Sankhyā : The Indian Journal of Statistics
Special Issue on Quantile Regression and Related Methods
2005, Volume 67, Part 2, pp 200-226

Bootstrap in Detection of Changes in Linear Regression

Marie Hušková
Charles University in Prague, and Czech Academy of Sciences, Czech Republic
Jan Picek
Technical University of Liberec, Czech Republic

Abstract

Applications of bootstrap with and without replacement in change point analysis in linear regression models are discussed. Particularly, bootstrap based approximations for critical values for two classes of M-type test procedures are treated. As a particular case, we obtain L_1 procedures and regression quantile procedures. Their asymptotic performance is investigated and finite sample properties are checked in a simulation study.

AMS (2000) *subject classification.* 62G20, 62E20, 60F17.
Keywords and phrases. Linear regression, structural changes, M- test procedures, bootstrap with and without replacement.

1 Introduction

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

$$Y_{in} = \boldsymbol{h}^T(i/n)\boldsymbol{\beta} + \boldsymbol{h}^T(i/n)\boldsymbol{\delta}_n I\{i > m_n\} + e_i, \quad i = 1, \ldots, n, \tag{1.1}$$

where $m_n\,(\le n)$, $\boldsymbol{\beta} = (\beta_1, \ldots, \beta_p)^T$ and $\boldsymbol{\delta}_n = (\delta_{1n}, \ldots, \delta_{pn})^T \neq \mathbf{0}$ are unknown parameters, $\boldsymbol{h}^T(t) = (h_1(t), \ldots, h_p(t))^T$, $h_1(t) = 1$, $t \in [0,1]$ and $h_j(.)$, $j = 2, \ldots, p$, are continuously differentiable functions on $[0,1]$. Finally, $e_1, \ldots, e_n$ are independent identically distributed (i.i.d.) random errors fulfilling regularity conditions specified below. The function $I\{A\}$ denotes the indicator of the set 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 $n-m_n$ observations follow the linear regression model with the parameter $\boldsymbol{\beta}+\boldsymbol{\delta}_n$. The parameter m_n is usually called *the change point.* This is one of the basic models for change in linear regression. The main problems are to test whether a change has occurred and to estimate location of the change point m_n. There are many papers and even books and survey papers concerning such problems, e.g. Andrews (1993), Csörgő and Horváth (1997), Bai and Perron (1999) and Antoch, Hušková and Jarušková (2003) among others. Mostly, it is assumed that the design points $\boldsymbol{h}(i/n)$, $i = 1, \ldots, n$ have neither deterministic nor stochastic trend. However, it does not cover cases when the functions h_j, $j = 2, \ldots, p$, are smooth but nonconstant. We focus on such situations. In the past, they were treated in papers by Jandhyala and McNeill (1989, 1991,1997), Jandhyala and Minogue (1993) and Bischoff (1998).

We are interested in testing "no change" (H_0) versus "there is a change" (H_1), i.e.,

$$H_0 : m_n = n \qquad \text{against} \qquad H_1 : m_n < n \tag{1.2}$$

with particular attention to approximations of critical values based on residual bootstrap without or with replacement. Mostly, critical values based on limit distributions of the test statistics under H_0 are recommended. But in some cases the convergence is rather slow, e.g. if the limit distribution belongs to the extreme value type (see, e.g. Antoch and Hušková (2001, 2003)) or its explicit form is unknown or it depends on unknown parameters. It appears that bootstrap without replacement that coincides with permutation principles provides reasonable and simple approximations for critical values, see, e.g. Antoch and Hušková (2001, 2003), Hušková (2003) and Hušková and Picek (2002, 2004). We discuss this possibility in the model (1.1).

We consider M-test procedures based on the partial sums

$$\boldsymbol{S}_k(\psi) = \sum_{i=1}^{k} \boldsymbol{h}(i/n)\psi(Y_{in} - \boldsymbol{h}^T(i/n)\boldsymbol{\beta}_n(\psi)), \quad k = 1, ..., n, \tag{1.3}$$

and

$$S_{1,k}(\psi) = \sum_{i=1}^{k} \psi(Y_{in} - \boldsymbol{h}^T(i/n)\boldsymbol{\beta}_n(\psi)), \quad k = 1, \ldots, n, \tag{1.4}$$

where ψ is a score function and $\boldsymbol{\beta}_n(\psi)$ is an M- estimator of $\boldsymbol{\beta}$ in the model (1.1) with $m_n = n$ generated by a score function ψ, particularly, we assume that $\boldsymbol{\beta}_n(\psi)$ is a solution of the equation

$$\sum_{i=1}^{n} \boldsymbol{h}(i/n)\psi(Y_{in} - \boldsymbol{h}^T(i/n)\boldsymbol{z}) = \boldsymbol{0}. \tag{1.5}$$

The likelihood ratio (LR-) type test statistics (for details see Andrews, 1993) for our testing problem are

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

where

$$\boldsymbol{C}_k=\sum_{i=1}^{k}\boldsymbol{h}(i/n)\boldsymbol{h}^T(i/n), \quad \boldsymbol{C}_k^0=\boldsymbol{C}_n-\boldsymbol{C}_k, \quad k=1,...,n, \tag{1.7}$$

and $\widehat{\sigma}_n^2(\psi)$ is an estimator of $\sigma^2(\psi)=\int\psi^2(x)dF(x)$ with the property

$$\widehat{\sigma}_n^2(\psi)-\sigma^2(\psi)=o_p(1), \quad n\to\infty. \tag{1.8}$$

However, the test statistic (1.6) tends to infinity in probability even under H_0 and therefore Andrews (1993) recommends to take maximum over a suitably smaller set such that the test statistics is bounded in probability under H_0. One of the possibilities is to use the trimmed version of (1.6) defined as

$$T_n(\varepsilon)=\max_{\varepsilon n\le k<(1-\varepsilon)n}\left\{\boldsymbol{S}_k^T(\psi)\boldsymbol{C}_k^{-1}\boldsymbol{C}_n(\boldsymbol{C}_k^0)^{-1}\boldsymbol{S}_k(\psi)\frac{1}{\widehat{\sigma}_n^2(\psi)}\right\}, \tag{1.9}$$

where $0<\varepsilon<1/2$, however, the eventual change point m_n has to lie in $(\varepsilon n,(1-\varepsilon)n)$. Another possibility is to use different standardization, e.g.,

$$T_n(\gamma)=\max_{1\le k<n}\left\{\frac{1}{\widehat{\sigma}_n^2(\psi)}\Big(\frac{k}{n}\big(1-\frac{k}{n}\big)\Big)^{-2\gamma}\boldsymbol{S}_k^T(\psi)\boldsymbol{C}_n^{-1}\boldsymbol{S}_k(\psi)\right\}, \tag{1.10}$$

$$T_n^*(\gamma)=\max_{1\le k<n}\left\{\frac{1}{n\widehat{\sigma}_n^2(\psi)}\Big(\frac{k}{n}\big(1-\frac{k}{n}\big)\Big)^{-2\gamma}\boldsymbol{S}_k^T(\psi)\boldsymbol{S}_k(\psi)\right\}, \tag{1.11}$$

where $\gamma\in[0,1/2)$. All the mentioned test statistics have the form

$$T_n(\psi,\boldsymbol{Q}_n)=\sup_{1\le k<n}\left\{\boldsymbol{S}_k^T(\psi)\boldsymbol{Q}_{k,n}\boldsymbol{S}_k(\psi)\frac{1}{\widehat{\sigma}_n^2(\psi)}\right\}, \tag{1.12}$$

where $\boldsymbol{Q}_{k,n}$, $k=1,\dots,n$, are symmetric positive semidefinite matrices. Under quite weak assumptions such test statistics converge under H_0 to almost surely finite random variables.

Analogously Bayesian like test statistics can be can introduced as

$$Z_n(\psi;\boldsymbol{Q}_n)=\frac{1}{\widehat{\sigma}_n^2(\psi)}\frac{1}{n}\sum_{i=1}^{n}\boldsymbol{S}_k^T(\psi)\boldsymbol{Q}_{k,n}\boldsymbol{S}_k(\psi), \tag{1.13}$$

where $\boldsymbol{Q}_{k,n}$, $k = 1, \dots,$ are as above.

The simpler test procedures are based on $S_{11}(\psi), \dots, S_{1n}(\psi)$ and they have the forms

$$T_n^0(\psi, q_{k,n}) = \sup_{0<t<1} \left\{ \frac{1}{\widehat{\sigma}_n(\psi)} |S_{1,\lfloor (n+1)t \rfloor}(\psi)| q_{\lfloor (n+1)t \rfloor, n} \right\}, \tag{1.14}$$

$$\text{and } Z_n^0(\psi, q_{k,n}) = \frac{1}{\widehat{\sigma}_n(\psi) n} \sum_{k=1}^{n} \big| S_{1,k}(\psi) \big| q_{k,n}, \tag{1.15}$$

where $q_{k,n}$ are nonnegative numbers with properties specified below.

Test statistics $T_n(\psi, \boldsymbol{Q}_{k,n})$ and $T_n^0(\psi, q_{k,n})$ are *weighted max type test statistics*, while $Z_n(\psi, \boldsymbol{Q}_{k,n})$ and $Z_n^0(\psi, q_{k,n})$ are *Bayesian type test statistics.*

Notice that for $\psi(x) = x$, we have classical L_2-procedures related to the least squares estimators, for $\psi(x) = \operatorname{sign} x$, it leads to L_1 -type test. Finally, for

$$\psi_\beta(x) = \beta - I\{x < 0\}, \quad x \in R^1,\ \beta \in (0,1), \tag{1.16}$$

we obtain procedures related to the β-regression quantiles.

Large values of the above test statistics indicate that the null hypothesis is violated. Therefore critical regions have the forms

$$T_n(\psi, \boldsymbol{Q}_n) > c_n(\alpha, \psi, \boldsymbol{Q}_n) \tag{1.17}$$

and

$$Z_n(\psi, \boldsymbol{Q}_n) > d_n(\alpha, \psi, \boldsymbol{Q}_n), \tag{1.18}$$

where $c_n(\alpha, \psi, \boldsymbol{Q}_n)$ and $d_n(\alpha, \psi, \boldsymbol{Q}_n)$ are critical values corresponding to the level α. Denote the critical values corresponding to $T_n^0(\psi, q_n)$ and $Z_n^0(\psi, q_n)$ by $c_n^0(\alpha, \psi, q_n)$ and $d_n^0(\alpha, \psi, q_n)$, respectively.

Approximations to these critical values can be obtained through the limit distributions of these test statistics under the null hypothesis, however, their explicit forms are generally unknown. As follows from Theorem 2.1 below, the limit distributions under H_0 are distributions of various functionals of Gaussian processes. For some particular cases of the test statistics, Jandhyala and MacNeill (1989, 1991, 1997) proposed numerical approximations for critical values.

Here we propose another possibility based on an application of permutation arguments modified for the situation of regression models. It coincides with bootstrapping based on residuals without replacement. Bootstrap based residuals with replacement provides approximation for the critical values also. Both methods provide good approximations for the critical values and are relatively simple.

Next, we explain shortly the application of the permutation arguments in our problem. For more detailed explanation see, e.g. Antoch and Hušková (2003), Hušková (2004) and Hušková and Picek (2002). By permutation principle, since the random errors $e_1, \dots, e_n$ are i.i.d. random variables, the permutation version of our test statistics should be obtained by replacing $e_1, \dots, e_n$ by $e_{R_1}, \dots, e_{R_n}$, where $R_1, \dots, R_n$ is a random permutation of $1, \dots, n$ independent of $Y_{1n}, \dots, Y_{nn}$. However, $e_1, \dots, e_n$ are not known, and therefore we permute $\widehat{e}_1(\psi), \dots, \widehat{e}_n(\psi)$, where $\widehat{e}_i(\psi)$ is the M-residual defined by (1.20). We have to retain the basic properties of the partial sums $\boldsymbol{S}_k(\psi)$. In particular, we try to retain the property:

$$\boldsymbol{S}_k(\psi) = \sum_{i=1}^{k} \boldsymbol{h}(i/n)\widehat{e}_i(\psi) - \boldsymbol{C}_k \boldsymbol{C}_n^{-1} \sum_{j=1}^{n} \boldsymbol{h}(j/n)\widehat{e}_j(\psi),\ k = 1, \dots, n, \quad (1.19)$$

where

$$\widehat{e}_i(\psi) = \psi(Y_{in} - \boldsymbol{h}^T(i/n)\boldsymbol{\beta}_n(\psi)), \quad i = 1, \dots, n. \quad (1.20)$$

Therefore we define the permutation version $T_n(\psi, \boldsymbol{Q}_n; \boldsymbol{R})$ of $T_n(\psi, \boldsymbol{Q}_n)$ by (1.12) with $\boldsymbol{S}_k(\psi, \gamma)$, $k = 1, \dots, n$, replaced by

$$\boldsymbol{S}_k(\psi, \boldsymbol{R}) = \sum_{i=1}^{k} \boldsymbol{h}(i/n)\widehat{e}_{R_i}(\psi) - \boldsymbol{C}_k \boldsymbol{C}_n^{-1} \sum_{j=1}^{n} \boldsymbol{h}(j/n)\widehat{e}_{R_j}(\psi), \quad k - 1, \dots, n. \quad (1.21)$$

Critical values $c_n(\alpha, \psi, \boldsymbol{Q}_n)$ and $d_n(\alpha, \psi, \boldsymbol{Q}_n)$ are then approximated by conditional (given $\boldsymbol{Y}_n = (Y_{1n}, \dots, Y_{nn})$) $100(1-\alpha)\%$ quantiles of $T_n(\psi, \boldsymbol{Q}_n; \boldsymbol{R})$ and $Z_n(\psi, \boldsymbol{Q}_n; \boldsymbol{R})$, respectively. Denoting these conditional quantiles by $c_n(\alpha, \psi, \boldsymbol{Q}_n; \boldsymbol{Y}_n)$ and $d_n(\alpha, \psi, \boldsymbol{Q}_n; \boldsymbol{Y}_n)$, respectively, the resulting procedures have the critical regions:

$$T_n(\psi, \boldsymbol{Q}_n) > c_n(\alpha, \psi, \boldsymbol{Q}_n; \boldsymbol{Y}_n) \quad (1.22)$$

and

$$Z_n(\psi, \boldsymbol{Q}_n) > d_n(\alpha, \psi, \boldsymbol{Q}_n; \boldsymbol{Y}_n). \quad (1.23)$$

The permutation versions $T_n^0(\psi, q_n; \boldsymbol{R}), Z_n^0(\psi, q_n; \boldsymbol{R})$ and related test procedures are defined accordingly. Notice that the above described permutation procedure can be also called bootstrap based on residuals without replacement.

It is necessary to show that the above proposed approximations for the critical values are at least asymptotically correct when data follow either H_0 or some alternatives. Toward this it suffices to show that given $\boldsymbol{Y}_n$ the asymptotic conditional distribution of $T_n(\psi, \boldsymbol{Q}_n; \boldsymbol{R})$ and $Z_n(\psi, \boldsymbol{Q}_n; \boldsymbol{R})$ coincide with the unconditional limit distributions of $T_n(\psi, \boldsymbol{Q}_n)$ and $Z_n(\psi, \boldsymbol{Q}_n)$ under H_0, respectively. This is a direct consequence of Theorems 2.1 and 2.3 below. We discuss also closeness of the distributions of $T_n(\psi, \boldsymbol{Q}_n)$ and $Z_n(\psi, \boldsymbol{Q}_n; \boldsymbol{R})$ under H_0.

Towards computational aspects, notice that the conditional distribution of $T_n(\psi, \boldsymbol{Q}_n; \boldsymbol{R})$ and of the other permutation versions of the test statistics, given the original observations $\boldsymbol{Y}_n$, is determined by the random permutation $\boldsymbol{R} = (R_1, \ldots, R_n)$ whose distribution is also known and therefore the conditional distribution (given $\boldsymbol{Y}_n$) of $T_n(\psi, \boldsymbol{Q}_n; \boldsymbol{R})$ is known and can be calculated, which amounts to calculating $T_n(\psi, \boldsymbol{Q}_n; \boldsymbol{r})$ for all permutations $\boldsymbol{r}$ of $(1, \ldots, n)$. Since the number of possible permutations $\boldsymbol{r}$ is $n!$, one can calculate $T_n(\psi, \boldsymbol{Q}_n; \boldsymbol{r})$ for all permutations only for very small n. However, one can calculate a reasonable approximation for large or moderately large n. Specifically, one chooses independent and random permutations $\boldsymbol{r}_1, \ldots, \boldsymbol{r}_B$, where B is large enough but still affordable with computers. Details are in Section 3. With the other test statistics, one proceeds analogously.

In Section 2, the limiting behaviours of the proposed test statistics both under H_0 and the alternatives are derived. Also, the conditional limiting behaviours of $T_n(\psi, \boldsymbol{Q}_n; \boldsymbol{R})$, $T_n^0(\psi, q_n; \boldsymbol{R})$, $Z_n(\psi, \boldsymbol{Q}_n; \boldsymbol{R})$ and $Z_n^0(\psi, q_n; \boldsymbol{R})$ are investigated under quite general assumptions. Section 3 contains simulation study while the proofs are in Section 4.

In the rest of the paper, we omit the index n whenever possible, e.g. we write Y_i and m instead of Y_{in} and m_n, respectively.

2 Asymptotic results

In this section, we investigate and discuss the limit distributions of the considered test statistics and their permutation versions.

• We assume that the vector of functions $\boldsymbol{h}(.) = (h_1(.), \ldots, h_p(.))^T$ on $[0,1]$ and the matrices $\boldsymbol{Q}_{k,n}$, $k = 1, \ldots, n$ satisfy:

(A.1) $h_1(x) = 1$, $x \in [0,1]$,

(A.2) $h_2(.), \ldots, h_p(.)$ are continuously differentiable functions on $[0,1]$ such that $\int_0^1 h_j(t)dt = 0$, $j = 2, \ldots, p$ and the $p \times p$ matrix functions $\boldsymbol{C}(t) = \left(\int_0^1 h_j(x) h_v(x) dx\right)_{j,v=1,\ldots,p}$, $t \in [0,1]$ and $\boldsymbol{C}(1) - \boldsymbol{C}(t)$ are regular for each $t \in (0,1]$ and $t \in [0,1)$, respectively.

(A.3) there are symmetric positive semidefinite matrices $\boldsymbol{Q}(t)$, $t \in (0,1)$ such that

$$\lim_{n\to\infty} \boldsymbol{Q}_{\lfloor (n+1)t\rfloor,n} = \boldsymbol{Q}(t),\ t \in (0,1), \tag{2.1}$$

$$\lim_{t\to 0^+} t^\eta \boldsymbol{Q}(t) = \mathbf{0}, \quad \lim_{t\to 0^+} (1-t)^\eta \boldsymbol{Q}(1-t) = \mathbf{0}$$

for some $\eta \in [0,1/2)$, $\boldsymbol{Q}(t) = \boldsymbol{Q}(1-t)$, $t \in (0,1)$ and on each closed subinterval of $(0,1)$, the elements of $\boldsymbol{Q}(t)$ as a function of t are bounded and the convergence in (2.1) is uniform.

(A.4) there is a nonnegative function q on $(0,1)$ such that

$$\lim_{n\to\infty} q_{\lfloor (n+1)t\rfloor,n} = q(t),\ t \in (0,1) \tag{2.2}$$

$$\lim_{t\to 0^+} t^\eta q(t) = 0, \quad \lim_{t\to 0^+} (1-t)^\eta q(1-t) = 0$$

for some $\eta \in [0,1/2)$, $q(t) = q(1-t)$, $t \in (0,1)$ and on each closed subinterval of $(0,1)$, $q(t)$ is bounded and the convergence in (2.2) is uniform.

- The distribution of the error terms e_i's satisfies the following assumption:

(B.1) $e_1, \dots, e_n$ are i.i.d. random variables with common symmetric distribution function F.

- The score function ψ and the function $\lambda(t) = -\int \psi(e-t)dF(e)$, $t \in R^1$, satisfy:

(C.1) ψ is non-decreasing, antisymmetric;

(C.2) the derivative $\lambda'(.)$ of the function $\lambda(.)$ exists and is Lipschitz in a neighbourhood of zero, $\lambda(0) = 0$ and $\lambda'(0) > 0$;

(C.3) $\int \psi^2(x)dF(x) \in (0,\infty)$, and for some $a > 0$ and $D_1 > 0, D_2 > 0$,

$$\int (\psi(x+t_2) - \psi(x-t_1))^2 dF(x) \le D_1|t_2 - t_1|^a, \quad |t_j| \le D_2,\ j = 1,2.$$

The above assumptions on the distribution of the error terms and the score function are traditional in nature. Concerning the design points $\boldsymbol{h}(i/n)$, $i = 1, \ldots, n$, quite often one assumes that, as $s \to \infty$, $(\boldsymbol{C}_{k+s} - \boldsymbol{C}_k)/s$ is close to a regular matrix $\boldsymbol{C}$ uniformly in k which is not generally satisfied under (A.2). The assumption (A.2) covers important situations like polynomial and harmonic polynomial regression. Such situations were studied by MacNeill and Jandhyala (1989, 1991, 1997) and Bischoff (1998). Particular cases of $\boldsymbol{Q}_{k,n}$, $k = 1, \ldots, n-1$, defined in (1.9)-(1.11) fulfill the assumption (A.3). Particular choices of $q_{k,n}$, $k = 1, \ldots, n-1$, fulfilling (A.4) are

$$q_{k,n}(\varepsilon, \gamma) = \frac{1}{\sqrt{n}} \left(\frac{k(n-k)}{n^2} \right)^{-\gamma}, \ \varepsilon n < k < (1-\varepsilon)n,$$

$\varepsilon \in (0, 1/2)$ and $\gamma \in [0, 1/2]$ or $\varepsilon = 0$ and $\gamma \in [0, 1/2)$.

Now, we formulate the main results on limiting behaviour of the test statistics both under H_0 and local alternatives and conditional limiting behaviour of their permutation versions.

THEOREM 2.1 *Let $\boldsymbol{Y} = (Y_1, \ldots, Y_n)$ follow the model (1.1) with $m = n$, i.e. "no change." Moreover, let assumptions (A.1)–(A.2), (B.1) and (C.1)-(C.3) be satisfied. Let $\widehat{\sigma}_n^2(\psi)$ satisfy (1.8).*

(i) If moreover (A.3) is satisfied, we have as $n \to \infty$,

$$T_n(\psi, \boldsymbol{Q}_n) \to^d \sup_{0<t<1} \boldsymbol{S}^T(t, \psi) \boldsymbol{Q}(t) \boldsymbol{S}(t, \psi) \tag{2.3}$$

and

$$Z_n(\psi, \boldsymbol{Q}_n) \to^d \int_{0<t<1} \boldsymbol{S}^T(t, \psi) \boldsymbol{Q}(t) \boldsymbol{S}(t, \psi) dt. \tag{2.4}$$

(ii) Moreover if (A.4) is satisfied, we have as $n \to \infty$,

$$T_n^0(\psi, q_n) \to^d \sup_{0<t<1} |S_1(t, \psi)| q(t) \tag{2.5}$$

and

$$Z_n^0(\psi, q_n) \to^d \int_{0<t<1} |S_1(t, \psi)| q(t) dt, \tag{2.6}$$

where

$$\boldsymbol{S}(t, \psi) = \int_0^t \boldsymbol{h}(x) dB(x) - \boldsymbol{C}(t) \boldsymbol{C}^{-1}(1) \int_0^1 \boldsymbol{h}(x) dB(x), \quad t \in [0, 1] \tag{2.7}$$

with $\{B(x),\, x \in [0,1]\}$ being a Brownian bridge, and $S_1(t,\psi)$ is the first component of $\boldsymbol{S}(t,\psi)$, $t \in [0,1]$. The assertions on $T_n^0(\psi, q_n)$ and $Z_n^0(\psi, q_n)$ remain true if the assumption (A.2) is replaced by
(A.2)* $\lim_{n\to\infty} \frac{1}{n}\boldsymbol{C}_n = \boldsymbol{C}$, *where* $\boldsymbol{C} > 0$.

THEOREM 2.2 *Let $\boldsymbol{Y}$ follow the model (1.1) with $m = \lfloor \kappa n \rfloor$, $\kappa \in (0,1)$ and with $\boldsymbol{\delta} = \boldsymbol{\delta}_n$ satisfying*

$$\lim_{n\to\infty} ||\boldsymbol{\delta}_n|| = 0, \quad \lim_{n\to\infty} ||\boldsymbol{\delta}_n||\sqrt{n} = \infty. \tag{2.8}$$

Let there exist $\lambda \in (0, 1/2)$ and $D > 0$ such that for all $\boldsymbol{\Delta} \neq 0$

$$\inf_{t\in(\lambda,1-\lambda)} \boldsymbol{\Delta}^T \boldsymbol{Q}(t)\boldsymbol{\Delta} \geq D||\boldsymbol{\Delta}||^2.$$

Then, under the assumptions (A.1)-(A.3), (B.1) and (C.1)-(C.3), as $n \to \infty$,

$$T_n(\psi, \boldsymbol{Q}_n) \to^P \infty \tag{2.9}$$

and

$$Z_n(\psi, \boldsymbol{Q}_n) \to^P \infty. \tag{2.10}$$

If instead of (A.3) the assumption (A.4) holds and there are $\lambda > 0$ and $d > 0$ such that $q(t) \geq d$ for $t \in (\kappa - \lambda, \kappa + \lambda)$, and if

$$\left| \int_0^{\kappa} \boldsymbol{h}^T(t) dt \boldsymbol{C}^{-1}(1) \boldsymbol{C}^0(\kappa) \boldsymbol{\delta}_n \right| \sqrt{n} \to \infty,$$

then, as $n \to \infty$,

$$T_n^0(\psi, q_n) \to^P \infty \tag{2.11}$$

and

$$Z_n^0(\psi, q_n) \to^P \infty. \tag{2.12}$$

THEOREM 2.3 *Let $\boldsymbol{Y}$ follow the model (1.1) with $\lim_{n\to\infty} ||\delta_n|| = 0$. Then, under the assumptions (A.1) – (A.3),(B.1) and (C.1)-(C.3) for arbitrary $y \in R^1$, as $n \to \infty$,*

$$P\big(T_n(\psi, \boldsymbol{Q}_n; \boldsymbol{R}) \leq y | \boldsymbol{Y}_n\big) \to^P P\left(\sup_{0<t<1} \boldsymbol{S}^T(t,\psi)\boldsymbol{Q}(t)\boldsymbol{S}(t,\psi) \leq y \right), \tag{2.13}$$

$$P\big(Z_n(\psi, \boldsymbol{Q}_n; \boldsymbol{R}) \le y|\boldsymbol{Y}_n\big) \to^P P\left(\int_{0<t<1} \boldsymbol{S}^T(t,\psi)\boldsymbol{Q}(t)\boldsymbol{S}(t,\psi)dt \le y\right). \tag{2.14}$$

If the assumptions (A.2) and (A.3) are replaced by the assumptions (A.2) and (A.4), then for all* $y \in R^1$, *as* $n \to \infty$,

$$P(T_n^0(\psi, q_n; \boldsymbol{R})) \le y|\boldsymbol{Y}_n) \to^P P\left(\sup_{0<t<1} |S_1(t,\psi)q(t)| \le y\right) \tag{2.15}$$

and

$$P((Z_n^0(\psi, \lambda; \boldsymbol{R}))^{1/2} \le y|\boldsymbol{Y}_n) \to^P P\left(\int_0^1 |S_1(t,\psi)q(t)|dt \le y\right). \tag{2.16}$$

Proofs are postponed to Section 4.

By Theorem 2.1 under H_0 the proposed test statistics are asymptotically distributed as functionals of Gaussian processes $\boldsymbol{S}(t,\psi)$, $t \in [0,1]$, defined in (2.7). The explicit forms of their distributions are known only for few particular cases. Approximations to these distributions can be obtained through simulations, e.g., simulating limit distributions through the respective Gaussian processes. Under the assumptions of Theorem 2.1, the limit random variables, i.e., random variables on the rhs of (2.3) – (2.6) are a.s. finite.

Jandhyala and MacNeill (1989, 1991, 1997) treated Bayesian type test statistics with $\psi(x) = x$, $x \in R^1$ and derived numerical approximations for critical values based on limit distributions for some particular choices of the functions h_j, $j = 1, \dots, p$. Bischoff (1998) considered $T_n^0(\psi, q_n)$ with $\psi(x) = x$, $x \in R^1$.

Concerning a suitable estimator of $\sigma^2(\psi)$, under the considered setup

$$\widehat{\sigma}_n^2(\psi) = \frac{1}{2n}\sum_{i=2}^n (\widehat{e}_i(\psi) - \widehat{e}_{i-1}(\psi))^2 \tag{2.17}$$

has the desired property (1.8) under H_0 and local as well fixed alternatives while

$$\widetilde{\sigma}_n^2(\psi) = \frac{1}{n}\sum_{i=1}^n \widehat{e}_i^2(\psi) \tag{2.18}$$

has this property only under H_0 and local alternatives but not under fixed ones.

Theorems 2.1 and 2.3 immediately imply that the permutation procedure provides asymptotically correct approximations for the null distributions of the test statistics when data follow either H_0 or local alternatives. Therefore the values $c_n(\alpha, \psi, \boldsymbol{Q}_n, \boldsymbol{Y})$ and $d_n(\alpha, \psi, \boldsymbol{Q}_n, \boldsymbol{Y})$, defined in (1.22) and (1.23), also provide asymptotically correct approximations for the desired critical values. Similar results hold true for the tests based on $T_n^0(\psi, q_n)$ and $Z_n^0(\psi, q_n)$.

By Theorems 2.1, 2.2 and 2.3, the tests with critical regions (1.17) and (1.18) are consistent under local alternatives formulated in Theorem 2.2.

The question that remains to be answered is whether the permutation arguments provide asymptotically correct or at least reasonable approximation even under fixed alternatives, i.e. $\boldsymbol{\delta} \neq \mathbf{0}$ fixed. In this case, under additional assumptions on the score function ψ (e.g., ψ bounded or $\psi(x) = x$, $x \in R^1$), one can show that the assertions of Theorem 2.3 remain true if the permutation versions of test statistics are differently standardized, e.g., the assertion (2.13) remains true even for fixed alternatives if $T_n(\psi, \boldsymbol{Q}_n; \boldsymbol{R})$ is replaced by

$$T_n(\psi, \boldsymbol{Q}_n; \boldsymbol{R})\widehat{\sigma}_n^2(\psi)/\widetilde{\sigma}_n^2(\psi),$$

where $\widehat{\sigma}_n^2(\psi)$ and $\widetilde{\sigma}_n^2(\psi)$ are defined by (2.17) and (2.18), respectively. It can be shown that for $\psi(x) = x$, $x \in R^1$, as $n \to \infty$,

$$\frac{\widetilde{\sigma}_n^2(\psi)}{\widehat{\sigma}_n^2(\psi)} \to^P 1 + \Big(\boldsymbol{\delta}^T \boldsymbol{C}(\kappa)\boldsymbol{C}^{-1}(1)\boldsymbol{C}^0(\kappa)\boldsymbol{C}^{-1}(1)\boldsymbol{C}(\kappa)\boldsymbol{\delta}\Big)/\sigma^2(\psi).$$

Another possibility to get asymptotically correct approximations of critical values when data follow fixed alternatives is to permute residuals that take into account a possible change point. Specifically, one estimates the change point m, e.g., the estimator $\breve{m}$ defined as the index k maximizing $\boldsymbol{S}_k^T(\psi)\boldsymbol{Q}_{k,n}\boldsymbol{S}_k(\psi)$ can be used. Then in (1.21), the permutations of the residuals $\widehat{e}_i(\psi)$'s are replaced by permutations of residuals $\breve{e}_i(\psi)$, $i = 1, \dots, n$, where $\breve{e}_i(\psi)$, $i = 1, \dots, \breve{m}$ and $\breve{e}_i(\psi)$, $i = \breve{m} + 1, \dots, n$ are M- residuals corresponding to the observations $Y_1, \dots, Y_{\breve{m}}$ and $Y_{\breve{m}+1}, \dots, Y_n$. This provides asymptotically correct approximation even for fixed alternatives.

Going carefully through the proofs, one finds that the bootstrap with replacement based on residuals also provides asymptotic approximation for the null distribution of the test statistics and the corresponding critical values.

The above theorems ensure that the simulations of the limit distribution, permutation procedure and bootstrap with replacement provide asymptotically correct approximation for the null distribution of our test statistics.

However, the permutation procedure provides better approximation under H_0 than the remaining two methods. Particularly, applying the same tools as in the proofs of Theorems 2.1 and 2.3 we can show that under the assumptions of Theorem 2.1 with $m = n$, for $\gamma < 1/2$, as $n \to \infty$,

$$\max_{1 \le k < n} \frac{1}{\sqrt{n}} \Big(\frac{k(n-k)}{n^2}\Big)^{-\gamma} || \sum_{i=1}^{k} \boldsymbol{x}_i(\widehat{e}_i(\psi) - \psi(e_i)) || = O_P(n^{-b}) \tag{2.19}$$

and

$$\max_{1 \le k < n} \frac{1}{\sqrt{n}} \Big(\frac{k(n-k)}{n^2}\Big)^{-\gamma} \Big\| \sum_{i=1}^{k} \boldsymbol{x}_i(\widehat{e}_{R_i}(\psi) - \psi(e_{R_i})) - \boldsymbol{C}_k \boldsymbol{C}_n^{-1} \sum_{j=1}^{n} \boldsymbol{x}_i(\widehat{e}_{R_i}(\psi) - \psi(e_{R_i})) \Big\| = o_P(n^{-b})$$

with some $b > 0$, which implies that under H_0, as $n \to \infty$,

$$T_n(\psi, \boldsymbol{Q}_n) =^d T_n(\psi, \boldsymbol{Q}_n; \boldsymbol{R}) + O_P(n^{-b})$$

with some $b > 0$ and similarly for the other test statistics.

Typically, the case of regression quantile type procedures is not covered by assumptions (B.1) and (C.1)-(C.3). However, using the results on the regression quantiles (e.g., Jurečková and Sen, 1996) we find that Theorems 2.1-2.3 remain true even for case $\psi_\beta(.)$ defined by (1.16) when the underlying distribution F has continuous second derivative in a neighbourhood of $F^{-1}(\beta)$ and, moreover, $F^{(1)}(F^{-1}(\beta)) > 0$.

Test procedures and approximations for critical values can be developed along the same lines for testing problems where the change concerns only some components of the parameter $\boldsymbol{\beta}$, as well as for problems that involve multiple changes.

3 Simulations

In order to check how the proposed test procedures perform for finite sample situations, we have conducted a simulation study.

We considered the model

$$Y_i = \beta_1 + \beta_2(i/n - 1/2) + (\delta_1 + \delta_2(i/n - 1/2))I\{i > m\} + e_i, \quad i = 1 \ldots, n, \tag{3.1}$$

where the errors were simulated from the normal $N(0,1)$ and Laplace $L(0,1)$ distributions.

The following parameter values were used:

- sample sizes: $n = 60,\ 200$;
- $\boldsymbol{\beta} = (1,1)$;
- $\boldsymbol{\delta}$= (0,1), (5,5);
- change points: $m = 3, 6, 9, \dots, 60$ for $n = 60$ and $m = 50, 100, 150, 200$ for $n = 200$;
- score functions: $\psi(x) = \text{sign}(x)$ and the Huber score generating function
$$\psi_H(x) = \begin{cases} x & \text{if } |x| \le 1.345 \\ 1.345 \,\text{sign}\, x & \text{if } |x| > 1.345. \end{cases}$$

Note that the choice $m = n$ corresponds to the null hypothesis H_0, i.e. "no change." We take $\boldsymbol{\delta} = (0,1)$ to represent smaller changes while $\boldsymbol{\delta} = (5,5)$ corresponds to larger changes.

We consider the test statistics:

$$T_n = \max_{1\le k<n} \frac{1}{\sqrt{n}\widehat{\sigma}_n}\Big(\boldsymbol{S}_k^T \boldsymbol{S}_k\Big), \quad Z_n = \max_{1\le k<n} \frac{1}{\sqrt{n}\widehat{\sigma}_n}\frac{1}{n}\sum_{i=1}^{n}\Big(\boldsymbol{S}_k^T \boldsymbol{S}_k\Big), \tag{3.2}$$

$$T_n^0 = \max_{1\le k<n} \frac{1}{\sqrt{n}\widehat{\sigma}_n}|S_{1,k}|, \quad Z_n^0 = \frac{1}{\widehat{\sigma}_n n^{3/2}}\sum_{i=1}^{n}|S_{1,k}| \tag{3.3}$$

i.e. we chose $\boldsymbol{Q}_{k,n} = n^{-1}\boldsymbol{I}_2$, $k = 1, \dots, n$, in (1.12) and (1.13) and $q_{k,n} = n^{-1/2}$, $k = 1, \dots, n$, in (1.14) and (1.15). $T_n(\boldsymbol{R})$, $Z_n(\boldsymbol{R})$ $T_n^0(\boldsymbol{R})$ and $Z_n^0(\boldsymbol{R})$ denote their respective permutation versions.

For each combination of the parameters, we proceeded as follows:
(1) we computed $\boldsymbol{\beta}_n(\psi)$ and $\hat{e}_i(\psi)$, $i = 1, \dots, n$.;
(2) a random permutation $\mathbf{r}$ of $(1, \dots, n)$ is generated;
(3) $T_n(\boldsymbol{R})$, $Z_n(\boldsymbol{R})$ $T_n^0(\boldsymbol{R})$ and $Z_n^0(\boldsymbol{R})$ with $\mathbf{R} = \mathbf{r}$ are calculated;
(4) steps (2) and (3) are repeated 1000 times;
(5) sample quantiles are computed.
Selected sample quantiles of $T_n(\boldsymbol{R})$, $Z_n(\boldsymbol{R})$, $T_n^0(\boldsymbol{R})$ and $Z_n^0(\boldsymbol{R})$ are summarized in Tables 1-4 and plotted in Figures 1-2 for the sample size $n = 60$.

Table 1. Simulated quantiles for $T_n(\mathbf{R})$ for normally and Laplace distributed errors.

n	m	$\boldsymbol{\delta}$	$\psi(x) = \operatorname{sign} x$				Huber ψ			
			normal		Laplace		normal		Laplace	
			95%	99%	95%	99%	95%	99%	95%	99%
60	15	(0,1)	0.868	1.207	0.921	1.390	0.803	1.098	0.865	1.127
60	30	(0,1)	0.861	1.242	0.838	1.135	0.857	1.180	0.840	1.282
60	45	(0,1)	0.870	1.200	0.907	1.196	0.797	1.202	0.843	1.190
60	60	(0,1)	0.855	1.251	0.906	1.329	0.863	1.189	0.856	1.218
60	15	(5,5)	0.846	1.285	0.870	1.266	0.819	1.167	0.880	1.141
60	30	(5,5)	0.882	1.257	0.884	1.326	0.828	1.182	0.860	1.181
60	45	(5,5)	0.876	1.369	0.845	1.167	0.893	1.261	0.865	1.300
60	60	(5,5)	0.950	1.452	0.926	1.334	0.881	1.172	0.867	1.181
200	50	(0,1)	0.859	1.239	0.918	1.258	0.869	1.185	0.828	1.096
200	100	(0,1)	0.933	1.203	0.874	1.122	0.865	1.253	0.858	1.219
200	150	(0,1)	0.842	1.151	0.822	1.230	0.848	1.176	0.844	1.161
200	200	(0,1)	0.834	1.148	0.857	1.150	0.866	1.224	0.858	1.209
200	50	(5,5)	0.909	1.210	0.893	1.193	0.856	1.206	0.828	1.125
200	100	(5,5)	0.882	1.255	0.876	1.199	0.903	1.174	0.884	1.162
200	150	(5,5)	0.875	1.244	0.871	1.232	0.900	1.185	0.879	1.214
200	200	(5,5)	0.852	1.165	0.866	1.236	0.874	1.289	0.833	1.082

Table 2. Simulated quantiles for $Z_n(\boldsymbol{R})$ for normally and Laplace distributed errors.

n	m	$\boldsymbol{\delta}$	$\psi(x) = \operatorname{sign} x$				Huber ψ			
			normal		Laplace		normal		Laplace	
			95%	99%	95%	99%	95%	99%	95%	99%
60	15	(0,1)	0.188	0.283	0.205	0.292	0.187	0.257	0.184	0.305
60	30	(0,1)	0.194	0.291	0.166	0.261	0.193	0.270	0.205	0.301
60	45	(0,1)	0.179	0.271	0.194	0.275	0.175	0.268	0.196	0.285
60	60	(0,1)	0.185	0.272	0.200	0.299	0.184	0.278	0.186	0.287
60	15	(5,5)	0.185	0.273	0.199	0.301	0.185	0.276	0.194	0.285
60	30	(5,5)	0.196	0.319	0.186	0.337	0.188	0.286	0.181	0.285
60	45	(5,5)	0.178	0.269	0.179	0.264	0.198	0.288	0.193	0.323
60	60	(5,5)	0.194	0.328	0.192	0.294	0.191	0.275	0.195	0.294
200	50	(0,1)	0.180	0.279	0.180	0.261	0.175	0.265	0.170	0.244
200	100	(0,1)	0.191	0.273	0.173	0.245	0.174	0.267	0.171	0.270
200	150	(0,1)	0.179	0.260	0.167	0.255	0.170	0.248	0.169	0.270
200	200	(0,1)	0.166	0.278	0.177	0.257	0.174	0.276	0.186	0.263
200	50	(5,5)	0.184	0.280	0.176	0.255	0.176	0.271	0.166	0.229
200	100	(5,5)	0.178	0.283	0.174	0.261	0.185	0.260	0.182	0.258
200	150	(5,5)	0.181	0.269	0.188	0.282	0.177	0.270	0.178	0.289
200	200	(5,5)	0.176	0.260	0.181	0.308	0.177	0.282	0.170	0.253

TABLE 3. SIMULATED QUANTILES FOR $T_n^0(\boldsymbol{R})$ FOR NORMALLY AND LAPLACE DISTRIBUTED ERRORS.

n	m	$\boldsymbol{\delta}$	$\psi(x) = \operatorname{sign} x$				Huber ψ			
			normal		Laplace		normal		Laplace	
			95%	99%	95%	99%	95%	99%	95%	99%
60	15	(0,1)	0.885	1.041	0.911	1.112	0.858	1.004	0.881	1.010
60	30	(0,1)	0.880	1.081	0.863	1.014	0.874	1.044	0.879	1.088
60	45	(0,1)	0.886	1.022	0.899	1.041	0.846	1.022	0.872	1.033
60	60	(0,1)	0.882	1.041	0.910	1.067	0.870	1.047	0.872	1.038
60	15	(5,5)	0.874	1.078	0.901	1.082	0.860	1.012	0.887	1.008
60	30	(5,5)	0.887	1.074	0.895	1.081	0.862	1.017	0.885	1.031
60	45	(5,5)	0.884	1.088	0.872	1.029	0.906	1.065	0.882	1.064
60	60	(5,5)	0.921	1.139	0.909	1.115	0.882	1.035	0.881	1.032
200	50	(0,1)	0.875	1.042	0.905	1.050	0.887	1.022	0.872	1.003
200	100	(0,1)	0.897	1.046	0.885	0.995	0.878	1.053	0.878	1.035
200	150	(0,1)	0.866	0.992	0.865	1.045	0.865	1.028	0.879	1.021
200	200	(0,1)	0.872	1.034	0.879	1.017	0.878	1.030	0.890	1.034
200	50	(5,5)	0.904	1.032	0.886	1.022	0.867	1.004	0.869	1.006
200	100	(5,5)	0.883	1.096	0.883	1.044	0.901	1.024	0.896	1.017
200	150	(5,5)	0.894	1.047	0.894	1.040	0.896	1.020	0.876	1.054
200	200	(5,5)	0.871	1.052	0.891	1.038	0.894	1.096	0.871	0.988

TABLE 4. SIMULATED QUANTILES FOR $Z_n^0(\lambda, \mathbf{R})$ FOR NORMALLY AND LAPLACE DISTRIBUTED ERRORS.

n	m	$\boldsymbol{\delta}$	$\psi(x) = \operatorname{sign} x$				Huber ψ			
			normal		Laplace		normal		Laplace	
			95%	99%	95%	99%	95%	99%	95%	99%
60	15	(0,1)	0.334	0.409	0.344	0.424	0.326	0.399	0.330	0.436
60	30	(0,1)	0.342	0.414	0.315	0.400	0.342	0.402	0.345	0.417
60	45	(0,1)	0.327	0.407	0.339	0.406	0.319	0.407	0.346	0.420
60	60	(0,1)	0.324	0.413	0.339	0.422	0.334	0.410	0.331	0.413
60	15	(5,5)	0.336	0.409	0.341	0.414	0.335	0.410	0.337	0.426
60	30	(5,5)	0.340	0.434	0.335	0.442	0.330	0.418	0.330	0.421
60	45	(5,5)	0.325	0.396	0.325	0.400	0.342	0.416	0.345	0.453
60	60	(5,5)	0.343	0.439	0.336	0.419	0.334	0.407	0.341	0.406
200	50	(0,1)	0.330	0.417	0.335	0.405	0.327	0.405	0.317	0.400
200	100	(0,1)	0.340	0.400	0.321	0.395	0.327	0.411	0.318	0.404
200	150	(0,1)	0.326	0.399	0.314	0.392	0.319	0.391	0.320	0.403
200	200	(0,1)	0.316	0.417	0.327	0.400	0.323	0.411	0.330	0.400
200	50	(5,5)	0.339	0.416	0.323	0.396	0.329	0.409	0.318	0.374
200	100	(5,5)	0.331	0.420	0.320	0.409	0.330	0.405	0.330	0.396
200	150	(5,5)	0.325	0.399	0.336	0.408	0.329	0.415	0.328	0.408
200	200	(5,5)	0.331	0.392	0.331	0.431	0.327	0.422	0.318	0.382

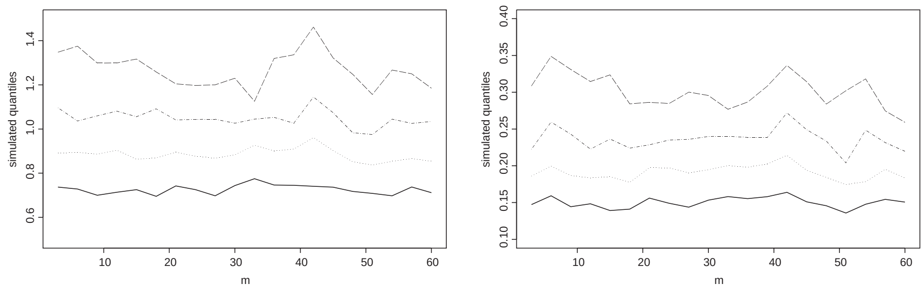

Figure 1. Simulated quantiles (90%-solid, 95%-dotted, 97.5%-dash-and-dot, 99%-dashed) for $T_n(\boldsymbol{R})$(left), $Z_n(\mathbf{R})$(right), $\boldsymbol{\delta}$= (0,1), $n = 60$, $\psi(x) = \text{sign}(x)$, the errors were simulated from the normal distributions.

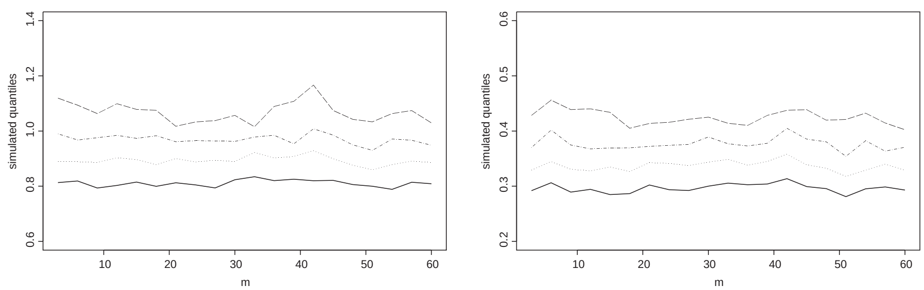

Figure 2. Simulated quantiles (90%-solid, 95%-dotted, 97.5%-dash-and-dot, 99%-dashed) for $T_n^0(\boldsymbol{R})$(left), $Z_n^0(\lambda, \mathbf{R})$(right), $\boldsymbol{\delta}$= (0,1), $n = 60$, $\psi(x) = \text{sign}(x)$, the errors were simulated from the normal distributions.

Figures 3-6 illustrate dependence of the sample 95% quantile on the change point m for $n = 60$ under various conditions. In particular, left parts of figures illustrate the influence of $\boldsymbol{\delta}$ and the distributions of errors for fixed ψ ($\psi(x) = \text{sign}\, x$). In these figures, the different lines represent different choices of $\boldsymbol{\delta}$ and error distributions, as follows.

(i) solid line: $\boldsymbol{\delta}$= (0,1), normally distributed errors,
(ii) dotted line: $\boldsymbol{\delta}$= (5,5), normally distributed errors,
(iii) dash-and-dot line: $\boldsymbol{\delta}$= (0,1), Laplace distributed errors,
(iv) dashed line: $\boldsymbol{\delta}$= (5,5), Laplace distributed errors.

Right parts of figures 3-6 illustrate the influence of choice of ψ and the distributions of errors for fixed $\boldsymbol{\delta}$ (= (0,1)). Here, the different lines represent different choices of $\psi(x)$ and error distributions, as follows.

(i) solid line: $\psi(x) = \text{sign}(x)$ and normally distributed errors,
(ii) dotted line: $\psi(x) = \text{sign}(x)$ and Laplace distributed errors,
(iii) dash-and-dot line: Huber ψ and normally distributed errors,
(iv) dashed line: Huber ψ and Laplace distributed errors.

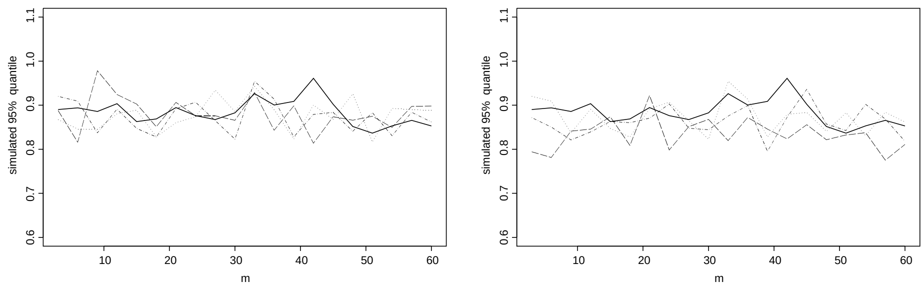

Figure 3. Simulated 95% quantile for $T_n(\mathbf{R})$ for sample size $n = 60$.

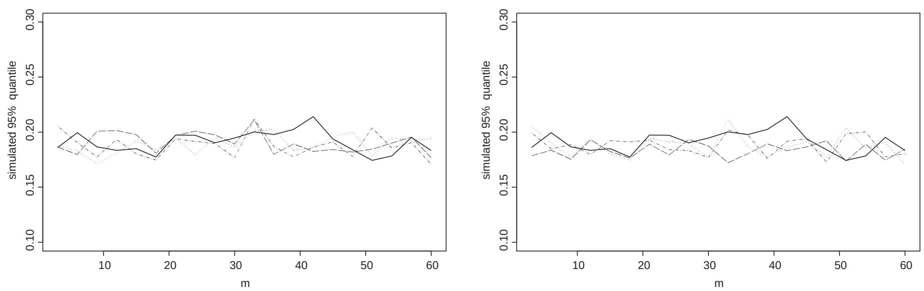

Figure 4. Simulated 95% quantile for $Z_n(\mathbf{R})$ for sample size $n = 60$.

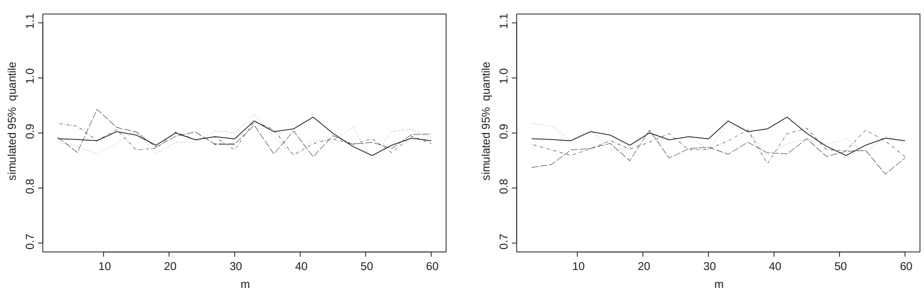

Figure 5. Simulated 95% quantile for $T_n^0(\mathbf{R})$ for sample size $n = 60$.

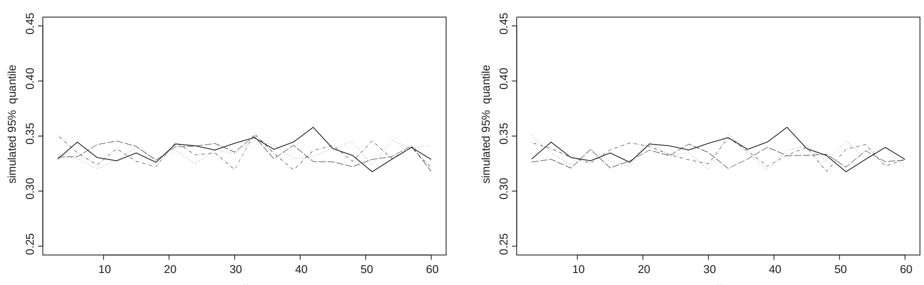

Figure 6. Simulated 95% quantile for $Z_n^0(\mathbf{R})$ for sample size $n = 60$.

TABLE 5. SIMULATED QUANTILES OF THE TEST STATISTICS AND THEIR PERMUTATION VERSIONS FOR NORMALLY AND LAPLACE DISTRIBUTED ERRORS UNDER THE NULL HYPOTHESIS.

Test statistic	n	$\psi(x) = \operatorname{sign} x$				Huber ψ			
		normal		Laplace		normal		Laplace	
		95%	99%	95%	99%	95%	99%	95%	99%
T_n	60	0.826	1.197	0.913	1.251	0.828	1.161	1.033	1.528
$T_n(\mathbf{R})$	60	0.855	1.251	0.906	1.329	0.863	1.189	0.856	1.218
T_n	200	0.857	1.185	0.850	1.188	0.869	1.202	1.045	1.499
$T_n(\mathbf{R})$	200	0.834	1.148	0.857	1.150	0.866	1.224	0.858	1.209
Z_n	60	0.185	0.277	0.193	0.270	0.191	0.294	0.249	0.341
$Z_n(\mathbf{R})$	60	0.185	0.272	0.200	0.299	0.184	0.278	0.186	0.287
Z_n	200	0.173	0.272	0.170	0.268	0.173	0.269	0.224	0.295
$Z_n(\mathbf{R})$	200	0.166	0.278	0.177	0.257	0.174	0.276	0.186	0.263
T_n^0	60	0.854	1.069	0.910	1.078	0.875	1.017	0.964	1.151
$T_n^0(\boldsymbol{R})$	60	0.882	1.041	0.910	1.067	0.870	1.047	0.872	1.038
T_n^0	200	0.864	1.029	0.879	1.035	0.887	1.029	0.973	1.140
$T_n^0(\boldsymbol{R})$	200	0.872	1.034	0.879	1.017	0.878	1.030	0.890	1.034
$Z_n^0(\lambda)$	60	0.326	0.399	0.334	0.413	0.336	0.427	0.382	0.464
$Z_n^0(\lambda, \mathbf{R})$	60	0.324	0.413	0.339	0.422	0.334	0.410	0.331	0.413
$Z_n^0(\lambda)$	200	0.322	0.408	0.318	0.401	0.320	0.394	0.363	0.425
$Z_n^0(\lambda, \mathbf{R})$	200	0.316	0.417	0.327	0.400	0.323	0.411	0.330	0.400

Sample quantiles have appeared quite stable and very similar.

Table 5 compares the simulated quantiles of the test statistics and their permutation versions for normally and Laplace distributed errors when data follow the null hypothesis. There is almost a perfect fit. These simulated quantiles were obtained by Monte Carlo procedure, i.e., the errors were simulated 1000 times and then for each case the test statistic was computed. The sample quantiles of these results provide sample critical values.

Finally, we study rejection rate for our testing problem in the model (3.1) with $n = 60$ (the sample size), the errors with normal $N(0,1)$, $\psi(x) = \operatorname{sign} x$, $\boldsymbol{\delta} = (5,5)$ and with various change points m. Notice that the null hypothesis corresponds to the last case, i.e. $m = n = 60$. We simulated such samples $Y_1, \ldots, Y_n$ 1000 times for the change point $m = 1, \ldots, 60$. Then the test statistics (3.2, 3.3) were computed and stored for each replication and for every m. The rejection rule was based on the simulated quantiles of the permutation version of the test statistics under null hypothesis, see Tables 1-4. The number of rejections was calculated for $m = 1, \ldots, n$. The results are plotted in Figures 7-8 together with the dash line corresponding to the number of 50 rejections of null hypothesis that corresponds to 5% level of the test.

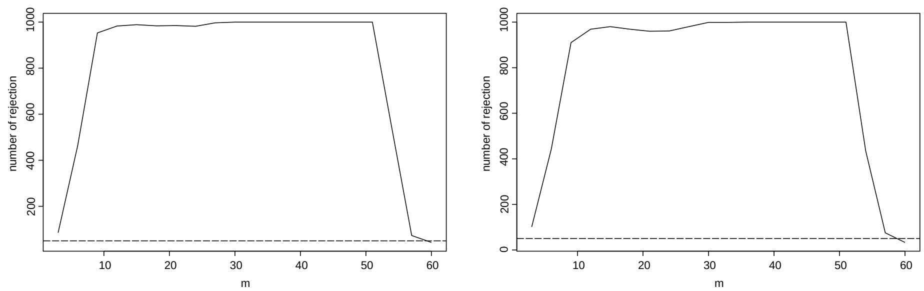

Figure 7. Simulated power of test based on T_n(left) and Z_n(right), 0.05-level.

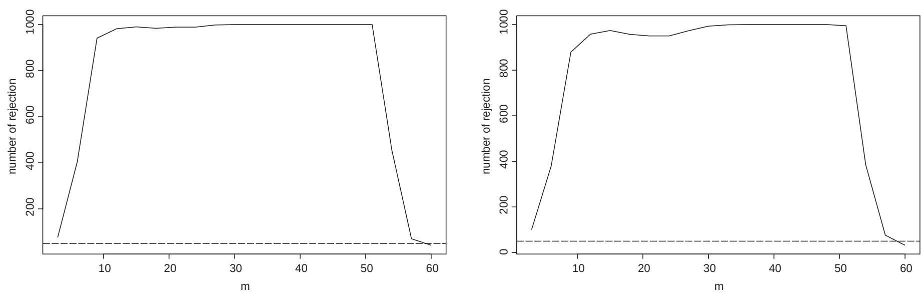

Figure 8. Simulated power of test based on T_n^0(left) and Z_n^0(right), 0.05-level.

Conclusion. The simulation study indicates:

(i) Sample quantiles are quite stable for various magnitudes of a change $\boldsymbol{\delta}$ and various locations of the change point m. The influence of both $\boldsymbol{\delta}$ and m is quite negligible even for larger $\boldsymbol{\delta}$.

(ii) There is no (or very small) influence of the location of the true change point m and size n.

(iii) Rejection rates are negligibly influenced by the location of the change.

We have made more extensive simulation experiments. In particular, we considered various functions $\boldsymbol{h}$, various score functions, various matrices $\boldsymbol{Q}_{k,n}$ and also other underlying distributions of the error terms. The corresponding sample quantiles have appeared to be either not influenced or very slightly influenced by varying choices of these parameters.

For the simulations we used S+ v. 6.0 running on 1800 MHz AMD Athlon XP with 256 MB memory. The time necessary for one simulation experiment was between 2-8 minutes.

4 Proofs

The proofs are along the lines of those in Antoch and Hušková (2003), Hušková and Picek (2002), and therefore we give only the main steps with emphasis on differences. In the present paper, the assumptions on the design matrix differ from those in the above papers. Also different test statistics are considered. The limit distributions are not of extreme value type any more and they are distributions of functionals of Gaussian processes that are almost surely finite. While in the papers Antoch and Hušková (2003) and Hušková and Picek (2002) higher order moments of $\psi(e_i)$ were supposed to be finite and the Komlós-Major-Tusnády arguments were employed, here we assume only finite second moment and the basic problem is to derive the limit distribution of the process

$$\boldsymbol{V}_n(t) = \frac{1}{\sqrt{n}} \sum_{i=1}^{\lfloor nt \rfloor} \boldsymbol{h}(i/n)\widehat{e}_i(\psi), \quad t \in [0,1],$$

and given $\boldsymbol{Y}$, the conditional limit distribution of the process

$$\boldsymbol{V}_n(t, \boldsymbol{R}) = \frac{1}{\sqrt{n}} \sum_{i=1}^{\lfloor nt \rfloor} \boldsymbol{h}(i/n)\widehat{e}_{R_i}(\psi), \quad t \in [0,1].$$

We focus on $T_n(\psi, \boldsymbol{Q}_n)$ and $T_n(\psi, \boldsymbol{Q}_n; \boldsymbol{R})$ as the proofs of the results on the other test statistics are quite similar. The latter are omitted.

Proof of Theorem 2.1. By results in Jurečková and Sen (1996) among others, as $n \to \infty$,

$$\boldsymbol{C}_n^{1/2}(\boldsymbol{\beta}_n(\psi) - \boldsymbol{\beta}) = \boldsymbol{C}_n^{-1} \frac{1}{\lambda'(0)} \sum_{i=1}^{n} \boldsymbol{h}(i/n)\psi(e_i) + o_P(n^{-1/2}). \tag{4.1}$$

Next, by the Hájek-Rényi-Chow inequality (Chow and Teicher, 1988) for each $A > 0$, $\gamma \in (0, 1/2]$ and $\boldsymbol{t} \in R^p$

$$P\Bigg(\max_{1\le k\le n/2} n^{-1/2+\gamma} k^{-\gamma} \Bigg| \sum_{i=1}^{k} h_j(i/n)\Big(\psi\big(e_i - \boldsymbol{h}^T(i/n)\boldsymbol{t}n^{-1/2}\big) - \psi(e_i) \quad (4.2)$$

$$+ \lambda \boldsymbol{h}^T(i/n)\boldsymbol{t}n^{-1/2}\Big)\Bigg| \ge A\Bigg)$$

$$\le D_1 A^{-2} n^{-1+2\gamma} \sum_{i=1}^{\lfloor n/2\rfloor} k^{-2\gamma} \int \big(\psi\big(e - \boldsymbol{h}^T(i/n)\boldsymbol{t}n^{-1/2}\big) - \psi(e)\big)^2 dF(e)$$

$$\le D_2 A^{-2}\big(||t||n^{-1/2}\big)^a, \quad j = 1, \ldots, p,$$

with some $D_1 > 0, D_2 > 0$, where a is a constant from the assumption (C.3), and similarly

$$P\Bigg(\max_{n/2\le k\le n} n^{-1/2+\gamma} (n-k)^{-\gamma} \Bigg| \sum_{i=k+1}^{n} h_j(i/n)\Big(\psi\big(e_i - \boldsymbol{h}^T(i/n)\boldsymbol{t}n^{-1/2}\big) \quad (4.3)$$

$$- \psi(e_i) + \lambda\Big(\boldsymbol{h}^T(i/n)\boldsymbol{t}n^{-1/2}\Big)\Big)\Bigg| \ge A\Bigg)$$

$$\le D_3 A^{-2}\big(||t||n^{-1/2}\big)^a, \quad j = 1, \ldots, p$$

with some $D_3 > 0$. Combining (4.1)-(4.3), we get, using standard tools that as $n \to \infty$,

$$\max_{1\le k\le n} \frac{1}{\sqrt{n}} \left(\frac{k(n-k)}{n^2}\right)^{-\gamma} \Bigg\| \sum_{i=1}^{k} \boldsymbol{h}(i/n)\widehat{e}_i(\psi)$$

$$- \left(\sum_{i=1}^{k} \boldsymbol{h}(i/n)\psi(e_i) - \boldsymbol{C}_k \boldsymbol{C}_n^{-1} \sum_{i=1}^{n} \boldsymbol{h}(i/n)\psi(e_i) \right) \Bigg\| = o_P(1).$$

Using again the same arguments, we also have as $n \to \infty$,

$$\widehat{\sigma}_n^2(\psi) = \frac{1}{2n} \sum_{i=2}^{n} \big(\psi(e_i) - \psi(e_{i-1})\big)^2 + o_P(1) = \int x^2 dF(x) + o_P(1). \quad (4.4)$$

Hence, the limit distribution of $T_n(\psi, \boldsymbol{Q}_n)$ is the same as that of

$$\max_{1\le k\le n}\left\{\frac{1}{\widehat{\sigma}^2(\psi)}\left(\sum_{i=1}^{k}\boldsymbol{h}(i/n)\psi(e_i)-\boldsymbol{C}_k\boldsymbol{C}_n^{-1}\sum_{i=1}^{n}\boldsymbol{h}(i/n)\psi(e_i)\right)^T\boldsymbol{Q}_{k,n}\right.$$
$$\left.\times\left(\sum_{i=1}^{k}\boldsymbol{h}(i/n)\psi(e_i)-\boldsymbol{C}_k\boldsymbol{C}_n^{-1}\sum_{i=1}^{n}\boldsymbol{h}(i/n)\psi(e_i)\right)\right\},$$

which in combination with Theorem 1 in Jandhyala and MacNeill (1997) implies assertion (2.3). The remaining assertions can be proved in the same way hence their proofs are omitted. □

PROOF OF THEOREM 2.2. After few standard steps using the assumption (2.8), we get as $n \to \infty$,

$$\boldsymbol{C}_n(\widehat{\boldsymbol{\beta}}_n(\psi)-\boldsymbol{\beta}) = (\boldsymbol{C}_n-\boldsymbol{C}_m)\boldsymbol{\delta}_n(1+o_P(1)), \tag{4.5}$$

which implies

$$\boldsymbol{S}_m = \lambda'(0)\boldsymbol{C}_m\boldsymbol{C}_n^{-1}(\boldsymbol{C}_n-\boldsymbol{C}_m)\boldsymbol{\delta}_n(1+o_P(1)),$$

and therefore

$$T_n(\psi,\boldsymbol{Q}_n)\ge(\lambda'(0))^2\boldsymbol{\delta}_n^T\boldsymbol{C}(\kappa)\boldsymbol{C}^{-1}(1)\boldsymbol{C}^0(t)\boldsymbol{Q}(\kappa)$$
$$\boldsymbol{C}^0(\kappa)\boldsymbol{C}^{-1}(1)\boldsymbol{C}(\kappa)\boldsymbol{\delta}_n(1+o_P(1)) \to^P \infty.$$

This proves (2.9). Again the assertion (2.10) can be obtained quite similarly, and therefore its proof is omitted. □

PROOF OF THEOREM 2.3. At first, notice that treating conditional distribution of $\boldsymbol{S}_k(\psi;\boldsymbol{R})$, $k = 1,\dots,n$, defined by (1.21), is equivalent to treating

$$\boldsymbol{S}_k^*(\psi,\boldsymbol{R}) = \sum_{i=1}^{k}\boldsymbol{h}(i/n)\widehat{e}_{(R_i)}(\psi)-\boldsymbol{C}_k\boldsymbol{C}_n^{-1}\sum_{j=1}^{n}\boldsymbol{h}(i/n)\widehat{e}_{(R_j)}(\psi),\ k-1,\dots,n, \tag{4.6}$$

where $\widehat{e}_{(1)}(\psi) \le \dots \le \widehat{e}_{(n)}(\psi)$ are order statistics corresponding to $\widehat{e}_1(\psi)$, $\dots,\widehat{e}_n(\psi)$. The proof relies on the fact that given $\boldsymbol{Y}$ the partial sums $\boldsymbol{S}_k^*(\psi,\boldsymbol{R})$, $k = 1,\dots,n$ can be viewed as vectors of linear rank statistics and the proofs reduce to treating functionals of linear rank statistics, where $\boldsymbol{R}$ is considered to be the ranks of random variables $U_1,\dots,U_n$ that are i.i.d. with uniform $(0,1)$ distribution.

We use the results on approximations of functionals of the rank statistics

$$\boldsymbol{V}_k = (V_{k1}, \dots, V_{kp})^T = \sum_{i=1}^{k} \boldsymbol{h}(i/n) a_n(R_i), \quad k = 1, \dots, n \tag{4.7}$$

by functionals of weighted sums of independent random variables

$$\boldsymbol{L}_k = (L_{k1}, \dots, L_{kp})^T = \sum_{i=1}^{k} \boldsymbol{h}(i/n)\big(a_n(\lfloor nU_i \rfloor + 1) - \overline{a}_n(\boldsymbol{U})\big), \quad k = 1, \dots, n, \tag{4.8}$$

where

$$\overline{a}_n(\boldsymbol{U}) = \frac{1}{n} \sum_{i=1} a_n(\lfloor nU_i \rfloor + 1). \tag{4.9}$$

The following lemma is the key tool in the proof of Theorem 2.3.

LEMMA 4.1 *Let scores $a_n(i)$'s be monotone and satisfy*

$$\sum_{i=1}^{n} a_n(i) = 0, \quad \frac{1}{n} \sum_{i=1}^{n} a_n^2(i) = 1. \tag{4.10}$$

Let the functions $\boldsymbol{h}$ satisfy (A.1) - (A.2). Then for any $\gamma < 1/2$, as $n \to \infty$,

$$\max_{1<k<n} \frac{1}{\sqrt{n}} \left(\frac{k(n-k)}{n^2} \right)^{-\gamma} \| \boldsymbol{V}_k - \boldsymbol{L}_k \| = O_P \left(\left(n^{-1} \max_{1 \le i \le n} a_n^2(i) \right)^{1/2} \right). \tag{4.11}$$

If moreover,

$$\lim_{n \to \infty} n^{-1} \max_{1 \le i \le n} a_n^2(i) = 0, \tag{4.12}$$

we have

$$\max_{1<k<n} \frac{1}{\sqrt{n}} \left(\frac{k(n-k)}{n^2} \right)^{-\gamma} \| \boldsymbol{V}_k - \boldsymbol{L}_k \| = o_P(1). \tag{4.13}$$

PROOF OF LEMMA 4.1. Since $\{\frac{1}{n-k}\boldsymbol{V}_k, B_{nk}; 1 \le k < n\}$ and $\{\frac{1}{k}(\boldsymbol{V}_n - \boldsymbol{V}_k), B^0_{nk}; n \ge k > 1\}$ are martingales, where B_{nk} and B^0_{nk} are σ-fields generated by $R_1, \dots, R_k$ and $R_n, \dots, R_{k+1}$, respectively, we have by the Hájek-Rényi-Chow inequality and by Lemma 2.1 in Hájek (1961) that for any $A > 0$ and any $\gamma < 1/2$:

$$P\left(\max_{1 \le k \le n/2} n^{-1/2+\gamma} k^{-\gamma} |V_{k1} - L_{k1}| \ge A \right) \ge D_5 A^{-2} \left(\max_{1 \le i \le n} a_n^2(i) \frac{1}{n^2} \sum_{i=1}^{n} a_n^2(i) \right)^{1/2}$$

and

$$P\Big(\max_{n/2\le k\le n} n^{-1/2+\gamma}(n-k)^{-\gamma}|V_{k1}-L_{k1}|\ge A\Big)$$
$$\ge D_6 A^{-2}\left(\max_{1\le i\le n} a_n^2(i)\frac{1}{n^2}\sum_{i=1}^n a_n^2(i)\right)^{1/2}$$

with some $D_5>0$ and $D_6>0$. This together with the assumption on function h_1 implies the assertion for first component of $\boldsymbol{V}_k$ and $\boldsymbol{L}_k$, i.e., as $n\to\infty$,

$$\max_{1\le k\le n} n^{-1/2}\left(\frac{k(n-k)}{n^2}\right)^{-\gamma}|V_{k1}-L_{k1}| = O_P\left(\left(\max_{1\le i\le n}|a_n(i)|n^{-1/2}\right)^{1/2}\right). \tag{4.14}$$

To prove the assertion for all components of $\boldsymbol{V}_k-\boldsymbol{L}_k$, we notice that by the Abel summation, for $k=1,\dots,n$,

$$\sum_{i=1}^k \boldsymbol{h}(i/n)\big(a_n(R_i)-(a_n(\lfloor nU_i\rfloor+1)-\overline{a}_n(\boldsymbol{U})\big)$$
$$=\boldsymbol{h}(k/n)\big(V_{k1}-L_{k1}\big)-\sum_{i=1}^{k-1}\big(\boldsymbol{h}((i+1)/n)-\boldsymbol{h}(i/n)\big)\big(V_{i1}-L_{i1}\big),$$

and then applying (4.14) we get the desired assertion after a few standard steps. □

CONTINUATION OF THE PROOF OF THEOREM 2.3. We apply Lemma 4.1 with

$$a_n(i)=\widehat{e}_{(i)}(\psi)\Big(\sum_{j=1}^n \widehat{e}_j^2(\psi)\Big)^{-1/2},\quad i=1,\dots,n. \tag{4.15}$$

By (1.5) and (1.20), we have

$$\sum_{i=1}^n \widehat{e}_i(\psi)=0.$$

Therefore it suffices to show that given $\boldsymbol{Y}$, our scores fulfill (4.12) in probability. In particular, we show that, as $n\to\infty$,

$$\frac{1}{n}\sum_{i=1}^n \widehat{e}_i^2(\psi)\to^P \int \psi^2(x)dF(x) \tag{4.16}$$

and

$$\frac{1}{n}\max_{1\le i\le n}\widehat{e}_i^2(\psi)\to^P 0. \tag{4.17}$$

The former assertion is an easy consequence of the assumptions on the score function ψ, weak law of large numbers and the fact that under the considered assumptions, as $n\to\infty$,

$$||\boldsymbol{\beta}_n(\psi)-\boldsymbol{\beta}||=o_P(1). \tag{4.18}$$

Since (4.18) holds, in order to prove (4.17), it suffices to show that for any sequence $\{\epsilon_n\}$ satisfying $\lim_{n\to\infty}\epsilon_n=0$,

$$n^{-1}\max_{1\le i\le n}\big(\psi^2(e_i+\epsilon_n)\big)\to^P 0.$$

However, this is implied by the Borel-Cantelli arguments and by the assumption (C.3). Hence, the assumptions of Lemma 4.1 are fulfilled in probability, and therefore given $\boldsymbol{Y}$, the conditional limiting behaviour of $T_n(\psi,\gamma;\boldsymbol{R})$ is the same as that of $T_n(\psi,\gamma;\boldsymbol{U})$, which is defined as $T_n(\psi,\gamma;\boldsymbol{R})$ with $\boldsymbol{S}_k(\psi;\boldsymbol{R})$ replaced by

$$\boldsymbol{S}_k(\psi,\boldsymbol{U})=\sum_{i=1}^k\boldsymbol{h}(i/n)(\widehat{e}_{\lfloor nU_i\rfloor+1}(\psi)-\widehat{e}_n(\psi,\boldsymbol{U}))$$

$$-\boldsymbol{C}_k\boldsymbol{C}_n^{-1}\sum_{i=1}^n\boldsymbol{h}(i/n)(\widehat{e}_{\lfloor nU_i\rfloor+1}(\psi)-\widehat{e}_n(\psi,\boldsymbol{U})),\quad k=1,\dots,n,$$

where

$$\widehat{e}(\psi,\boldsymbol{U})=\frac{1}{n}\sum_{i=1}^n\widehat{e}_{\lfloor nU_i\rfloor+1}(\psi).$$

Given $\boldsymbol{Y}$, the random vectors $\boldsymbol{M}_k(\psi)=\sum_{i=1}^k\boldsymbol{h}(i/n)\widehat{e}_{\lfloor nU_i\rfloor+1}(\psi)$, $k=1,\dots,$ n, are partial sums of independent random vectors with zero means and variance matrices

$$var\,\{\boldsymbol{M}_k(\psi)\}=\boldsymbol{C}_k\frac{1}{n}\sum_{i=1}^n\widehat{e}^2_{\lfloor nU_i\rfloor+1}(\psi),\quad k=1,\dots,n.$$

Then, taking into account (4.16) and proceeding similarly as in the proof of Theorem 2.1, we get the assertion of Theorem 2.3. □

Acknowledgement. The work of the first author was supported by grants GAČR 201/04/0945, PST.EAP.CLG980599 and MSM 0021620839. Research of the second author was supported by grants GAČR 201/05/2340 and MSM4674788501.

References

Andrews, D.W.K. (1993). Testing for structural instability and structural change with unknown change point. *Econometrica*, **61**, 821-856.

Antoch J., Hušková, M. and Jarušková, D. (2001). Off-line quality control. In *Multivariate Total Quality Control: Foundation and Recent Advances*, Lauro N. C. et al. eds., Springer-Verlag, Heidelberg, 1-86.

Antoch, J., Hušková, M. (2001). Permutation tests for change point analysis. *Statist. Probab. Lett.*, **53**, 37-,46.

Antoch, J., Hušková, M. (2003). Detection of structural changes in regression, *Tatra Mountains Publ.*, **26**, 201-215.

Bischoff, W. (1998). A functional central limit theorem for regression models. *Ann. Statist.*, **26**, 1398-1410.

Chow, Y.S and Teicher, H. (1988). *Probability Theory.* Springer Verlag, New York.

Csörgő, M. and Horváth, L. (1997). *Limit Theorems in Change-point Analysis.* J. Wiley, New York.

Hájek, J. (1961). Some extensions of the Wald-Wolfowitz-Noether theorem. *Ann. Math. Statist.*, **32** , 506-523.

Hušková, M. (1997). L_1-test procedures for detection of change. In *L_1-Statistical Procedures and Related Topics*, Y. Dodge ed., *IMS Lecture Notes-Monograph Series,* **31**, 56-70.

Hušková, M. (2000). Some invariant test procedures for detection of structural changes. *Kybernetika*, **36**, 401-414.

Hušková, M. (2004). Permutation principle in change point analysis. In *Asymptotic Methods in Stochastics*, Fields Institute Communications **44**, L. Horváth and B. Szyszkowicz, eds., 273-291.

Hušková, M. and Picek, J. (2002). M-tests for detection of structural changes in regression. In *Statistical Data Analysis Based on L_1-Norm and Related Methods.* ed. Y. Dodge, Birkhäuser, 213-228.

Hušková, M. and Picek, J. (2004). Some remarks on permutation type tests in linear models. *Discuss. Math. Probab. Stat.*, **24**, 151-182.

MacNeill, I.B. (1978). Properties of sequences of partial sums of polynomial regression residuals with applications to tests for change in regression at unknown times. *Ann. Statist.*, **6**, 422-433.

Jandhyala, V.K. and MacNeill, I.B. (1989). Residual partial sum limit processes for regression models with application to detecting parameter changes in unknown times. *Stochastic Process. Appl.*, **33**, 309-323.

Jandhyala V.K. and MacNeill, I.B. (1991). Tests for parameter changes at unknown times in linear regression models. *J. Statist. Plann. Inference*, **27**, 291-316.

Jandhyala V.K. and MacNeill, I.B. (1997). Iterated partial sum sequences of regression residuals and tests for change-points with continuity constraints. *J. Roy. Statist. Soc., Ser. B*, **59**, 147-156.

Jandhyala V.K. and Minogue C.D. (1993). Distributions of Bayes-type change-point statistics under polynomial regression. *J. Statist. Plann. Inference*, **37**, 271-290.

Jurečková, J. and Sen, P.K. (1996). *Robust Statistical Procedures: Asymptotics and Interrelations*. J. Wiley, New York.

Marie Hušková
Department of Statistics
Charles University of Prague
Sokolovská 83, CZ-186 75 Praha 8
and
ÚTIA Czech Academy of Sciences
Czech Republic
E-mail: marie.huskova@karlin.mff.cuni.cz

Jan Picek
Technical University of Liberec
Department of Applied Mathematics
Hálkova 6, CZ-461 17 Liberec
Czech Republic
E-mail: jan.picek@vslib.cz

Paper received: August 2004; revised March 2005.