KYBERNETIKA — VOLUME 36 (2000), NUMBER 4, PAGES 401–414

# SOME INVARIANT TEST PROCEDURES FOR DETECTION OF STRUCTURAL CHANGES[1]

MARIE HUŠKOVÁ

Regression and scale invariant $M$-test procedures are developed for detection of structural changes in linear regression model. Their limit properties are studied under the null hypothesis.

## 1. INTRODUCTION

In applications one meets quite often the problem to detect structural changes. Typically, one observes a sequence of variables and might be interested to known 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.

Statisticians have developed a number of test procedures for various models. For recent references, see, e. g. Csörgő and Horváth [3].

Here we focus on a class of $M$-type test statistics that are regression- and scale-invariant. It is well known the $M$-test procedures are generally developed to be insensitive to a certain violation of the normality.

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

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

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 $A$.

The model under consideration corresponds to the so called two phase regression, 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

[1]Partially supported by Grant 161/1999/BMAT/MFF from the Grant Agency of Charles University and by Grant MSM 113200008.

$\boldsymbol{\beta}+\boldsymbol{\delta}_n$. This means that the difference between these two regression parameters is $\boldsymbol{\delta}_n$. We write the index $n$ with 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 $n$. The parameter $m_n$ is usually called *the change point*.

The problem of our interest is to construct a $M$-type test for

$$H_0 : m = n \quad \text{against} \quad H_1 : m < n. \tag{1.2}$$

The null hypothesis is saying that "no change has occurred" and the alternative states "a change has occurred".

This testing problem is both regression- and scale-invariant, which means that our testing problem does not change if we transform the observations $\boldsymbol{Y}_n = (Y_1, .., Y_n)^T$ into $\boldsymbol{Z}_n = (Z_1, .., Z_n)^T = (\boldsymbol{Y}_n + \boldsymbol{X}_n \boldsymbol{b})^T s$, where $\boldsymbol{X}_n = (\boldsymbol{x}_1, \dots, \boldsymbol{x}_n)^T$, $\boldsymbol{b} \neq \boldsymbol{0}$ and $s > 0$ otherwise arbitrary. Therefore it is desirable to construct tests that are both regression- and scale-invariant.

It is known that the $L_2$ and $L_1$ procedures are regression- and scale-invariant, however we focus here on a class of $M$-test procedures that have the desired properties. We remind that the $L_2$ test procedures are related to the likelihood ratio tests when the random errors $e_i$'s have normal distribution $N(0, \sigma^2)$ while the $L_1$ test procedures are related to the likelihood ratio tests when the random errors $e_i$'s have double exponential distribution.

The $L_2$ procedures for testing $H_0$ against $H_1$ are based on either of the following test statistics:

$$T_{n,L_2} = \max_{p<k<n-p} \left\{ \boldsymbol{S}_{k,L_2}^T (\boldsymbol{C}_k^{-1} \boldsymbol{C}_n (\boldsymbol{C}_k^0)^{-1}) \boldsymbol{S}_{k,L_2} \right\} / \widehat{\sigma}_n^2 \tag{1.3}$$

$$T_{n,L_2}(q) = \sup_{0<t<1} \left\{ \frac{\boldsymbol{S}_{[nt],L_2}^T \boldsymbol{C}_n^{-1} \boldsymbol{S}_{[nt],L_2}}{q^2(t)\, \widehat{\sigma}_n^2} \right\} \tag{1.4}$$

where $[a]$ denotes the integer part of $a$,

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

$$\boldsymbol{S}_{k,L_2} = \sum_{i=1}^k \boldsymbol{x}_i (Y_i - \boldsymbol{x}_i^T \boldsymbol{\beta}_{n,L_2}), \qquad k = 1, \dots, n, \tag{1.6}$$

$$\boldsymbol{\beta}_{n,L_2} = \boldsymbol{C}_n^{-1} \sum_{i=1}^n \boldsymbol{x}_i Y_i \tag{1.7}$$

and $\widehat{\sigma}_n^2$ is a scale-equivariant and regression-invariant estimator of $\sigma^2$ with the property

$$\widehat{\sigma}_n^2 - \sigma^2 = o_p((\log\log n)^{-1/2}), \quad n \to \infty,$$

and $q$ is a positive weight function.

Notice that $\boldsymbol{\beta}_{n,L_2}$ is the least squares estimator of the vector parameter $\boldsymbol{\beta}$ in the model (1.1) with $m = n$ and the differences $Y_i - \boldsymbol{x}_i^T \boldsymbol{\beta}_{n,L_2}$, $i = 1, \dots, n$, are residuals. Since under the null hypothesis $H_0$ the random vector $\boldsymbol{S}_{k,L_2}$ has distribution

$N_p(\mathbf{0}, \sigma^2 C_k C_n^{-1} C_k^0)$, $k = 1, \ldots, n$, we realize that under $H_0$ the random variable $T_{n,L_2}\widehat{\sigma}_n^2/\sigma^2$ has the distribution as maximum of $n-2p$ (dependent) random variables with $\chi^2$-distribution with $p$ degrees of freedom.

Some authors, mostly working in the area of detection structural changes in econometrics, suggest to apply the procedures based on the properly standardized maximum of the first components of $\boldsymbol{S}_{k,L_2}, k = 1, \ldots, n$, which leads to computationally simpler procedures, however the resulting test is not sensitive with respect to some particular changes. The test procedures are based either on

$$T_{n,L_2}^0 = \max_{1 \le k < n} \left\{ \frac{|S_{1k,L_2}|}{\sqrt{n}\widehat{\sigma}_n} \right\} \tag{1.8}$$

or on

$$T_{n,L_2}^0(q) = \sup_{0<t<1} \left\{ \frac{|S_{1,[nt],L_2}|}{\sqrt{n}q(t)\widehat{\sigma}_n} \right\}, \tag{1.9}$$

where

$$S_{1k,L_2} = \sum_{i=1}^{k} (Y_i - \boldsymbol{x}_i^T \boldsymbol{\beta}_{n,L_2}), \quad k = 1, \ldots, n.$$

These procedures has been studied for example by Jandhyala and MacNeill [9] and by Ploberger, Krämer and Kontrus [12].

The null hypothesis is rejected for large values of the above test statistics. The $L_1$ procedures can be obtained by replacing the $L_2$-estimators of $\boldsymbol{\beta}$, the $L_2$-residuals and the $L_2$-estimator of $\sigma^2$ by their $L_1$-counterparts. It appears that under the null hypothesis the limit distributions of $L_2$- and the corresponding $L_1$-test statistics coincide, see Hušková [8].

Various approximations to the critical values have been developed. The test statistics (1.3), (1.5) were widely studied in the literature, e. g. Quandt [13], Worlsey [15]. More information about recent development can be found, e. g. in Horváth [4] and Csörgő and Horváth [3]. The $L_1$-procedures were developed along the line of $L_2$-procedures and studied by Hušková [8] and Víšek [14].

In the present paper we construct $M$-test procedures for the problem (1.2) that are regression- and scale-invariant.

Generally, the $M$-type test procedures generated by a score function $\psi$ can be proposed along the line of $L_2$-procedures. We can formally replace the least squares estimators $\boldsymbol{\beta}_{n,L_2}$, residuals $Y_i - \boldsymbol{x}_i^T \boldsymbol{\beta}_{n,L_2}$ and variance estimators $\widehat{\sigma}_n^2$ by their $M$-type counterparts. Then the resulting $M$-test procedures generated by a score function $\psi$ are

$$T_n(\psi) = \max_{p<k<n-p} \left\{ \boldsymbol{S}_k(\psi)^T (\boldsymbol{C}_k^{-1} \boldsymbol{C}_n (\boldsymbol{C}_k^0)^{-1}) \boldsymbol{S}_k(\psi) \right\} / \widehat{\sigma}_n^2(\psi) \tag{1.10}$$

$$T_n(\psi, q) = \sup_{0<t<1} \left\{ \frac{\boldsymbol{S}_{[nt]}(\psi)^T \boldsymbol{C}_n^{-1} \boldsymbol{S}_{[nt]}(\psi)}{q^2(t)\, \widehat{\sigma}_n^2(\psi)} \right\} \tag{1.11}$$

$$T_n^0(\psi) = \max_{1 \le k < n} \left\{ \frac{|S_{1k}(\psi)|}{\sqrt{n}\widehat{\sigma}_n(\psi)} \right\} \tag{1.12}$$

or on

$$T_n^0(\psi;\, q) = \sup_{0<t<1} \left\{ \frac{|S_{1,[nt]}(\psi)|}{\sqrt{n}q(t)\widehat{\sigma}_n(\psi)} \right\}, \tag{1.13}$$

where

$$S_k(\psi) \;=\; \sum_{i=1}^{k} \boldsymbol{x}_i \psi\left(Y_i - \boldsymbol{x}_i^T \widehat{\boldsymbol{\beta}}_n(\psi)\right), \quad k = 1, \dots, n, \tag{1.14}$$

$$S_{1k}(\psi) \;=\; \sum_{i=1}^{k} \psi\left(Y_i - \boldsymbol{x}_i^T \widehat{\boldsymbol{\beta}}_n(\psi)\right), \quad k = 1, \dots, n \tag{1.15}$$

with $\widehat{\boldsymbol{\beta}}_n(\psi)$ being the $M$-estimator with the score function $\psi$ based on $X_1, \dots, Y_n$, with $\widehat{\sigma}_n^2(\psi)$ being an scale-equivariant and regression-invariant estimator of $\sigma^2(\psi) = \int \psi(x)^2 \, \mathrm{d}F(x)$ with the property

$$\widehat{\sigma}_n^2(\psi) - \sigma^2(\psi) = o_p((\log\log n)^{-1/2}), \quad n \to \infty, \tag{1.16}$$

and $q$ is a positive weight function. It is known that

$$\widehat{\sigma}_n^2(\psi) = \frac{1}{n} \sum_{i=1}^{n} \psi^2(Y_i - \boldsymbol{x}_i^T \widehat{\boldsymbol{\beta}}_n(\psi)) \tag{1.17}$$

has the desired property (1.16) for a quite broad spectrum of $\psi$. However, since this estimator can behave quite poorly under alternatives (usually, it becomes too large and negatively influences the resulting test statistics) it is recommended to use a modified estimator, namely, make it dependent on $k$, e. g. the $k$th term should be standardized by

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

that has the desired property (1.16) even under alternatives and works well even for finite sample sizes.

However, these resulting test procedures are regression-invariant but generally not scale-invariant. To develop a scale invariant $M$-test procedure one can proceed similarly as in the construction of scale invariant $M$-estimators. A number of possibilities is discussed in detail in Jurečková and Sen [11]. For our testing problem either studentization or application of the adaptive version of the Huber $\psi$ function, proposed by Jurečková and Sen [10] seems to be reasonable.

In principle, studentization means that instead of working with the original score function $\psi$ we apply its so called studentized version $\psi(./s_n)$, where $s_n$ is a regression- and scale-invariant estimator of the scale, e. g. $s_n$ can be based on a suitably chosen functional of the regression quantiles.

We will concentrate here on the procedures based on the *adaptive Huber score function* proposed by Jurečková and Sen [10].

In the following $F^{-1}(\alpha)$ denotes the $\alpha$-quantile of the distribution function $F$ and for every $K > 0$ and $\alpha \in (0,1)$ we set

$$\psi(x;K) = \begin{cases} x & |x| \le K \\ K \text{ sign } x & |x| \ge K, \end{cases} \tag{1.18}$$

$$\phi_\alpha(x) = \alpha - I\{x \le 0\}, \quad x \in R^1, \tag{1.19}$$

$$\rho_\alpha(x) = x\phi_\alpha(x), \quad x \in R^1. \tag{1.20}$$

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

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

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

Jurečková and Sen [10] proposed an adaptive estimator $\psi(.; K_n(\alpha))$, $\alpha \in (0, 1/2)$, where

$$K_n(\alpha) = K_n(\alpha, \boldsymbol{Y}_n) = \frac{1}{2}(\tilde{\beta}_{n1}(1-\alpha) - \tilde{\beta}_{n1}(\alpha)), \tag{1.21}$$

with $\tilde{\beta}_{n1}(\alpha)$ and $\tilde{\beta}_{n1}(1-\alpha)$ being the first components of the $\alpha$th and $(1-\alpha)$th regression quantiles based on $Y_1, \dots, Y_n$. This score function is called *adaptive Huber score function* and it is related to the score function $\psi(x; F^{-1}(1-\alpha))$. Jurečková and Sen [10] showed that the $M$-estimator of the parameter $\boldsymbol{\theta}$ generated by the score function $\psi(.; K_n(\alpha))$ with a proper choice of $\alpha$ leads to the estimator that is regression- and scale-invariant and also minimax in the contaminated normal model

$$\mathcal{F} = \{F;\ F = (1-\epsilon)\Phi + \epsilon H;\ H \in \mathcal{H}\}$$

where $\Phi$ is the distribution of $N(0,1)$, $\epsilon \in (0,1)$ represents level of contamination and $\mathcal{H}$ is the family of symmetric distributions on $R^1$. In this case for the considered contamination level $\epsilon$, our $\alpha$ fulfills

$$\alpha = (1-\epsilon)(1-\Phi(K)) + \epsilon/2$$

with $K$ satisfying

$$2\phi(K)/K - 2\Phi(-K) = \epsilon/(1-\epsilon),$$

where $\phi$ denotes the density of $N(0,1)$.

The resulting regression- and scale-invariant $M$-tests are based on the test statistics defined in (1.10)–(1.13) with $\psi(.) = \psi(.; K_n(\alpha))$. We should note that these procedures are regression- and scale-invariant for each $\alpha \in (0, 1/2)$. In the following we write shortly

$$\widehat{\psi}_n(.) = \psi(.; K_n(\alpha)). \tag{1.22}$$

## 2. MAIN RESULTS

First, we formulate the assumptions. The assumptions on the distribution function $F$ of the error terms are identical with those considered by Jurečková and Sen [10] while the assumptions on the design points coincide with those on design points for $L_2$ procedures for detection of a change.

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 $\boldsymbol{C}_k$ is defined in (1.5).

(A.2) There exist $\epsilon \in (0, 1/2)$ and $\gamma > 0$ such, as $n \to \infty$,

$$\left\| \frac{1}{k} \boldsymbol{C}_k - \boldsymbol{C} \right\| = O(k^{-\gamma})$$

and

$$\left\| \frac{1}{n-k} \boldsymbol{C}_k^0 - \boldsymbol{C} \right\| = O((n-k)^{-\gamma})$$

uniformly for $1 \le k \le n\epsilon$, where $\boldsymbol{C}$ is the same as in (A.2).

(A.4) As, $n \to \infty$,

$$\max_{1 \le k \le n} \left\{ \frac{1}{k} \sum_{i=1}^{k} ||\boldsymbol{x}_i||^3 + \frac{1}{n-k} \sum_{i=k+1}^{n} ||\boldsymbol{x}_i||^3 \right\} = O(1).$$

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).

Assumptions on the weight function $q$ are the following:

(C.1) $q$ is positive on $(0, 1)$, nondecreasing in a neighborhood of 0, nonincreasing in a neighborhood of 1, $\inf\{q(t); t \in (\eta, 1-\eta)\} > 0$ for all $\eta \in (0, 1/2)$ and

$$\int_0^1 \frac{1}{s(1-s)} \exp\left\{ -\frac{cq^2(s)}{s(1-s)} \right\} \mathrm{d}s < \infty$$

for some $c > 0$.

Now, we formulate the main results. They are confirming what can be anticipated that under the null hypothesis the limit behavior of the developed $M$- test statistics is the same as that of the corresponding $L_2$ statistics.

**Theorem 2.1.** Let $Y_1, \dots, Y_n$ follow the model (1.1) with $m = n$ and let assumptions (A.1) - (A.4), (B.1) - (B.2) and (B.3) with $K = F^{-1}(1-\alpha)$ for $\alpha \in (0, 1/2)$ be satisfied, then

$$\lim_{n\to\infty} P(a(\log n)(T_n(\widehat{\psi}_n))^{1/2} \le t + b_p(\log n)) = \exp\{-2\exp\{-t\}\},\ t \in R^1, \tag{2.1}$$

and

$$\lim_{n\to\infty} P(a(\log n)T_n^0(\widehat{\psi}_n) \le t + b_1(\log n)) = \exp\{-2\exp\{-t\}\},\ t \in R^1, \tag{2.2}$$

where $\widehat{\psi}_n$ is defined by (1.22),

$$a(y) = (2\log y)^{1/2}, \quad b_p(y) = 2\log y + \frac{p}{2}\log\log y - \log(2\Gamma(p/2)), y > 1, \tag{2.3}$$

and

$$\Gamma(p) = \int_0^\infty t^{p-1}\exp\{-t\}\,dt.$$

**Theorem 2.2.** Let $Y_1, \dots, Y_n$ follow the model (1.1) with $m = n$ and let assumptions (A.1), (A.2), (A.4) and (B.1) - (B.3) with $K = F^{-1}(1-\alpha)$ be satisfied, then, as $n \to \infty$,

$$(T_n(\widehat{\psi}_n, q))^{1/2} \to^{\mathcal{D}} \sup_{0<t<1} \left\{ \frac{\left(\sum_{i=1}^p B_i^2(t)\right)^{1/2}}{q(t)} \right\} \tag{2.4}$$

and

$$T_n^0(\widehat{\psi}_n, q) \to^{\mathcal{D}} \sup_{0<t<1} \left\{ \frac{|B_1(t)|}{q(t)} \right\}, \tag{2.5}$$

where $\{B_j(t); t \in (0,1)\}$, $j = 1, \dots, p$, are independent Brownian bridges and $q$ is a weight function fulfilling (C.1).

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 $\psi(.; K)$ with arbitrary $K > 0$. The assertions hold true even for unbounded score function $\psi$ satisfying some smoothness assumptions, however the proofs become still more cumbersome.

**Remark 2.2.** Notice that under the null hypothesis the limit behavior of the considered test statistics does not depend on the particular choice of the score function and, moreover, it coincides with the limit behavior of the corresponding $L_2$ and $L_1$ test statistics.

**Remark 2.3.** The limit distributions in Theorem 2.1 belong to the extreme value family. The distributions of the limiting random variables in Theorem 2.2 are known only for particular choices of the weight function $q$. For more information, consult, e. g. Csörgő and Horváth [3].

Limit behavior of the proposed test statistics will be studied elsewhere.

## 3. PROOFS

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

First we formulate auxiliary lemmas mostly proved elsewhere.

**Lemma 3.1.** Let $Y_1, \dots, Y_n$ follow the model (1.1) with $m = n$ and let assumptions (A.1) – (A.4), (B.1) – (B.2) and (B.3(K))) for a $K > 0$ be satisfied. Then for any $\eta > 0$ there exist $A_\eta > 0$ and $n_\eta$ such that for all $n \geq n_\eta$

$$P\Big(|(\tilde{\beta}_{k1}(1-\alpha) - \tilde{\beta}_{k1}(\alpha)) - (F^{-1}(1-\alpha) - F^{-1}(\alpha)) + \frac{1}{f(F^{-1}(\alpha))}(\boldsymbol{C}_k^{-1})_1$$

$$\sum_{i=1}^{k} \boldsymbol{x}_i(\phi_{1-\alpha}(e_i - F^{-1}(1-\alpha)) - \phi_\alpha(e_i - F^{-1}(\alpha)))| \geq A_\eta k^{-v}\Big) < k^{-\eta},\ k \leq n,$$

and

$$P\Big(\Big|(\tilde{\beta}^0_{k1}(1-\alpha) - \tilde{\beta}^0_{k1}(\alpha)) - (F^{-1}(1-\alpha) - F^{-1}(\alpha))$$

$$+ \frac{1}{f(F^{-1}(\alpha))}((\boldsymbol{C}_n - \boldsymbol{C}_k)^{-1})_1 \sum_{i=k+1}^{n} \boldsymbol{x}_i(\phi_{1-\alpha}(e_i - F^{-1}(1-\alpha)) - \phi_\alpha(e_i - F^{-1}(\alpha)))\Big|$$

$$\geq A_\eta (n-k)^{-v}\Big) < (n-k)^{-\eta},\ k < n,$$

with some $v > 0$ and arbitrary $D > 0$, where $\tilde{\beta}_{k1}(\alpha)$ and $\tilde{\beta}^0_{k1}(\alpha)$ are the first components of the $\alpha$-regression quantiles $\tilde{\boldsymbol{\beta}}_k(\alpha)$, based on $Y_1, \dots, Y_k$, and of the $\alpha$-regression quantiles $\tilde{\boldsymbol{\beta}}^0_k(\alpha)$, based on $Y_{k+1}, \dots, Y_n$ and $(\boldsymbol{A})_1$ denotes the first row of the matrix $\boldsymbol{A}$.

Proof. The first assertion is a consequence of Theorem 4 in Hušková [6]. The second assertion follows in the same way if we realize that the distribution of $(e_1, \dots e_n)^T$ is the same as that of $(e_n, \dots e_1)^T$. □

We should note that this assertion is slightly stronger than is needed. However, it enables to improve the estimator of the score function $\widehat{\psi}_n$.

**Lemma 3.2.** Let the assumptions of Lemma 3.1 be satisfied. Then, as $n \to \infty$,

$$||\widehat{\boldsymbol{\beta}}_n(\psi(.;K)) - \boldsymbol{\beta}|| = O_p(n^{-1/2})$$

and

$$C_n^{1/2}(\widehat{\boldsymbol{\beta}}_n(\psi(.;K)) - \boldsymbol{\beta}) = -\frac{1}{\int \psi'(x)\,\mathrm{d}F(x)} C_n^{-1/2} \sum_{i=1}^{k} \boldsymbol{x}_i \psi(e_i;K) + O_p(n^{-v}))$$

and

$$\widehat{\sigma}_n^2(\psi(.;K)) - \sigma^2(\psi(.;K)) = O_p(n^{-v})$$

for some $v > 0$, where $\widehat{\sigma}_n^2(\psi)$ is defined in (1.17) and

$$\sigma^2(\psi) = \int \psi^2(x)\,\mathrm{d}F(x). \tag{3.1}$$

Proof. These results belong to standard results on the $M$-estimators. The proof is omitted. □

**Lemma 3.3.** Let the assumptions of Lemma 3.1 be satisfied. Then as $n \to \infty$

$$P\left(a(\log n)\left(\max_{p<k<n-p} (L_k^T C_k^{-1} C_n (C_n - C_k)^{-1} L_k)^{1/2} \frac{1}{\sigma(\psi(.;K))}\right) \tag{3.2}$$
$$\leq t + b_p(\log n)\Big) \to \exp\{-2\exp\{-t\}\},\ t \in R^1$$

and

$$\sup_{0<t<1} \frac{\left(L_{[nt]}^T C_n^{-1} L_{[nt]}\right)^{1/2}}{q(t)\sigma(\psi(.;K))} \to^{\mathcal{D}} \sup_{0<t<1} \left\{ \frac{\left(\sum_{i=1}^{p} B_i^2(t)\right)^{1/2}}{q(t)} \right\} \tag{3.3}$$

where

$$L_k = \sum_{i=1}^{k} \boldsymbol{x}_i \psi(e_i;K) - C_k C_n^{-1} \sum_{i=1}^{n} \boldsymbol{x}_i \psi(e_i;K)$$

and where $K > 0$ arbitrary, $\psi(.;K)$ and $\sigma(\psi(.;K))$ are defined by (1.18) and by (3.1), respectively, and $\{B_1(t); t \in (0,1)\}, \dots, \{B_p(t); t \in (0,1)\}$ are independent Brownian bridges.

Proof. Since $\psi(e_1;K), \dots, \psi(e_n;K)$ are iid bounded random variables the assertion (3.2) is a consequence of Theorem 3.1.5 in Csörgő and Horváth [3]. The assertion (3.3) is a consequence of Lemma 3.1.6 in Csörgő and Horváth [3] and results in Chapter 4 in Csörgő and Horváth [2]. □

**Lemma 3.4.** Let the assumptions of Theorem 2.1 be satisfied. Then as $n \to \infty$

$$\left\|\widehat{\boldsymbol{\beta}}_n(\psi(.;K_n(\alpha))) - \widehat{\boldsymbol{\beta}}_n(\psi(.;F^{-1}(\alpha)))\right\| = O_p(n^{-3/4}), \quad \alpha \in (0,1/2).$$

Proof. The assertion is a consequence of Theorem 4.1 in Jurečková and Sen [10]. □

**Lemma 3.5.** Let the assumptions of Lemma 3.1 be satisfied. Then for any $\eta > 0$ and $D > 0$ there exist $A_\eta > 0$ and $n_\eta$ such that for all $n \geq n_\eta$

$$P\left(\sup_{\|\boldsymbol{t}\|\leq D}\left\|\sum_{i=1}^{k}\boldsymbol{x}_i(\psi(e_i - n^{-1/2}\boldsymbol{x}_i^T\boldsymbol{t};K) - \psi(e_i;K)) + n^{-1/2}\boldsymbol{C}_k\boldsymbol{t}\int\psi'(x;K)\,\mathrm{d}F(x)\right\| \geq A_\eta(k/n)^{1/2}\sqrt{\log n}\right) < n^{-\eta}, \quad \alpha \in (0,1/2) \tag{3.4}$$

$$P\left(\sup_{\|\boldsymbol{t}\|\leq D}\left\|\sum_{i=k+1}^{n}\boldsymbol{x}_i(\psi(e_i - n^{-1/2}\boldsymbol{x}_i^T\boldsymbol{t};K) - \psi(e_i;K)) + n^{-1/2}(\boldsymbol{C}_n - \boldsymbol{C}_k)\boldsymbol{t}\int\psi'(x;K)\,\mathrm{d}F(x)\right\| \geq A_\eta((n-k)/n)^{1/2}\sqrt{\log n}\right) < n^{-\eta}, \quad \alpha \in (0,1/2) \tag{3.5}$$

for $1 \leq k \leq n$, where $\psi(.;K)$ is defined by (1.18).

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

$$Z_i(\boldsymbol{t}) = \psi(e_i - n^{-1/2}\boldsymbol{x}_i^T\boldsymbol{t};K) - \psi(e_i;K) - E\psi(e_i - n^{-1/2}\boldsymbol{x}_i^T\boldsymbol{t};K), \quad i = 1,\dots n$$

Then by the Markov inequality for each $\boldsymbol{t}$, $z > 0$ and $A > 0$

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

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

$$EZ_i^2(\boldsymbol{t}) \leq n^{-1}(\boldsymbol{x}_i^T\boldsymbol{t})^2 D_1$$

with some $D_1 > 0$ we obtain after few standard steps for $0 < z \leq \sqrt{n/k}$

$$P\left(\left|\sum_{i=1}^{k}\boldsymbol{x}_{ij}Z_i(\boldsymbol{t})\right| \geq A\right) \leq 2\exp\left\{-zA + z^2D_2k/n\right\}.$$

We want the right hand side smaller than $n^{-\eta}$ for an arbitrary but fixed $\eta > 0$. This will be obtained for $z = \sqrt{n/k}$ and any $A > \eta\sqrt{k/n}\log n$. Moreover,

$$\sum_{i=1}^{k} \boldsymbol{x}_i E\psi(e_i - n^{-1/2}\boldsymbol{x}_i^T\boldsymbol{t}; K) = -\int \psi'(x;K)\,\mathrm{d}F(x) n^{-1/2}\boldsymbol{C}_k\boldsymbol{t} + O_P(||\boldsymbol{t}||^2 k/n),$$

uniformly for $1 \le k \le n$. Hence for any $\eta > 0$ and $D > 0$ there exists $A_\eta > 0$ and $n_\eta$ such that for all $n \ge n_\eta$

$$P\left(\left\|\sum_{i=1}^{k} \boldsymbol{x}_i(\psi(e_i - n^{-1/2}\boldsymbol{x}_i^T\boldsymbol{t}; K) - \psi(e_i;K)) + n^{-1/2}\boldsymbol{C}_k\boldsymbol{t} \int \psi'(x;K)\,\mathrm{d}F(x)\right\| \ge A_\eta (k/n)^{1/2}\sqrt{\log n}\right) < n^{-\eta}, \quad \alpha \in (0,1/2) \tag{3.6}$$

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

$$P\left(\left|\sum_{i=1}^{k} \boldsymbol{x}_{ij}(Z_i(\boldsymbol{t}_1) - Z_i(\boldsymbol{t}_2))\right| \ge A\right) \le 2\exp\left\{-zA + z^2||\boldsymbol{t}_1 - \boldsymbol{t}_2||^2 D_3 k/n\right\}$$

with some $D_3 > 0$ and

$$\sum_{i=1}^{k} \boldsymbol{x}_i E(\psi(e_i - n^{-1/2}\boldsymbol{x}_i^T\boldsymbol{t}_1; K) - \psi(e_i - n^{-1/2}\boldsymbol{x}_i^T\boldsymbol{t}_2; K)) = -\int \psi'(x;K)\,\mathrm{d}F(x) n^{-1/2}\boldsymbol{C}_k(\boldsymbol{t}_1 - \boldsymbol{t}_2) + O(||\boldsymbol{t}_1 - \boldsymbol{t}_2||^2 k/n),$$

uniformly for $1 \le k \le n$. To finish the proof we apply Theorem 12.1 of Billingsley [1]. □

**Lemma 3.6.** Let the assumptions of Lemma 3.1 be satisfied. Then for any $\eta > 0$ and $D > 0$ there exist $A_\eta > 0$ and $n_\eta$ such that for all $n \ge n_\eta$

$$P\left(\sup_{||\boldsymbol{t}|| \le D;\, |u| \le D} \frac{1}{\sqrt{k}}\left\|\sum_{i=1}^{k} \boldsymbol{x}_i(\psi(e_i - n^{-1/2}\boldsymbol{x}_i^T\boldsymbol{t}; K + un^{-1/2}) - \psi(e_i - n^{-1/2}\boldsymbol{x}_i^T\boldsymbol{t}; K))\right\| \ge A_\eta n^{-1/4}\log n\right) < n^{-\eta}$$

$$P\left(\sup_{\|t\|\le D;|u|\le D}\frac{1}{\sqrt{n-k}}\left\|\sum_{i=k+1}^{n}\boldsymbol{x}_i(\psi(e_i-n^{-1/2}\boldsymbol{x}_i^T\boldsymbol{t};K+un^{-1/2})\right.\right.$$
$$\left.-\psi(e_i-n^{-1/2}\boldsymbol{x}_i^T\boldsymbol{t};K))\right\|$$
$$\left.\geq A_\eta n^{-1/4}\log n\right)<n^{-\eta}$$

for any $K>0$ and for $1\le k\le n$.

Proof. It is a modification of the proof of Theorem 4.1 in Jurečková and Sen [10] and Lemma 3.5 of the present paper, therefore it is omitted. □

Proof of Theorem 2.1. We show only (2.1) for the proof of (2.2) follows the same line.

We first notice that under the assumptions of Lemma 3.1 by Lemma 3.2 and Lemma 3.5 we have, as $n\to\infty$,

$$\max_{p<k<n-p}\frac{\sqrt{n}}{\sqrt{(n-k)k}}\left\|\sum_{i=1}^{k}\boldsymbol{x}_i(\psi(Y_i-\boldsymbol{x}_i^T\boldsymbol{\beta}_n(\psi(.;K)))-\psi(e_i;K))\right. \tag{3.7}$$
$$\left.-\boldsymbol{C}_k\boldsymbol{C}_n^{-1}\sum_{j=1}^{n}\boldsymbol{x}_j\psi(e_j;K))\right\|$$
$$=o_p((\log\log n)^{-1/2})$$

which in combination with (3.2) implies that (2.1) holds true for $\widehat{\psi}_n$ replaced by $\psi(.;K)$.

To finish the proof we notice that by Lemma 3.4, 3.5 and 3.6

$$\max_{p<k<n-p}\frac{1}{\sqrt{k}}\left\|\sum_{i=1}^{k}\boldsymbol{x}_i(\psi(Y_i-\boldsymbol{x}_i^T\boldsymbol{\beta}_n(\psi(.;F^{-1}(1-\alpha))\right. \tag{3.8}$$
$$\left.-\widehat{\psi}_n(Y_i-\boldsymbol{x}_i^T\boldsymbol{\beta}_n(\widehat{\psi}_n)))\right\|$$
$$=o_p((\log\log n)^{-1/2}).$$

and

$$\max_{p<k<n-p}\frac{1}{\sqrt{n-k}}\left\|\sum_{i=k+1}^{k}\boldsymbol{x}_i(\psi(Y_i-\boldsymbol{x}_i^T\boldsymbol{\beta}_n(\psi(.;F^{-1}(1-\alpha))\right. \tag{3.9}$$
$$\left.-\widehat{\psi}_n(Y_i-\boldsymbol{x}_i^T\boldsymbol{\beta}_n(\widehat{\psi}_n)))\right\|$$
$$=o_p((\log\log n)^{-1/2}).$$

This together with (3.2)–(3.3) and Lemma 3.3 implies that the assertion (2.1) holds true. □

Proof of Theorem 2.2. By (3.8)–(3.10) we have

$$\max_{p<k<n-p}\left(\frac{n}{(n-k)k}\right)^{1/2}\left\|C_n^{-1/2}\left(L_k-\sum_{i=1}^{k}\boldsymbol{x}_i(\psi(Y_i-\boldsymbol{x}_i^T\boldsymbol{\beta}_n(\psi(.;F^{-1}(1-\alpha))\right)\right\| \quad (3.10)$$
$$= o_p((\log\log n)^{-1/2}).$$

Moreover, for the weight function $q$ fulfilling (C.1) there exists a constant $D>0$ such that $q(s)\geq D$ for $s\in(\eta,1-\eta)$ and

$$\lim_{s\to 0+}\frac{q(s)}{\sqrt{s(1-s)}}=\infty,\quad \lim_{s\to 1-}\frac{q(s)}{\sqrt{s(1-s)}}=\infty.$$

The assertion (2.4) follows from this property, (3.10) and (3.3). The proof of the assertion (2.5) is the same and hence it is omitted. □

## ACKNOWLEDGEMENT

The author acknowledges very careful reading and helpful remarks of the referees.

(Received January 21, 2000.)

## 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 Series 31), Institute of Mathematical Statistics, Hayward, California 1997, pp. 56–70.

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

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

[11] 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.

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

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

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

[15] 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*