum_{i=1}^{n} E\psi^2(e_i + \boldsymbol{x}_i\Delta_{ni}(\boldsymbol{t})) = n\int \psi(x)\,\mathrm{d}F(x) + O_P\left(\left\|\sum_{i=1}^{n}\boldsymbol{x}_i\Delta_{ni}(\boldsymbol{t})\right\|\right) \\
= \;& n\int \psi^2(x)\,\mathrm{d}F(x) + O_P\left(n(\|\boldsymbol{\delta}_n\| + |q_n|)\right).
\end{aligned}$$

Also, it can be easily checked that (3.3) holds true if ψ is replaced by ψ^2. Combining these arguments together with (3.9) we find that (3.13) holds true. Property (3.12) of $\widehat{\sigma}_n^2(\psi(\cdot, K))$ can be shown in the same way and hence it is omitted. □

Lemma 3.5. Let assumptions of Theorem 2.1 be satisfied and let $q_n \to 0$ and $\kappa_n \to 0$. Then for any $K > 0$ and any $c \in (0,1)$, as $n \to \infty$,

$$\begin{aligned}
&\max_{1\le k<cn}\sup_{|s|\le D}\sup_{\|\boldsymbol{t}\|\le D}\frac{1}{\sqrt{k}}\Bigg\|\sum_{i=1}^{k}\boldsymbol{x}_i\big(\psi(e_i+\boldsymbol{x}_i^T\Delta_{ni}(\boldsymbol{t});K+s\kappa_n) \\
&\quad -\psi(e_i+\boldsymbol{x}_i^T\Delta_{ni}(\boldsymbol{t});K) - E\big(\psi(e_i+\boldsymbol{x}_i^T\Delta_{ni}(\boldsymbol{t});K+s\kappa_n) \\
&\quad -\psi(e_i+\boldsymbol{x}_i^T\Delta_{ni}(\boldsymbol{t});K)\big)\big)\Bigg\| \\
= \;& O_P(|\kappa_n|\sqrt{\log n}),
\end{aligned} \tag{3.15}$$

$$\begin{aligned}
&\max_{cn\le k<n}\sup_{|s|\le D}\sup_{\|\boldsymbol{t}\|\le D}\frac{1}{\sqrt{n-k}}\Bigg\|\sum_{i=k+1}^{n}\boldsymbol{x}_i\big(\psi(e_i+\boldsymbol{x}_i^T\Delta_{ni}(\boldsymbol{t});K+s\kappa_n) \\
&\quad -\psi(e_i+\boldsymbol{x}_i^T\Delta_{ni}(\boldsymbol{t});K) - E\big(\psi(e_i+\boldsymbol{x}_i^T\Delta_{ni}(\boldsymbol{t});K+s\kappa_n) \\
&\quad -\psi(e_i+\boldsymbol{x}_i^T\Delta_{ni}(\boldsymbol{t});K)\big)\big)\Bigg\| \\
= \;& O_P(|\kappa_n|\sqrt{\log n})
\end{aligned} \tag{3.16}$$

$$\begin{aligned}
&\sup_{|s|\le D}\sup_{\|\boldsymbol{t}\|\le D}\frac{1}{\sqrt{n}}\Bigg\|\sum_{i=1}^{n}\boldsymbol{x}_i\big(\psi(e_i+\boldsymbol{x}_i^T\Delta_{ni}(\boldsymbol{t});K+s\kappa_n) \\
&\quad -\psi(e_i+\boldsymbol{x}_i^T\Delta_{ni}(\boldsymbol{t});K) - E\big(\psi(e_i+\boldsymbol{x}_i^T\Delta_{ni}(\boldsymbol{t});K+s\kappa_n) \\
&\quad -\psi(e_i+\boldsymbol{x}_i^T\Delta_{ni}(\boldsymbol{t});K)\big)\big)\Bigg\| \\
= \;& O_P(|\kappa_n|),
\end{aligned} \tag{3.17}$$

$$\begin{aligned}
&\max_{1\le k<cn}\sup_{|s|\le D}\sup_{\|\boldsymbol{t}\|\le D}\frac{1}{k}\sum_{i=1}^{k}\boldsymbol{x}_i E\big(\psi(e_i+\boldsymbol{x}_i^T\Delta_{ni}(\boldsymbol{t});K+s\kappa_n) \\
&\quad -\psi(e_i+\boldsymbol{x}_i^T\Delta_{ni}(\boldsymbol{t});K)\big) \\
= \;& O_P(|\kappa_n|^3 + |\kappa_n|(\|\boldsymbol{\delta}_n\| + q_n))
\end{aligned} \tag{3.18}$$

and

$$\max_{cn\le k<n}\sup_{|s|\le D}\sup_{\|\boldsymbol{t}\|\le D}\frac{1}{n-k}\sum_{i=k+1}^{n}\boldsymbol{x}_i E\left(\psi(e_i+\boldsymbol{x}_i^T\boldsymbol{\Delta}_{ni}(\boldsymbol{t});K+s\kappa_n)\right.$$
$$\left.-\psi(e_i+\boldsymbol{x}_i^T\boldsymbol{\Delta}_{ni}(\boldsymbol{t});K)\right)$$
$$= O_P(|\kappa_n|^3+|\kappa_n|(\|\boldsymbol{\delta}_n\|+q_n)) \tag{3.19}$$

for any $D>0$, where $\boldsymbol{\Delta}_{ni}(\boldsymbol{t})$ is defined in (3.6).

Proof. We follow the line of the proof of Lemma 3.1 and we point out differences only. Namely, in our situation for $s>0$

$$\psi(e;K+s\kappa_n)-\psi(e;K)=\begin{cases}-s\kappa_n & e<-K-s\kappa_n\\ e+K & -K-s\kappa_n\le e<-K\\ 0 & -K\le e<K\\ e-K & K\le e<K+s\kappa_n\\ s\kappa_n & e\ge K+s\kappa_n\end{cases}$$

and similar expression holds also for $s<0$. Then standard arguments give

$$\frac{1}{k}\sum_{i=1}^{k}E\left(\psi(e_i+\boldsymbol{x}_i^T\boldsymbol{\Delta}_{ni}(\boldsymbol{t});K+s\kappa_n)-\psi(e_i+\boldsymbol{x}_i^T\boldsymbol{\Delta}_{ni}(\boldsymbol{t});K)\right)$$
$$= O\left(\kappa_n^3+\kappa_n(\|\boldsymbol{\delta}_n\|+|q_n|)\right)$$

uniformly for $1\le k\le n$ and

$$\max_i E\left(\psi(e_i+\boldsymbol{x}_i^T\boldsymbol{\Delta}_{ni}(\boldsymbol{t});K+s\kappa_n)-\psi(e_i+\boldsymbol{x}_i^T\boldsymbol{\Delta}_{ni}(\boldsymbol{t});K)\right)^2=O\left(|s|^2\kappa_n^2\right).$$

The rest of the proof is the same as that of Lemma 3.1. □

Proof of Theorem 2.1. At first we prove that Theorem 2.1 holds true if $\widehat{\psi}_n$ is replaced by $\psi=\psi(\cdot;K)$, $K>0$.

By Lemma 3.4 $\boldsymbol{S}_k(\psi)$ can be decomposed into random and nonrandom part and the nonrandom part dominates the random one, namely, (3.10) implies

$$\boldsymbol{S}_k(\psi)=\boldsymbol{\xi}_k(\psi)-\boldsymbol{H}_k\boldsymbol{\delta}_n\int\psi'(x)\,\mathrm{d}F(x)+O_P\left(\sqrt{k\log n}\|\boldsymbol{\delta}_n\|+k\|\boldsymbol{\delta}_n\|^3\right) \tag{3.20}$$

uniformly for $1\le k\le m$, where

$$\boldsymbol{\xi}_k(\psi)=\sum_{i=1}^{k}\boldsymbol{x}_i\psi(e_i)-\boldsymbol{C}_k\boldsymbol{C}_n^{-1}\sum_{i=1}^{n}\boldsymbol{x}_i\psi(e_i),\quad k=1,\dots,n. \tag{3.21}$$

Then using (3.14) and (3.20) we have

$$\begin{aligned}
&\boldsymbol{S}_k^T(\psi)\boldsymbol{C}_k^{-1}\boldsymbol{C}_n(\boldsymbol{C}_k^0)^{-1}\boldsymbol{S}_k(\psi) = \boldsymbol{\xi}_k^T(\psi)\boldsymbol{C}_k^{-1}\boldsymbol{C}_n(\boldsymbol{C}_k^0)^{-1}\boldsymbol{\xi}_k(\psi)\\
&\quad -2\boldsymbol{\delta}_n^T\boldsymbol{H}_k^T\boldsymbol{C}_k^{-1}\boldsymbol{C}_n(\boldsymbol{C}_k^0)^{-1}\boldsymbol{\xi}_k(\psi)\int\psi'(x)\,\mathrm{d}F(x)\\
&\quad +\boldsymbol{\delta}_n^T\boldsymbol{H}_k^T\boldsymbol{C}_k^{-1}\boldsymbol{C}_n(\boldsymbol{C}_k^0)^{-1}\boldsymbol{H}_k\boldsymbol{\delta}_n(\int\psi'(x)\,\mathrm{d}F(x))^2\\
&\quad +O_P\left\{\frac{n}{k(n-k)}\left((\|\boldsymbol{\xi}_k(\psi)\|+\|\boldsymbol{\delta}_n\|k)\left(\sqrt{k\log n}\|\boldsymbol{\delta}_n\|+k\|\boldsymbol{\delta}_n\|^3\right)\right.\right.\\
&\quad \left.\left. +k\log n\|\boldsymbol{\delta}_n\|^2+k^2\|\boldsymbol{\delta}_n\|^6\right)\right\}\\
&= -2\boldsymbol{\delta}_n^T\boldsymbol{C}_m^0(\boldsymbol{C}_k^0)^{-1}\boldsymbol{\xi}_k(\psi)\int\psi'(x)\,\mathrm{d}F(x)+\boldsymbol{\delta}_n^T\boldsymbol{G}_k\boldsymbol{\delta}_n\left(\int\psi'(x)\,\mathrm{d}F(x)\right)^2\\
&\quad +O_P\left(\log\log n+\sqrt{\log\log n}\sqrt{k}\|\boldsymbol{\delta}_n\|^3+\|\boldsymbol{\delta}_n\|^2(\sqrt{k\log n}+\log n)+k\|\boldsymbol{\delta}_n\|^6\right)
\end{aligned}$$

uniformly for $1 \le k \le m$, where

$$\boldsymbol{G}_k = \boldsymbol{H}_k^T\boldsymbol{C}_k^{-1}\boldsymbol{C}_n(\boldsymbol{C}_k^0)^{-1}\boldsymbol{H}_k.$$

Clearly,

$$\boldsymbol{G}_k = -\boldsymbol{C}_m^0\boldsymbol{C}_n^{-1}\boldsymbol{C}_m^0+\boldsymbol{C}_m^0(\boldsymbol{C}_k^0)^{-1}\boldsymbol{C}_m^0,\quad 1\le k\le m,$$

and therefore the sequence $\boldsymbol{G}_k$, $1 \le k \le m$ is monotone (natural ordering of symmetric positive definite matrices) in k, the maximum is attained for $k = m$ and

$$\boldsymbol{G}_m = \boldsymbol{C}_m^0\boldsymbol{C}_n^{-1}\boldsymbol{C}_m.$$

Moreover, by the assumption (D.1)

$$\boldsymbol{\delta}_n^T\boldsymbol{G}_{[nt]}\boldsymbol{\delta}_n(\log n)^{-1}\to\infty$$

for any $t \in (0,1)$, while standard arguments give

$$\|\boldsymbol{\delta}_n^T\boldsymbol{C}_m^0(\boldsymbol{C}_k^0)^{-1}\boldsymbol{\xi}_k(\psi)\| = O_P\left(\|\boldsymbol{\delta}_n\|\sqrt{k\log\log n}\right) = o_P(\|\boldsymbol{\delta}_n\|^2 n)$$

uniformly for $1 \le k \le m$. This implies that, as $n \to \infty$,

$$P\left(\max_{1\le k\le m}\boldsymbol{S}_k^T(\psi)\boldsymbol{C}_k^{-1}\boldsymbol{C}_n(\boldsymbol{C}_k^0)^{-1}\boldsymbol{S}_k(\psi)=\max_{m(1-\epsilon)\le k\le m}\boldsymbol{S}_k^T(\psi)\boldsymbol{C}_k^{-1}\boldsymbol{C}_n(\boldsymbol{C}_k^0)^{-1}\boldsymbol{S}_k(\psi)\right)\to 1$$

for any $\epsilon \in (0,1)$. Analogous results holds true for the maximum over $m < k < n$. Moreover,

$$\begin{aligned}
&\max_{|k-m|\le\epsilon m}|\boldsymbol{\xi}_k^T(\psi)\boldsymbol{C}_k^{-1}\boldsymbol{C}_n(\boldsymbol{C}_k^0)^{-1}\boldsymbol{\xi}_k(\psi)-\boldsymbol{\xi}_m^T(\psi)\boldsymbol{C}_m^{-1}\boldsymbol{C}_n(\boldsymbol{C}_m^0)^{-1}\boldsymbol{\xi}_m(\psi)|\\
&= O_p\left(\max_{|k-m|\le\epsilon m}\{\|\boldsymbol{\xi}_k(\psi)-\boldsymbol{\xi}_m(\psi)\|\cdot\|\boldsymbol{C}_k^{-1}\boldsymbol{C}_n(\boldsymbol{C}_k^0)^{-1}\|\cdot\|\boldsymbol{\xi}_k(\psi)\|\right.\\
&\quad \left. +\|\boldsymbol{\xi}_m(\psi)\|^2\|\boldsymbol{C}_k^{-1}\|\cdot\|\boldsymbol{C}_m-\boldsymbol{C}_k\|\cdot\|\boldsymbol{C}_m^{-1}\boldsymbol{C}_n(\boldsymbol{C}_k^0)^{-1}\|\}\right)\\
&= O_P(\epsilon),
\end{aligned}$$

which means that choosing ϵ positive small enough the maximum can be made arbitrary small. Combining the above relations together with (3.13) we can infer that the limit behavior of $T_n(\psi)$ is the same as that of

$$S_m^T(\psi)C_m^{-1}C_n(C_m^0)^{-1}S_m(\psi)\frac{1}{\sigma^2(\psi)}.$$

The assertion (2.2) can be concluded just inserting (3.15) with $k = m$ into (3.20) and noticing that $\xi_m(\psi)$ is a sum of independent random vectors with zero means and the variance matrix

$$\sigma^2(\psi)C_mC_n^{-1}C_m^0$$

and the assumptions of central limit theorem are satisfied.

The proof for $T_n(\psi(\cdot;K))$ is finished.

Now, we show that the assertion of our theorem is true, i.e. for $T_n(\widehat{\psi}_n)$. At first by Lemma 3.3, (3.3) and (3.16) we find that

$$\widehat{\beta}_n(\widehat{\psi}_n)-\beta = C_n^{-1}C_m^0\delta_n - C_n^{-1}\frac{1}{\int \psi'(x)\,\mathrm{d}F(x)}\sum_{i=1}^n x_i\psi(e_i) + O_P\left(\|\delta_n\|^3 + n^{-1/2}\|\delta_n\|\right).$$

Then going step by step through the proof for $T_n(\psi)$ we observe that

$$S_k(\widehat{\psi}_n) = \xi_k(\psi) - H_k\delta_n\int \psi'(x)\,\mathrm{d}F(x) + O_P\left(\sqrt{k\log n}\|\delta_n\| + k\|\delta_n\|^3\right)$$

and

$$\widehat{\sigma}_n^2(\widehat{\psi}_n) - \sigma^2(\psi(\cdot;F^{-1}(1-\alpha))) = O_p(\|\delta_n\|)$$

which means that the asymptotic representations remains the same if $\psi(\cdot;F^{-1}(1-\alpha))$ is replaced by its estimator $\widehat{\psi}_n$. The rest of the proof is identical with that for $T_n(\psi)$. □

Proof of Theorem 2.2. Similarly as in the proof of Theorem 2.1 we show the assertion for $T_n(\psi, q_\eta)$ and then for $T_n(\widehat{\psi}_n, q_\eta)$. By (3.10) and (3.11) we have

$$\begin{aligned} & S_k^T(\psi)C_n^{-1}S_k(\psi) \qquad\qquad (3.22)\\ = & \left(\xi_k^T(\psi) - H_k\delta_n\int \psi'(x)\,\mathrm{d}F(x) + O_P(\sqrt{k\log n}\|\delta_n\| + k\|\delta_n\|^3)\right)^T C_n^{-1} \\ & \cdot\left(\xi_k(\psi) - H_k\delta_n\int \psi'(x)\,\mathrm{d}F(x) + O_P(\sqrt{k\log n}\|\delta_n\| + k\|\delta_n\|^3)\right) \\ = & -2\xi_k^T(\psi)C_n^{-1}H_k\delta_n + \delta_nH_k^TC_n^{-1}H_k\delta_n\left(\int \psi'(x)\,\mathrm{d}F(x)\right)^2 \\ & + O_P\left(1 + \|\delta_n\|^3k^{3/2}/n + \|\delta_n\|^6\right)k^2/n \end{aligned}$$

uniformly for $1 \le k \le m$. The sequence

$$\delta_nH_k^TC_n^{-1}H_k\delta_n\left(\frac{k(n-k)}{n^2}\right)^{-2\eta}$$

is monotone in k for $1 \le k \le m$ and arbitrary $0 < \eta < 1/2$ and maximum is reached for $k = m$. We set $\boldsymbol{\xi}_0(\psi) = \mathbf{0}$ and $\boldsymbol{H}_0 = \mathbf{0}$. Moreover, uniformly for $0 < t \le \lambda$

$$\left|\boldsymbol{\xi}_{[nt]}^T(\psi)\,\boldsymbol{C}_n^{-1}\boldsymbol{H}_{[nt]}\boldsymbol{\delta}_n\right| = O_P\left(\|\boldsymbol{\delta}_n\|\sqrt{n}\right) = o_P\left(\|\boldsymbol{\delta}_n\|^2 n\right)$$

and

$$\liminf_{n\to\infty} \frac{\boldsymbol{\delta}_n\boldsymbol{H}_m^T\boldsymbol{C}_n^{-1}\boldsymbol{H}_m\boldsymbol{\delta}_n}{\|\boldsymbol{\delta}_n\|^2 n} > 0.$$

Analogous results can be derived for $m < k \le n$. Thus arguing as in the proof of Theorem 2.1 one can conclude that the limit behavior of $T_n(\psi, q_\eta)$ is the same as that of

$$\frac{\boldsymbol{S}_m^T(\psi)\,\boldsymbol{C}_n^{-1}\boldsymbol{S}_m(\psi)}{q_\eta^2(\lambda)\sigma^2(\psi)}.$$

The first part of our theorem for $T_n(\psi, q_\eta)$ follows while the second one is a consequence of (3.22) for $k = m$.

The proof for the $T_n(\widehat{\psi}_n, q_\eta)$ can be derived proceeding in quite the same way as at the end of the proof of Theorem 2.1 and therefore is omitted. □

ACKNOWLEDGEMENT

The author expresses her sincere thanks to the referees for very careful reading and helpful remarks on the manuscript.

(Received March 1, 2001.)

REFERENCES

[1] P. Billingsley: Convergence of Probability Measures. Wiley, New York 1968.

[2] M. Csörgő and L. Horváth: Weighted Approximations in Probability and Statistics. Wiley, New York 1993.

[3] M. Csörgő and L. Horváth: Limit Theorems in Change-point Analysis. Wiley, New York 1997.

[4] L. Horváth: Detecting changes in linear regressions. Statistics *26* (1995), 189–208.

[5] P. J. Huber: Robust Statistics. Wiley, New York 1981.

[6] M. Hušková: Some sequential procedures based on regression rank scores. Nonparametric Statistics *3* (1994), 285–298.

[7] M. Hušková: Limit theorems for M-processes via rank statistics processes. In: Advances in Combinatorial Methods with Applications to Probability and Statistics (N. Balakrishnan, ed.), Birkhäuser, Boston 1997, pp. 521–534.

[8] M. Hušková: L_1-test procedures for detection of change. In: L_1-Statistical Procedures and Related Topics (IMS Lecture Notes – Monograph Ser. 31), Institute of Mathematical Statistics, Beachwood 1997, pp. 56–70.

[9] M. Hušková: Some invariant test procedures for detection of structural changes. Kybernetika *36* (2000), 401–414.

[10] V. K. Jandhyala and I. B. MacNeill: Residual partial sum limit process for regression models with applications to detecting parameter changes at unknown times. Stoch. Process. Appl. *33* (1989), 309–323.

[11] J. Jurečková and P. K. Sen: On adaptive scale-equivariant M-estimators in linear models. Statist. Decisions, Suplement Issue *1* (1984), 31–41.

[12] J. Jurečková and P. K. Sen: Regression rank scores scale statistics and studentization in linear models. In: Asymptotic Statistics (M. Hušková and P. Mandl, eds.), Physica-Verlag, Heidelberg, 1994, pp. 111–122.

[13] W. Ploberger, W. Krämer, and K. Kontrus: A new test for structural stability in linear regression model. J. Econometrics *40* (1989), 307–318.

[14] R. E. Quandt: Tests of hypothesis that a linear regression systems obeys two separate regimes. J. Amer. Statist. Assoc. *55* (1960), 324–330.

[15] T. Víšek: Detection of Changes in Econometric Models. Ph.D. Dissertation. Charles University, Prague 1999.

[16] K. J. Worsley: Testing for a two-phase multiple regression. Technometrics *25* (1983), 35–42.

Prof. RNDr. Marie Hušková, DrSc., Charles University – Faculty of Mathematics and Physics, Department of Statistics, Sokolovská 83, 186 00 Praha 8, and Institute of Information Theory and Automation – Academy of Sciences of the Czech Republic, Pod vodárenskou věží 4, 182 08 Praha 8. Czech Republic.
e-mail: huskova@karlin.mff.cuni.cz