Discussiones Mathematicae
Probability and Statistics 24 (2004) 151–181

SOME REMARKS ON PERMUTATION TYPE TESTS IN LINEAR MODELS

Marie Hušková

Charles University of Prague, Department of Statistics
Sokolovská 83, CZ – 186 75 Praha 8 and
UTIA Czech Academy of Sciences, Czech Republic

e-mail: marie.huskova@karlin.mff.cuni.cz

and

Jan Picek

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

Abstract

The paper discusses applications of permutation arguments in testing problems in linear models. Particular attention will be paid to the application in L_1-test procedures. Theoretical results will be accompanied by a simulation study.

Keywords: hypotheses testing, linear regression models, L_1- and L_2 - procedures.

2000 Mathematics Subject Classification: 62G20, 62E20, 60F17.

1. Introduction

We assume that the observations $Y_1, \dots, Y_n$ follow the linear model:

$$Y_i = \boldsymbol{x}_i^T \boldsymbol{\beta} + \boldsymbol{z}_i^T \boldsymbol{\gamma} + e_i, \quad i = 1, \dots, n, \tag{1}$$

where $\boldsymbol{x}_i = (x_{i1}, \dots, x_{ip})^T$, $i = 1, \dots, n$, and $\boldsymbol{z}_i = (z_{i1}, \dots, z_{iq})^T$, $i = 1, \dots, n$, are known regression vectors, $\boldsymbol{\beta} = (\beta_1, \dots, \beta_p)^T$ and $\boldsymbol{\gamma} = (\gamma_1, \dots, \gamma_q)^T$ are unknown parameters, the error terms $e_1, \dots, e_n$ are independent identically distributed random variables with some distributional properties specified below.

We are interested in the classical statistical testing problem on the linear hypothesis:

$$H_0 : \boldsymbol{\gamma} = \mathbf{0} \tag{2}$$

against the alternative that the model (1) with $\boldsymbol{\gamma} \neq \mathbf{0}$ holds true. The model (1) is called a full model, while the model (1) with $\boldsymbol{\gamma} = \mathbf{0}$ is called a reduced model. There are many books and an enormous amount of papers, where this problem is treated under various sets of assumptions on the distribution of the error term e_i's and the design points $\boldsymbol{x}_i, \boldsymbol{z}_i$, $i = 1, \dots, n$.

Assuming mild conditions on the distribution of the error terms e_i's we focus on the application of permutation arguments for the test statistics related to L_2- and L_1- procedures. Particularly, we consider the classical F-test statistic (L_2-type test statistic) and related L_1-test statistic.

We assume that the design points $\boldsymbol{x}_i, \boldsymbol{z}_i$, $i = 1, \dots, n$, satisfy:

A.1 $x_{i1} = 1, \quad i = 1, \dots, n, \ \sum_{i=1}^n x_{ij} = 0, \quad j = 2, \dots, p, \quad \sum_{i=1}^n z_{ij} = 0,$ $j = 1, \dots, q$, and $\sum_{i=1}^n \boldsymbol{z}_i \boldsymbol{x}_i^T = 0$.

A.2 There exist positive definite $p \times p$ matrices $\boldsymbol{C}_1$ and $\boldsymbol{C}_2$ such that, as $n \to \infty$,

$$\frac{1}{n} \sum_{i=1}^n \boldsymbol{x}_i \boldsymbol{x}_i^T \to \boldsymbol{C}_1,$$

$$\frac{1}{n} \sum_{i=1}^n \boldsymbol{z}_i \boldsymbol{z}_i^T \to \boldsymbol{C}_2$$

$$\max_{1\le i\le n} ||\boldsymbol{x}_i|| = O\left(n^{1/2-v}\right), \quad \max_{1\le i\le n} ||\boldsymbol{z}_i|| = O\left(n^{1/2-v}\right)$$

for some $v > 0$, where $||.||$ denotes the Euclidean norm.

Recall that the classical F-test statistic is equivalent (up to a multiplicative constant) to the statistic

$$T_n = \frac{1}{\widetilde{s}_n^2}\left(\sum_{i=1}^n \left(Y_i - \boldsymbol{x}_i^T\widehat{\boldsymbol{\beta}}_{n,2}\right)^2 - \sum_{i=1}^n \left(Y_i - \boldsymbol{x}_i^T\widetilde{\boldsymbol{\beta}}_{n,2} - \boldsymbol{z}_i^T\widetilde{\boldsymbol{\gamma}}_{n,2}\right)^2\right), \tag{3}$$

where $\widehat{\boldsymbol{\beta}}_{n,2}$ is the least square estimator of $\boldsymbol{\beta}$ in the reduced model and $\widetilde{\boldsymbol{\beta}}_{n,2}$ and $\widetilde{\boldsymbol{\gamma}}_{n,2}$ are the least square estimators of $\boldsymbol{\beta}$ and $\boldsymbol{\gamma}$, respectively, in the full model (1), i.e.,

$$\widehat{\boldsymbol{\beta}}_{n,2} = \widetilde{\boldsymbol{\beta}}_{n,2} = \left(\sum_{i=1}^n \boldsymbol{x}_i\boldsymbol{x}_i^T\right)^{-1} \sum_{i=1}^n \boldsymbol{x}_i Y_i, \tag{4}$$

$$\widetilde{\boldsymbol{\gamma}}_{n,2} = \left(\sum_{i=1}^n \boldsymbol{z}_i\boldsymbol{z}_i^T\right)^{-1} \sum_{i=1}^n \boldsymbol{z}_i Y_i. \tag{5}$$

Also,

$$\widetilde{s}_n^2 = \frac{1}{n-p-q}\sum_{i=1}^n \left(Y_i - \boldsymbol{x}_i^T\widetilde{\boldsymbol{\beta}}_{n,2} - \boldsymbol{z}_i^T\widetilde{\boldsymbol{\gamma}}_{n,2}\right)^2, \tag{6}$$

which is the classical estimator of $var\, Y_i = var\, e_i$. The test statistic T_n can be equivalently expressed as follows:

$$T_n = \frac{1}{\widetilde{s}_n^2}\left(\sum_{i=1}^n \boldsymbol{z}_i\widehat{e}_{i,2}\right)^T \left(\sum_{i=1}^n \boldsymbol{z}_i\boldsymbol{z}_i^T\right)^{-1} \left(\sum_{i=1}^n \boldsymbol{z}_i\widehat{e}_{i,2}\right), \tag{7}$$

where $\widehat{e}_{i,2}$, $i = 1, \ldots, n$, are the L_2-residuals in the reduced model, i.e.,

$$\widehat{e}_{i,2} = Y_i - \boldsymbol{x}_i^T \widehat{\boldsymbol{\beta}}_{n,2}, \quad i = 1, \ldots, n. \tag{8}$$

The so called L_1-analog of this test statistic T_n in the form (3) is defined as

$$T_{n,1} = \frac{1}{\widetilde{\sigma}_n} \left(\sum_{i=1}^n |Y_i - \boldsymbol{x}_i^T \widehat{\boldsymbol{\beta}}_{n,1}| - \sum_{i=1}^n |Y_i - \boldsymbol{x}_i^T \widetilde{\boldsymbol{\beta}}_{n,1} - \boldsymbol{z}_i^T \widetilde{\boldsymbol{\gamma}}_{n,1}| \right), \tag{9}$$

where $\widehat{\boldsymbol{\beta}}_{n,1}$ is the L_1-estimator of $\boldsymbol{\beta}$ in the reduced model (i.e., under H_0) and $\widetilde{\boldsymbol{\beta}}_{n,1}$ and $\widetilde{\boldsymbol{\gamma}}_{n,1}$ are L_1-estimators of $\boldsymbol{\beta}$ and $\boldsymbol{\gamma}$ in the full model (1), i.e., $\widehat{\boldsymbol{\beta}}_{n,1}$ is defined as a solution of the minimization problem

$$\min_{\boldsymbol{b}} \sum_{i=1}^n |Y_i - \boldsymbol{x}_i^T \boldsymbol{b}| \tag{10}$$

and $\tilde{\boldsymbol{\beta}}_{n,1}, \tilde{\boldsymbol{\gamma}}_{n,1}$ is defined as a solution of the minimization problem

$$\min_{\boldsymbol{b},\boldsymbol{d}} \sum_{i=1}^n |Y_i - \boldsymbol{x}_i^T \boldsymbol{b} - \boldsymbol{z}_i^T \boldsymbol{d}|. \tag{11}$$

Also

$$\widetilde{\sigma}_n = \frac{1}{n-p-q} \sum_{i=1}^n |Y_i - \boldsymbol{x}_i^T \widetilde{\boldsymbol{\beta}}_{n,1} - \boldsymbol{z}_i^T \widetilde{\boldsymbol{\gamma}}_{n,1}|. \tag{12}$$

For convenience we set

$$\widehat{\sigma}_n = \frac{1}{n-p} \sum_{i=1}^n |Y_i - \boldsymbol{x}_i^T \widehat{\boldsymbol{\beta}}_{n,1}|. \tag{13}$$

In its present form the statistic $T_{n,1}$ is an analog of the statistic T_n in the form (3). Sometimes $\widetilde{\sigma}_n$ in (9) is replaced either by $\widehat{\sigma}_n$ or by a consistent estimator of $2f(0)$. The last possibility leads to an asymptotically χ^2-distributed random variable under H_0, for discussion see, e.g. Cade and Richards(1996).

The L_1-analog of the test statistic F_n in the form (7) is defined as

$$T_{n,2} = \left(\sum_{i=1}^{n} \boldsymbol{z}_i \text{sign}\, \widehat{e}_{i,1}\right)^T \left(\sum_{i=1}^{n} \boldsymbol{z}_i \boldsymbol{z}_i^T\right)^{-1} \left(\sum_{i=1}^{n} \boldsymbol{z}_i \text{sign}\, \widehat{e}_{i,1}\right), \tag{14}$$

where $\widehat{e}_{i,1}$, $i = 1, \dots, n$, are the L_1-residuals in the reduced model, i.e.,

$$\widehat{e}_{i,1} = Y_i - \boldsymbol{x}_i^T \widehat{\boldsymbol{\beta}}_{n,1}, \quad i = 1, \dots, n. \tag{15}$$

The statistic $T_{n,1}$ is a loglikelihood type test statistic while $T_{n,2}$ is called the score (Rao) test statistic. The test statistics $T_{n,1}$ and $T_{n,2}$ are not identical, only asymptotically equivalent up to a multiplicative constant, for more details see, e.g. Koenker and Basset (1982).

Having the test statistics one needs a decision rule, more precisely one needs critical regions corresponding to the level α. Large values of the introduced test statistics indicate that the null hypothesis is violated. Therefore the problem reduces to finding the $100(1-\alpha)\%$-quantile of the distribution of test statistics under H_0. The corresponding critical regions are of the form:

$$T_n \geq t_n(1-\alpha), \tag{16}$$

$$T_{n,1} \geq t_{n,1}(1-\alpha), \tag{17}$$

$$T_{n,2} \geq t_{n,2}(1-\alpha), \tag{18}$$

where $t_n(1-\alpha), t_{n,1}(1-\alpha), t_{n,2}(1-\alpha)$ are the $100(1-\alpha)\%$-quantile of T_n, $T_{n,1}$, $T_{n,2}$, respectively, under H_0. An explicit form of these quantiles can be found exceptionally, e.g., when the error terms have $N(0, \sigma^2)$ distribution, otherwise one has to try to find an approximation. For a large number of observations approximations based on the asymptotic distribution of the test statistic can be used. It is known that if mild assumptions are fulfilled under H_0 all three test statistics have χ^2 distribution up to a multiplicative constants. Another possibility is to use resampling methods. We focus on the bootstrap without replacement, however the bootstrap with replacement can also be applied. The bootstrap without replacement can be viewed as an application of permutational arguments and we will formulate it in this way. Its advantage over the bootstrap with replacement is that it provides slightly better approximations to the critical values when H_0 holds true.

The application of resampling methods was proposed by De Angelis et al (1993) and by Edwards (1985) (bootstrap with replacement) for testing in some particular linear models. Some particular permutation type procedures were also considered by ter Braak (1992), Kennedy (1995), Freedman and Lane (1983) and Manly (1991). Cade and Richards (1996) present an extensive simulation study on the performance of permutation type test related to $T_{n,1}$. This paper also provides a list of related papers. Information on general aspects of permutation type tests can be found, e.g. in Lehmann (1991) and Good (2000). Permutation type tests for detection of changes in linear models were introduced and studied in Antoch and Hušková (2003) and Hušková and Picek (2002).

In the rest of the paper, we discuss the possibility of using the permutational principle to get approximations for the desired critical values. Theoretical results on the performance of the proposed test procedures are accompanied by a simulation study. In Section 2, the application of the permutational principle in a linear regression model is explained on the test statistic T_n. Section 3 is devoted to the procedures based on $T_{n,1}$ and $T_{n,2}$. The proofs of assertions from Section 3 are in Section 4. Results of the simulation study are presented in Section 5.

2. Permutation type procedures related to T_n

We start with an explanation of the ideas of permutation type procedures. Under the assumptions under consideration the vector of the random errors $(e_1, \ldots, e_n)$ has the same distribution as $(e_{R_1}, \ldots, e_{R_n})$, where $\boldsymbol{R} = (R_1, \ldots, R_n)$ is a random permutation of $(1, \ldots, n)$ independent of $e_1, \ldots, e_n$. According to the strict permutational arguments (see, e.g. Lehmann (1991) and Good (2000)) one should develop the permutation version of the test statistics under consideration replacing $(e_1, \ldots, e_n)$ by $(e_{R_1}, \ldots, e_{R_n})$ and considering only randomness induced by the random permutation $\boldsymbol{R}$. However, the error terms $e_1, \ldots, e_n$ are unknown, therefore we permute their "estimators"- the residuals. Notice also that under H_0:

$$T_n = \frac{1}{\widetilde{s}_n^2} \left(\sum_{i=1}^n \boldsymbol{z}_i e_i \right)^T \left(\sum_{i=1}^n \boldsymbol{z}_i \boldsymbol{z}_i^T \right)^{-1} \left(\sum_{i=1}^n \boldsymbol{z}_i e_i \right).$$

Then the permutation version $\widehat{T}_n(\boldsymbol{Y}, \boldsymbol{R})$ of T_n in (3) is defined as

$$\widehat{T}_n(\boldsymbol{Y}, \boldsymbol{R}) = \frac{1}{\widehat{s}_n^2} \left(\sum_{i=1}^n \boldsymbol{z}_i \widehat{e}_{R_i,2} \right)^T \left(\sum_{i=1}^n \boldsymbol{z}_i \boldsymbol{z}_i^T \right)^{-1} \left(\sum_{i=1}^n \boldsymbol{z}_i \widehat{e}_{R_i,2} \right), \tag{19}$$

where

$$\widehat{s}_n^2 = \frac{1}{n-p} \sum_{i=1}^n \widehat{e}_{i,2}^2. \tag{20}$$

The random permutation $\boldsymbol{R}$ has a uniform distribution over all permutations of $(1, \ldots, n)$, i.e., for any fixed permutation $\boldsymbol{r}$ of $1, \ldots, n$

$$P(\boldsymbol{R} = \boldsymbol{r}) = \frac{1}{n!}.$$

Therefore given the observations $\boldsymbol{Y} = (Y_1 \ldots, Y_n)$, the distribution of $\widehat{T}_n(\boldsymbol{Y}, \boldsymbol{R})$ is known and can be calculated. Hence also the $100(1-\alpha)\%$ conditional quantile $\widehat{t}_n(\boldsymbol{Y}, 1-\alpha)$ of $\widehat{T}_n(\boldsymbol{Y}, \boldsymbol{R})$ can be calculated. This conditional quantile $\widehat{t}_n(\boldsymbol{Y}, 1-\alpha)$ can be used as an approximation for the critical value $t_n(1-\alpha)$. The resulting test with the approximative level α has the critical region

$$T_n \geq \widehat{t}_n(\boldsymbol{Y}, 1-\alpha). \tag{21}$$

This means that the approximation for the quantiles is data based. The immediate question is whether this approximation is reasonable when the data follow either the full model or the reduced one. We discuss this issue. Notice

$$\sum_{i=1}^n \boldsymbol{z}_i \widehat{e}_{R_i,2} = \sum_{i=1}^n \boldsymbol{z}_i \left(e_{R_i} - \boldsymbol{x}_{R_i}^T \left(\widehat{\boldsymbol{\beta}}_{n,2} - \boldsymbol{\beta} \right) + \boldsymbol{z}_{R_i}^T \boldsymbol{\gamma} \right) = M_{n,1} + M_{n,2} + M_{n,3},$$

where

$$M_{n,1} = \sum_{i=1}^n \boldsymbol{z}_i e_{R_i} =^d \sum_{i=1}^n \boldsymbol{z}_i e_i,$$

$$M_{n,2} = \sum_{i=1}^n \boldsymbol{z}_i \boldsymbol{x}_{R_i}^T \left(\widehat{\boldsymbol{\beta}}_{n,2} - \boldsymbol{\beta} \right)$$

and

$$M_{n,3} = \sum_{i=1}^{n} \boldsymbol{z}_i \boldsymbol{z}_{R_i}^T \boldsymbol{\gamma}.$$

Here $=^d$ denotes equality in distribution. The terms $M_{n,2}$ and $M_{n,3}$ can be viewed as vectors of simple linear rank statistics. Direct calculations give

$$E\ (M_{n,2}|\boldsymbol{Y}) = \mathbf{0}, \quad E\ (M_{n,3}|\boldsymbol{Y}) = \mathbf{0},$$

$$var\ (M_{n,2}|\boldsymbol{Y}) = \sum_{i=1}^{n} \boldsymbol{z}_i \boldsymbol{z}_i^T \frac{1}{n-1} \sum_{i=1}^{n} \left((0, \boldsymbol{x}_i^{0T}) \left(\widehat{\boldsymbol{\beta}}_{n,2} - \boldsymbol{\beta} \right) \right)^2,$$

$$var\ (M_{n,3}|\boldsymbol{Y}) = \sum_{i=1}^{n} \boldsymbol{z}_i \boldsymbol{z}_i^T \frac{1}{n-1} \sum_{i=1}^{n} \left(\boldsymbol{z}_i^T \boldsymbol{\gamma} \right)^2,$$

$$cov\ (M_{n,2}, M_{n,3}|\boldsymbol{Y}) = \mathbf{0},$$

where $\boldsymbol{x}_i^0 = (x_{i2}, \ldots, x_{in})$, $i = 1, \ldots, n$. Hence for the unconditional distribution we have, as $n \to \infty$,

$$M_{n,2} = O_P(1), \qquad M_{n,3} = O_P(\sqrt{n}||\boldsymbol{\gamma}||).$$

From here one can infer that under H_0

$$T_n =^d \widehat{T}_n(\boldsymbol{Y}, \boldsymbol{R}) + O_P\left(n^{-1/2}\right)$$

which means that the permutational procedure mimics reasonably well the distribution under H_0 and n large while under alternatives we have

$$T_n =^d \widehat{T}_n\left(\boldsymbol{Y}, \boldsymbol{R}\right) + O_P\left(n^{-1/2}\right) + O_P\left(||\boldsymbol{\gamma}||\right)$$

and therefore the reasonable results are valid only for small γ. Alternatively, one can permute instead of the residuals $\widehat{e}_{i2}$'s residuals $\widetilde{e}_{i2}$'s corresponding to the model (1)

$$\widetilde{e}_{i,2} = Y_i - \boldsymbol{x}_i^T \widetilde{\boldsymbol{\beta}} - \boldsymbol{z}_i^T \widetilde{\boldsymbol{\gamma}}, \quad i = 1, \ldots, n,$$

that would work reasonably well even under alternatives. Then the corresponding permutation version $\widetilde{T}_n(\boldsymbol{Y}, \boldsymbol{R})$ of T_n is defined as $\widehat{T}_n(\boldsymbol{Y}, \boldsymbol{R})$ in (19) with $\widehat{e}_{R_i,2}$ replaced by $\widetilde{e}_{R_i,2}$. Clearly,

$$T_n =^d \widetilde{T}_n(\boldsymbol{Y}, \boldsymbol{R}) + O_P\left(n^{-1/2}\right)$$

both for the full model as well as the reduced one. Also, notice that, as $n \to \infty$,

$$\widetilde{s}_n^2 = \sigma^2 + O_P\left(n^{-1/2}\right)$$

and

$$\widehat{s}_n^2 = \sigma^2 + O_P\left(n^{-1/2}\right) + O_P\left(||\boldsymbol{\gamma}||^2\right),$$

where $\widehat{s}_n^2$ and $\widetilde{s}_n^2$ are defined by (20) and (6), respectively.

It is important to realize that the permutation based approximation for the critical values provides a reasonable approximation not only when the data follows H_0 but also when they follow the (full) model (1). It is a consequence of the following theorem.

Theorem A. *Let $Y_1, \ldots, Y_n$ follow model* (1)*, where the error terms $e_1, \ldots, e_n$ are i.i.d. random variables with zero mean and finite nonzero variance σ^2. Let the design points $\boldsymbol{x}_i, \boldsymbol{z}_i$, $i = 1, \ldots, n$, satisfy* (A.1)–(A.2).

(i) *Then under H_0 ($\boldsymbol{\gamma} = \boldsymbol{0}$) the test statistic T_n has asymptotically χ^2 distribution with q degrees of freedom.*

(ii) *Under H_0 and local alternatives ($\boldsymbol{\gamma} = \boldsymbol{\gamma}_n \to \boldsymbol{0}$) then, given $\boldsymbol{Y}$, the asymptotic distribution of the permutation test statistic $\widehat{T}_n(\boldsymbol{Y}, \boldsymbol{R})$ is χ^2 distribution with q degrees of freedom for almost all $\boldsymbol{Y}$.*

(iii) *Under H_0, local and fixed alternatives, given $\boldsymbol{Y}$, the asymptotic distribution of the permutation test statistic $\widetilde{T}_n(\boldsymbol{Y}, \boldsymbol{R})$ is χ^2 distribution with q degrees of freedom for almost all $\boldsymbol{Y}$.*

Proof. The assertion (i) is an easy consequence of the multivariate central limit theorem.

Concerning the assertion (ii) it suffices to realize that given $Y_1, \ldots, Y_n$ the statistics $\sum_{i=1}^n \boldsymbol{z}_i \widehat{e}_{R_i,2}$ can be viewed as a vector of simple linear rank statistics and their linear combinations satisfy the assumptions of Theorem 4.1 in Hájek (1961) and therefore given $Y_1, \ldots, Y_n$ the asymptotic distribution of $\sum_{i=1}^n \boldsymbol{z}_i \widehat{e}_{R_i,2}$ is asymptotically normal with zero mean and variance matrix

$$\sum_{i=1}^{n} \boldsymbol{z}_i \boldsymbol{z}_i^T \frac{1}{n-1} \sum_{i=1}^{n} \widehat{e}_{i,2}^2 = \sum_{i=1}^{n} \boldsymbol{z}_i \boldsymbol{z}_i^T \frac{n-p}{n-1} \widehat{s}_n^2.$$

The assertion (ii) is then an easy consequence. The proof (iii) is quite the same and therefore is omitted. ∎

Remark 2.1. By this theorem the limit distribution of T_n under H_0 and the conditional asymptotic distribution, given $\boldsymbol{Y}$, of $\widetilde{T}_n(\boldsymbol{Y}, \boldsymbol{R})$ coincide. In the case of $\widehat{T}_n(\boldsymbol{Y}, \boldsymbol{R})$ the assertion remains true if either the data follow H_0 or local alternatives.

Remark 2.2. It can be easily checked that the assertions (ii) and (iii) of Theorem A remain true if the permutation type procedure is replaced by bootstrapping the estimated residuals with replacement.

Remark 2.3. The parts (ii) and (iii) of Theorem A imply that under the respective conditions the conditional quantile $\widehat{t}_n(\boldsymbol{Y}, 1-\alpha)$ converges to the true one, i.e., as $n \to \infty$,

$$\widehat{t}_n(\boldsymbol{Y}, 1-\alpha) - t_n(1-\alpha) \to 0, \quad \text{a.s.},$$

which means that the permutation type procedure provides an asymptotically correct approximation for the critical value even if $\widehat{t}_n(\boldsymbol{Y}, 1-\alpha)$ corresponds to the data that do not follow the null hypothesis. This is very important for practical situations where we do not know whether the observed data follow the null hypothesis or the alternative one. The asymptotic power of the test with critical region (21) behavior is the same as that with the critical region $T_n \geq t_n(1-\alpha)$.

3. Permutation type procedures related to $T_{n,1}$ and $T_{n,2}$

The permutation versions of $T_{n,1}$ and $T_{n,2}$ are developed along the line for the statistic T_n. They are based on the L_1-residuals of the reduced model defined in (15) and the L_1-residuals of the full model defined as

$$\widetilde{e}_{i,1} = Y_i - \boldsymbol{x}_i^T \widetilde{\beta}_{n,1} - \boldsymbol{z}_i^T \widetilde{\gamma}_{n,1}, \quad i = 1, \ldots, n, \tag{22}$$

respectively. Notice that under H_0

$$\min_{\boldsymbol{b},\boldsymbol{d}} \sum_{i=1}^{n} |Y_i - \boldsymbol{x}_i^T \boldsymbol{b} - \boldsymbol{z}_i^T \boldsymbol{d}| = \min_{\boldsymbol{b},\boldsymbol{d}} \sum_{i=1}^{n} |e_i - \boldsymbol{x}_i^T \boldsymbol{b} - \boldsymbol{z}_i^T \boldsymbol{d}| \tag{23}$$

and

$$\min_{\boldsymbol{b}} \sum_{i=1}^{n} |Y_i - \boldsymbol{x}_i^T \boldsymbol{b}| = \min_{\boldsymbol{b}} \sum_{i=1}^{n} |e_i - \boldsymbol{x}_i^T \boldsymbol{b}|. \tag{24}$$

Then the permutation versions $\widehat{T}_{n,1}(\boldsymbol{Y},\boldsymbol{R})$ and $\widehat{T}_{n,2}(\boldsymbol{Y},\boldsymbol{R})$ based on the residuals $\widehat{e}_{i,1}$'s are defined by

$$\widehat{T}_{n,1}(\boldsymbol{Y},\boldsymbol{R}) = \frac{1}{\widetilde{\sigma}_n} \left(\min_{\boldsymbol{b}} \sum_{i=1}^{n} |\widehat{e}_{R_i,1} - \boldsymbol{x}_i^T \boldsymbol{b}| - \min_{\boldsymbol{b},\boldsymbol{d}} \sum_{i=1}^{n} |\widehat{e}_{R_i,1} - \boldsymbol{x}_i^T \boldsymbol{b} - \boldsymbol{z}_i^T \boldsymbol{d}| \right) \tag{25}$$

and

$$\widehat{T}_{n,2}(\boldsymbol{Y},\boldsymbol{R}) = \left(\sum_{i=1}^{n} \boldsymbol{z}_i \mathrm{sign}\, \widehat{e}_{R_i,1} \right)^T \left(\sum_{i=1}^{n} \boldsymbol{z}_i \boldsymbol{z}_i^T \right)^{-1} \left(\sum_{i=1}^{n} \boldsymbol{z}_i \mathrm{sign}\, \widehat{e}_{R_i,1} \right). \tag{26}$$

The permutation versions $\widetilde{T}_{n,1}(\boldsymbol{Y},\boldsymbol{R})$ and $\widetilde{T}_{n,2}(\boldsymbol{Y},\boldsymbol{R})$ based on the residuals $\widetilde{e}_{i,1}$'s are then defined by

$$\widetilde{T}_{n,1}(\boldsymbol{Y},\boldsymbol{R}) = \frac{1}{\widetilde{\sigma}_n} \left(\min_{\boldsymbol{b}} \sum_{i=1}^{n} |\widetilde{e}_{R_i,1} - \boldsymbol{x}_i^T \boldsymbol{b}| - \min_{\boldsymbol{b},\boldsymbol{d}} \sum_{i=1}^{n} |\widetilde{e}_{R_i,1} - \boldsymbol{x}_i^T \boldsymbol{b} - \boldsymbol{z}_i^T \boldsymbol{d}| \right) \tag{27}$$

and

$$\widetilde{T}_{n,2}(\boldsymbol{Y},\boldsymbol{R}) = \left(\sum_{i=1}^{n} \boldsymbol{z}_i \mathrm{sign}\, \widetilde{e}_{R_i,1} \right)^T \left(\sum_{i=1}^{n} \boldsymbol{z}_i \boldsymbol{z}_i^T \right)^{-1} \left(\sum_{i=1}^{n} \boldsymbol{z}_i \mathrm{sign}\, \widetilde{e}_{R_i,1} \right). \tag{28}$$

We do not apply the permutational principle to $\widetilde{\sigma}_n$ because

$$\widetilde{\sigma}_n = \frac{1}{n-p-q}\sum_{i=1}^{n}|e_i| + O_P(n^{-1/2})$$

which means that $\widetilde{\sigma}_n$ is only negligibly influenced by random permutations.

In the present section, we assume that the error terms satisfy:

(B.1) The error terms $e_1, \dots, e_n$ are i.i.d. random variables with zero median having a density f in a neighborhood of 0 with the property:

$$|f(x) - f(0)| \le c_1\sqrt{|x|} \quad \text{for} \quad |x| \le c_2$$

with some $c_1 > 0, c_2 > 0$.

Next we formulate main results on the test statistics $T_{n,1}, T_{n,2}$ and their permutation counterparts.

Theorem B. *Let $Y_1, \dots, Y_n$ follow model* (1)*, where the error terms $e_1, \dots, e_n$ are i.i.d. random variables with finite mean $E\,|e_1| < \infty$ and with common distribution function F satisfying* (B.1)*. Let the design points $\boldsymbol{x}_i, \boldsymbol{z}_i$, $i = 1, \dots, n$, satisfy* (A.1)–(A.2).

(i) *Then under H_0 ($\boldsymbol{\gamma} = \mathbf{0}$) the test statistic $T_{n,1}E|e_1|(2f(0))^{-1}$ has asymptotically χ^2 distribution with q degrees of freedom.*

(ii) *Under H_0 and local alternatives ($\boldsymbol{\gamma} = \boldsymbol{\gamma}_n$, $n^{1/2-v}\boldsymbol{\gamma}_n \to \mathbf{0}$, v is from the assumption* (A.2)*), given $\boldsymbol{Y}$, $\widehat{T}_{n,1}(\boldsymbol{Y}, \boldsymbol{R})E|e_1|(2f(0))^{-1}$ has asymptotically χ^2 distribution with q degrees of freedom for almost all $\boldsymbol{Y}$.*

(iii) *Under H_0 and local and fixed alternatives, given $\boldsymbol{Y}$, the asymptotic conditional distribution of $\widetilde{T}_{n,1}(\boldsymbol{Y}, \boldsymbol{R})E|e_1|(2f(0))^{-1}$ has asymptotically χ^2 distribution with q degrees of freedom for almost all $\boldsymbol{Y}$.*

(iv) *As $n \to \infty$,*

$$\tilde{\sigma}_n = E\,|e_1| + O_P(n^{-s}).$$

Theorem C. *Let $Y_1, \dots, Y_n$ follow model* (1)*, where the error terms $e_1, \dots, e_n$ are i.i.d. random variables common distribution function F satisfying* (B.1)*. Let the design points $\boldsymbol{x}_i, \boldsymbol{z}_i$, $i = 1, \dots, n$ satisfy* (A.1)–(A.2).

(i) *Then under* H_0 $(\boldsymbol{\gamma} = \mathbf{0})$ *the test statistic* $T_{n,2}$ *has asymptotically* χ^2 *distribution with* q *degrees of freedom.*

(ii) *Under* H_0 *and local alternatives* $(\boldsymbol{\gamma} = \boldsymbol{\gamma}_n,\ n^{1/2-v}\boldsymbol{\gamma}_n \to \mathbf{0}$, v *is from the assumption* (A.2)) *then, given* $\boldsymbol{Y}$, $\widehat{T}_{n,2}(\boldsymbol{Y},\boldsymbol{R})$ *has asymptotically* χ^2 *distribution with* q *degrees of freedom for almost all* $\boldsymbol{Y}$.

(iii) *Under* H_0 *and local and fixed alternatives, given* $\boldsymbol{Y}$, *the asymptotic conditional distribution of* $\widetilde{T}_{n,2}(\boldsymbol{Y},\boldsymbol{R})$ *has asymptotically* χ^2 *distribution with* q *degrees of freedom for almost all* $\boldsymbol{Y}$.

Proofs are posented in Section 4.

The assertions of both theorems say that permuting the residuals corresponding to the full model we always (both for the full and reduced models) get good approximations for the distribution of the test statistics under H_0. While permuting the residuals corresponding to the reduced model we get a reasonable approximation for the desired distribution only either under H_0 or under local alternatives.

The remarks formulated at the end of Section 2 remain true here also.

4. Proofs of Theorems B and C

The assertions (i) of Theorem B and (i) of Theorem C are well known results, see Jurečková and Sen (1996) and Babu (1989) among others. We concentrate on the proofs of the assertions (ii) of Theorem B and (ii) of Theorem C. The assertions (iii) of Theorem B and (iii) of Theorem C can be proved in the same way and therefore they are omitted.

First, we recall some known results for the estimators $\widehat{\boldsymbol{\beta}}_{n,1}$, $\widetilde{\boldsymbol{\beta}}_{n,1}$ and $\widetilde{\boldsymbol{\gamma}}_{n,1}$ and test statistics $T_{n,1}$ and $T_{n,2}$. Under the assumptions (A.1)-(A.2) and (B.1) in the reduced model, as $n \to \infty$,

$$(29)\quad \widehat{\boldsymbol{\beta}}_{n,1} - \boldsymbol{\beta} = \frac{1}{2f(0)}\left(\sum_{i=1}^{n} \boldsymbol{x}_i\boldsymbol{x}_i^T\right)^{-1}\sum_{i=1}^{n}\boldsymbol{x}_i \operatorname{sign} e_i + O(n^{-1/2-s}), \quad \text{a.s.},$$

with some $s > 0$, while in the full model we have , as $n \to \infty$,

$$\widetilde{\boldsymbol{\beta}}_{n,1} - \boldsymbol{\beta} = \frac{1}{2f(0)} \left(\sum_{i=1}^{n} \boldsymbol{x}_i \boldsymbol{x}_i^T \right)^{-1} \sum_{i=1}^{n} \boldsymbol{x}_i \operatorname{sign} e_i + O(n^{-1/2-s}), \quad \text{a.s.}, \tag{30}$$

$$\widetilde{\gamma}_{n,1} - \boldsymbol{\gamma} = \frac{1}{2f(0)} \left(\sum_{i=1}^{n} \boldsymbol{z}_i \boldsymbol{z}_i^T \right)^{-1} \sum_{i=1}^{n} \boldsymbol{z}_i \operatorname{sign} e_i + O(n^{-1/2-s}), \quad \text{a.s.}, \tag{31}$$

with some $s > 0$ and in the full model with $||\boldsymbol{\gamma}||n^{1/2-v} \to 0$

$$\sqrt{n}(\widehat{\boldsymbol{\beta}}_{n,1} - \boldsymbol{\beta}) = O(\log n), \quad \text{a.s.} \tag{32}$$

To prove the assertions on the asymptotic conditional distributions of $T_{n,1}(\boldsymbol{R})$ and $T_{n,2}(\boldsymbol{R})$ one follows the main step of the proofs for $T_{n,1}$ and $T_{n,2}$ but here we have the linear rank statistics instead of sums of independent random variables. So some steps need particular attention. The proofs are highly technical and we present here only main steps.

The fundamental issue is to investigate distributional properties of $\widehat{\boldsymbol{\beta}}_{n,1}(\boldsymbol{R})$, defined as a solution of the minimization problem

$$\min_{\boldsymbol{b}} \sum_{i=1}^{n} \Big| \widehat{e}_{R_i,1} - \boldsymbol{x}_i^T \boldsymbol{b} \Big|, \tag{33}$$

and also distributional properties of $\sum_{i=1}^{n} \Big| \widehat{e}_{R_i,1} - \boldsymbol{x}_i^T \widehat{\boldsymbol{\beta}}_{n,1}(\boldsymbol{R}) \Big|$ for a given $\boldsymbol{Y}$.

We will repeatedly use the relation

$$|v - a| - |v| + a \operatorname{sign} v = 2|a - v|(I\{v;\, a < v < 0\} + I\{v;\, 0 < v < a\}), \tag{34}$$

$v \in R^1$ and $a \in R^1$, where $I\{A\}$ denotes the indicator of the set A.

Denoting for each $\boldsymbol{b} \in R^p$

$$\widehat{S}_n(\boldsymbol{b}, \boldsymbol{R}) = \sum_{i=1}^{n} \left| \widehat{e}_{R_i,1} - \boldsymbol{x}_i^T \boldsymbol{b} n^{-1/2} \right| - \sum_{i=1}^{n} \left| \widehat{e}_{R_i,1} \right| + n^{-1/2} \sum_{i=1}^{n} \boldsymbol{x}_i^T \boldsymbol{b} n^{-1/2} \operatorname{sign} \widehat{e}_{R_i,1} \tag{35}$$

we find that by (34)

$$\widehat{S}_n(\boldsymbol{b}, \boldsymbol{R}) = \sum_{i=1}^{n} (a(R_i, i; \boldsymbol{b}) - \overline{a}(.,.; \boldsymbol{b})) - n\overline{a}(.,.; \boldsymbol{b})), \tag{36}$$

where

$$a(i, j; \boldsymbol{b}) = \left| \widehat{e}_{i,1} - \boldsymbol{x}_j^T \boldsymbol{b} n^{-1/2} \right| \tag{37}$$

$$\times \left(I \left\{ \boldsymbol{x}_j^T \boldsymbol{b} n^{-1/2} < \widehat{e}_{i,1} < 0 \right\} + I \left\{ 0 < \widehat{e}_{i,1} < \boldsymbol{x}_j^T \boldsymbol{b} n^{-1/2} \right\} \right), \quad i, j = 1, \ldots, n,$$

and

$$\overline{a}(.,.; \boldsymbol{b}) = \frac{1}{n} \sum_{i=1}^{n} \sum_{j=1}^{n} a(i, j; \boldsymbol{b}). \tag{38}$$

Notice that for $\boldsymbol{b} \in R^p$

$$|a(i, j; \boldsymbol{b})| \le |\boldsymbol{x}_j^T \boldsymbol{b} n^{-1/2}| \le D||\boldsymbol{b}||n^{-v}, \; i, j = 1, \ldots, n, \tag{39}$$

with some $D > 0$ and with v from the assumption (B.2). Given $\boldsymbol{Y}$, $\widehat{S}_n(\boldsymbol{b}, \boldsymbol{R})$, $\boldsymbol{b} \in R^p$ can be viewed as rank statistics and therefore we try to employ various results on rank statistics. We formulate them in the following lemma.

Lemma a. *Under the assumptions* (A.1) *and* (A.2) *we have*

$$E\left(\widehat{S}_n(\boldsymbol{b},\boldsymbol{R})|\boldsymbol{Y}\right) = \overline{a}(.,.;\boldsymbol{b}), \tag{40}$$

$$var\{\widehat{S}_n(\boldsymbol{b},\boldsymbol{R})|\boldsymbol{Y}\} = O\left(\frac{1}{n}\sum_{i=1}^{n}\sum_{j=1}^{n} a^2(i,j;\boldsymbol{b})\right), \quad \text{a.s.} \tag{41}$$

and

$$E\Big\{\left(\widehat{S}_n(\boldsymbol{b},\boldsymbol{R}) - E(\widehat{S}_n(\boldsymbol{b},\boldsymbol{R})|\boldsymbol{Y})\right)^{2K}|\boldsymbol{Y}\Big\} \tag{42}$$

$$= O\left(\sum_{L=1}^{2K}\left(||\boldsymbol{b}||n^{-v}\right)^{2K-2\sum_{s=2}^{L}k_s}\left(\frac{1}{n}\sum_{i=1}^{n}\sum_{j=1}^{n} a^2(i,j;\boldsymbol{b})\right)^{\sum_{s=2}^{L}k_s}\right), \quad \text{a.s.}$$

uniformly in $\boldsymbol{b} \in R^p$*, where* $K -$ *an arbitrary fixed positive integer.*

Proof. The assortions (40) and (41) are straightforward. As concerns (42), notice that for any positive integer $s < 2K$

$$\left|E\left(\sum_{i_1=1, i_1\neq i_2,\dots,i_s}^{n}(a(R_{i_1}, i_1, \boldsymbol{b}) - \overline{a}(.,.;\boldsymbol{b}))|\boldsymbol{Y}\right)\right|$$
$$= O(||\boldsymbol{b}||n^{-v}), \quad \text{a.s.}$$

Then using this relation and (39) we obtain after tedious but standard calculations

$$E\Big\{ \Big(\widehat{S}_n(\boldsymbol{b},\boldsymbol{R}) - E(\widehat{S}_n(\boldsymbol{b},\boldsymbol{R})|\boldsymbol{Y})\Big)^{2K} |\boldsymbol{Y}\Big\}$$

$$= O\left(\sum_{L=1}^{2K}\sum_{k_1,\dots,k_L}\sum_{i_{k_1+1},\dots,i_{\sum_{v=1}^{L}k_v}}\right.$$

$$\left. E\left(\prod_{s=1}^{L}\prod_{j=k_1+\dots+k_{s-1}+1}^{k_1+\dots+k_s}\left(a\left(R_{i_j},i_j;\boldsymbol{b}\right)-\overline{a}_n(.,.;\boldsymbol{b})\right)^s\right)^{k_s}|\boldsymbol{Y}\right)$$

$$= O\left(\sum_{L=1}^{2K}\left(||\boldsymbol{b}||n^{-v}\right)^{k_1+\sum_{s=2}^{L}(s-2)k_s}\sum_{k_2,\dots,k_L}\sum_{i_1,\dots,i_{\sum_{v=1}^{L}k_v}}\right.$$

$$\left. E\left(\prod_{s=2}^{L}\prod_{j=k_1+\dots+k_{s-1}+1}^{k_1+\dots+k_s}\left(a(R_{i_j},i_j;\boldsymbol{b})-\overline{a}_n(.,.;\boldsymbol{b})\right)^2\right)^{k_s}|\boldsymbol{Y}\right)$$

$$= O\left(\sum_{L=1}^{2K}\left(||\boldsymbol{b}||n^{-v}\right)^{2K-2\sum_{s=2}^{L}k_s}\left(\frac{1}{n}\sum_{i=1}^{n}\sum_{j=1}^{n}a^2(i,j;\boldsymbol{b})\right)^{\sum_{s=2}^{L}k_s}\boldsymbol{Y}\right),$$

where K is an arbitrary positive integer, $\sum_{k_1,\dots,k_L}$ extends over all nonnegative integers $k_1,\dots,k_L$ such that $\sum_{s=1}^{L} sk_s = 2K$ and $\sum_{i_1,\dots,i_{k_L}}$ extend over all $i_1,\dots,i_{k_L}$, such that $1 \le i_j \le n$ and $i_1,\dots,i_{k_L}$ are all different. ■

Clearly, we need to investigate

$$(43)\qquad \frac{1}{n}\sum_{i=1}^{n}\sum_{j=1}^{n}|a(i,j;\boldsymbol{b})|^2 \quad\text{and}\quad \frac{1}{n}\sum_{i=1}^{n}\sum_{j=1}^{n}a(i,j;\boldsymbol{b}),\quad \boldsymbol{b}\in R^p$$

that depends on $\widehat{e}_{1,1},\dots,\widehat{e}_{n,1}$. To this end it is useful to investigate

$$W_{n,s}(\boldsymbol{t}_1, \boldsymbol{t}_2, \boldsymbol{t}_3; \boldsymbol{Y}) = \sum_{i=1}^{n} Z_{in,s}(\boldsymbol{t}_1, \boldsymbol{t}_2, \boldsymbol{t}_3), \tag{44}$$
$$\boldsymbol{t}_j \in R^p,\ j = 1, 3,\ \boldsymbol{t}_2 \in R^q,\ s = 1, 2,$$

where

$$Z_{in,s}(\boldsymbol{t}_1, \boldsymbol{t}_2, \boldsymbol{t}_3) = \frac{1}{n} \sum_{j=1}^{n} \left| e_i - \boldsymbol{x}_i^T \boldsymbol{t}_1 n^{-1/2} - \boldsymbol{z}_i^T \boldsymbol{t}_2 n^{-1/2} - \boldsymbol{x}_j^T \boldsymbol{t}_3 n^{-1/2} \right|^s$$
$$\times \left(I \left\{ 0 < e_i - \boldsymbol{x}_i^T \boldsymbol{t}_1 n^{-1/2} - \boldsymbol{z}_i^T \boldsymbol{t}_2 n^{-1/2} < \boldsymbol{x}_j^T \boldsymbol{t}_3 n^{-1/2} \right\} \right.$$
$$\left. + I \left\{ \boldsymbol{x}_j^T \boldsymbol{t}_3 n^{-1/2} < e_i - \boldsymbol{x}_i^T \boldsymbol{t}_1 n^{-1/2} - \boldsymbol{z}_i^T \boldsymbol{t}_2 n^{-1/2} < 0 \right\} \right),\ s = 1, 2.$$

Lemma b. *Under the assumptions of Theorem* B(ii), *for any* $\beta > 0$ *there exists* $\lambda > 0$ *such that, as* $n \to \infty$,

$$\sup_{||\boldsymbol{t}_i|| \le D n^{v-\lambda}, i=1,2; ||\boldsymbol{t}_3|| \le D n^{\beta} n^{v/3-\lambda}} \left| W_{n,1}(\boldsymbol{t}_1, \boldsymbol{t}_2, \boldsymbol{t}_3; \boldsymbol{Y}) - \frac{f(0)}{2} \frac{1}{n} \sum_{i=1}^{n} \left(\boldsymbol{t}_3^T \boldsymbol{x}_i \right)^2 \right| \to 0, \tag{45}$$

a.s. *and*

$$\sup_{||\boldsymbol{t}_i|| \le D n^{v-\lambda}, i=1,2; ||\boldsymbol{t}_3|| \le D n^{\beta} n^{v/3-\lambda}} \left| W_{n,2}(\boldsymbol{t}_1, \boldsymbol{t}_2, \boldsymbol{t}_3; \boldsymbol{Y}) \right| \to 0, \quad \text{a.s.} \tag{46}$$

Proof. Clearly,

$$|Z_{in}(\boldsymbol{t}_1, \boldsymbol{t}_2, \boldsymbol{t}_3)| \le C ||\boldsymbol{t}_3|| n^{-v} \tag{47}$$

for some $C > 0$. Standard straightforward calculations give

$$
(48) \quad \begin{aligned} E\, W_{n,1}(\boldsymbol{t}_1, \boldsymbol{t}_2, \boldsymbol{t}_3; \boldsymbol{Y}) &= \frac{f(0)}{2} \frac{1}{n} \sum_{i=1}^{n} (\boldsymbol{x}_i^T \boldsymbol{t}_3)^2 \\ &\quad + O\left(n^{-v/2} ||\boldsymbol{t}_3||^2 (||\boldsymbol{t}_1|| + ||\boldsymbol{t}_2|| + ||\boldsymbol{t}_3||)^{1/2}\right), \end{aligned}
$$

$$
var\Big\{W_{n,1}(\boldsymbol{t}_1, \boldsymbol{t}_2, \boldsymbol{t}_3; \boldsymbol{Y})\Big\} = O\left(||\boldsymbol{t}_3||^3 n^{-v}\right)
$$

for $||\boldsymbol{t}_i|| \le Dn^{\beta}, \beta < v/3$. Since $W_{n,1}(\boldsymbol{t}_1, \boldsymbol{t}_2, \boldsymbol{t}_3; \boldsymbol{Y})$ is the sum of independent random variables we get for the $2K$th central moment:

$$
\begin{aligned}
&E\ (W_{n,1}(\boldsymbol{t}_1, \boldsymbol{t}_2, \boldsymbol{t}_3; \boldsymbol{Y}) - E\, W_{n,1}(\boldsymbol{t}_1, \boldsymbol{t}_2, \boldsymbol{t}_3; \boldsymbol{Y}))^{2K} \\
&= O\left(\sum_{k_2, \ldots, k_{2K}} \prod_{s=2}^{2K} \left(\sum_{i=1}^{n} E|Z_{in,1}(\boldsymbol{t}_1, \boldsymbol{t}_2, \boldsymbol{t}_3)|^s \right)^{k_s} \right) \\
&= O\left(\sum_{k_2, \ldots, k_{2K}} \left(\sum_{i=1}^{n} E|Z_{in,1}(\boldsymbol{t}_1, \boldsymbol{t}_2, \boldsymbol{t}_3)|^2 \right)^{\sum_{s=2}^{2K} k_s} \left(||\boldsymbol{t}_3|| n^{-v}\right)^{\sum_{s=2}^{2K} (s-2) k_s} \right) \\
&= O(||\boldsymbol{t}_3||^{3K} n^{-Kv})
\end{aligned}
$$

uniformly for $\boldsymbol{t}_1, \boldsymbol{t}_2, \boldsymbol{t}_3$ considered in the supremum in the assertion.

Since K can be chosen arbitrary large this relation together with (48) imply (45). The assertion (46) can be obtained quite analogously. Particularly, we get that

$$
E\, (W_{n,2}(\boldsymbol{t}_1, \boldsymbol{t}_2, \boldsymbol{t}_3; \boldsymbol{Y}) | \boldsymbol{Y}) = O(n^{-v} ||\boldsymbol{t}_3||^3)
$$

and

$$
\begin{aligned}
&E\left((W_{n,2}(\boldsymbol{t}_1, \boldsymbol{t}_2, \boldsymbol{t}_3; \boldsymbol{Y}) - E(W_{n,2}(\boldsymbol{t}_1, \boldsymbol{t}_2, \boldsymbol{t}_3; \boldsymbol{Y}) | \boldsymbol{Y}))^{2K} \, | \boldsymbol{Y} \right) \\
&= O(||\boldsymbol{t}_3||^{3K} n^{-Kv}).
\end{aligned}
$$

■

Since (29) and (32) and applying the Lemma b we can get

$$\overline{a}(.,.;\boldsymbol{b}) = f(0)\frac{1}{n}\sum_{i=1}^{n}(\boldsymbol{b}^T\boldsymbol{x}_i)^2 + O(n^{-s}), \quad \text{a.s.} \tag{49}$$

uniformly for $||\boldsymbol{b}|| \le n^{-w}$) with some $s > 0, w > 0$ and

$$\sup_{||\boldsymbol{b}||\le Dn^{-v/3-\beta}} \frac{1}{n}\sum_{i=1}^{n}\sum_{j=1}^{n} a^2(i,j;\boldsymbol{b}) = O(n^{-s}), \quad \text{a.s.} \tag{50}$$

with some $s > 0, \beta > 0$. Hence

$$\sup_{||\boldsymbol{b}||\le Dn^{-v/3-\beta}} E\left(\left(\widehat{S}_n(\boldsymbol{b},\boldsymbol{R}) - E(\widehat{S}_n(\boldsymbol{b},\boldsymbol{R})|\boldsymbol{Y})\right)^{2K}\Big|\boldsymbol{Y}\right) = O\left((||\boldsymbol{b}||n^{-s})^K\right).$$

Then after a few standard steps we get receive that in the full model (1) with $n^{\beta}||\boldsymbol{\gamma}|| \to 0$, for some $\beta > 0$

$$\widehat{\boldsymbol{\beta}}_{n,1}(\boldsymbol{R}) = \frac{1}{2f(0)}\left(\sum_{i=1}^{n}\boldsymbol{x}_i\boldsymbol{x}_i^T\right)^{-1}\sum_{i=1}^{n}\boldsymbol{x}_i\text{sign}\,\widehat{e}_{R_i} + O(n^{-1/2+v}), \quad \text{a.s.} \tag{51}$$

and for any $\boldsymbol{b} \in R^p$,

$$\sum_{i=1}^{n}\left|\widehat{e}_{R_i,1} - \boldsymbol{x}_i^T\boldsymbol{b}n^{-1/2}\right| - \sum_{i=1}^{n}\left|\widehat{e}_{R_i,1}\right|$$

$$= -n^{-1/2}\sum_{i=1}^{n}\boldsymbol{x}_i^T\boldsymbol{b}\text{sign}\,\widehat{e}_{R_i,1} + \frac{f(0)}{2}\frac{1}{n}\sum_{i=1}^{n}(\boldsymbol{x}_i^T\boldsymbol{b})^2 + O(n^{-s}), \quad \text{a.s.}$$

for some $s > 0$. Hence standard arguments give

$$\min_{\boldsymbol{b}}\sum_{i=1}^{n}\left|\widehat{e}_{R_i,1} - \boldsymbol{x}_i^T\boldsymbol{b}n^{-1/2}\right| - \sum_{i=1}^{n}\left|\widehat{e}_{R_i,1}\right|$$

$$= -\frac{1}{2f(0)}\left(\sum_{i=1}^{n}\boldsymbol{x}_i\text{sign}\ \widehat{e}_{R_i,1}\right)^T\left(\sum_{i=1}^{n}\boldsymbol{x}_i\boldsymbol{x}_i^T\right)^{-1}\left(\sum_{i=1}^{n}\boldsymbol{x}_i\text{sign}\ \widehat{e}_{R_i,1}\right)$$

$$+O(n^{-s}), \quad \text{a.s.}$$

for some $s > 0$.

Quite analogously we obtain

$$\min_{\boldsymbol{b},\boldsymbol{d}} \sum_{i=1}^{n} \left| \widehat{e}_{R_i,1} - \boldsymbol{x}_i^T \boldsymbol{b} n^{-1/2} - \boldsymbol{z}_i^T \boldsymbol{d} n^{-1/2} \right| - \sum_{i=1}^{n} \left| \widehat{e}_{R_i,1} \right|$$

$$= -\frac{1}{2f(0)} \left\{ \left(\sum_{i=1}^{n} \boldsymbol{x}_i \operatorname{sign} \widehat{e}_{R_i,1} \right)^T \left(\sum_{i=1}^{n} \boldsymbol{x}_i \boldsymbol{x}_i^T \right)^{-1} \left(\sum_{i=1}^{n} \boldsymbol{x}_i \operatorname{sign} \widehat{e}_{R_i,1} \right) \right.$$

$$\left. + \left(\sum_{i=1}^{n} \boldsymbol{z}_i \operatorname{sign} \widehat{e}_{R_i,1} \right)^T \left(\sum_{i=1}^{n} \boldsymbol{z}_i \boldsymbol{z}_i^T \right)^{-1} \left(\sum_{i=1}^{n} \boldsymbol{x}_i \operatorname{sign} \widehat{e}_{R_i,1} \right) \right\}$$

$$+ O(n^{-s}), \quad \text{a.s.}$$

Therefore for $\widehat{T}_{n,1}(\boldsymbol{Y}, \boldsymbol{R})$ we have, as $n \to \infty$,

$$\widehat{T}_{n,1}(\boldsymbol{Y}, \boldsymbol{R}) = \frac{1}{4f(0)\widetilde{\sigma}_n} \left(\sum_{i=1}^{n} \boldsymbol{z}_i \operatorname{sign} \widehat{e}_{R_i,1} \right)^T \left(\sum_{i=1}^{n} \boldsymbol{z}_i \boldsymbol{z}_i^T \right)^{-1} \left(\sum_{i=1}^{n} \boldsymbol{x}_i \operatorname{sign} \widehat{e}_{R_i,1} \right)$$
$$+ O(n^{-s}), \text{ a.s.}$$

and under the additional assumption $E\,|e_i| < \infty$ we have

$$\widetilde{\sigma}_n = E\,|e_1| + O(n^{-s}), \quad \text{a.s.}$$

for some $s > 0$. The asymptotic relation $\widehat{T}_{n,1}(\boldsymbol{Y}, \boldsymbol{R})$ and $\widehat{T}_{n,2}(\boldsymbol{Y}, \boldsymbol{R})$ is now visible. Clearly, given $\boldsymbol{Y}$, $\sum_{i=1}^{n} \boldsymbol{x}_i \operatorname{sign} \widehat{e}_{R_i}$ is a vector of linear rank statistics. Applying Theorem 4.1 by Hájek (1961) and noticing

$$E \left(\sum_{i=1}^{n} \boldsymbol{z}_i \operatorname{sign} \widehat{e}_{R_i} | \boldsymbol{Y} \right) = \sum_{i=1}^{n} \boldsymbol{z}_i \frac{1}{n} \sum_{j=1}^{n} \operatorname{sign} \widehat{e}_{j,1} = \boldsymbol{0}$$

and

$$var \left(\sum_{i=1}^{n} \boldsymbol{x}_i \operatorname{sign} \widehat{e}_{R_i} | \boldsymbol{Y} \right) = \sum_{i=1}^{n} \boldsymbol{z}_i \boldsymbol{z}_i^T \frac{1}{n-1} \left(n - \left(\frac{1}{n} \sum_{j=1}^{n} \operatorname{sign} \widehat{e}_{j,1} \right)^2 \right)$$

we get the asymptotic normality of $\sum_{i=1}^{n} \boldsymbol{x}_i \operatorname{sign} \widehat{e}_{R_i,1}$. Also, as soon as $n^{\beta}||\boldsymbol{\gamma}|| \to 0$ for some $\beta > 0$,

$$\frac{1}{n}\sum_{j=1}^{n} \operatorname{sign} \widehat{e}_{i,1} \to 0, \quad \text{a.s.}$$

The proof can be finished in a standard way. ■

5. Numerical illustration

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

We considered the model

$$Y_i = \beta_0 + x_i\beta_1 + z_i\gamma + e_i, \quad i = 1, \ldots, n,$$

where the errors were simulated from the normal, Laplace and Cauchy distributions. The following parameter values were used: sample sizes $n = 100$ and $n = 500$; $\beta_0 = 1$; $\beta_1 = 1$; $\gamma = 0$ and $\gamma = 1$.

Concerning the design points we consider matrices, where $x_1^1, \ldots, x_n^1$, $z_1^1, \ldots, z_n^1$ were generated from a uniform distribution on the interval (-0.1,0.1) (matrix 2), (-10,10) (matrix 1) and (-1000,1000) (matrix 3), and transformed to satisfy the assumption A.1.

We proceeded as follows:

(1) $\widehat{e}_{i,2}$, $\widetilde{e}_{i,2}$ $\widehat{e}_{i,1}$ and $\widetilde{e}_{i,1}$ are computed;

(2) a random permutation $\mathbf{r}$ of $(1, \ldots, n)$ is generated;

(3) $\widehat{T}_n(\boldsymbol{Y},\boldsymbol{r}), \widetilde{T}_n(\boldsymbol{Y},\boldsymbol{r}), \widehat{T}_{n,1}(\boldsymbol{Y},\boldsymbol{r}), \widetilde{T}_{n,1}(\boldsymbol{Y},\boldsymbol{r}), \widehat{T}_{n,2}(\boldsymbol{Y},\boldsymbol{r})$ and $\widetilde{T}_{n,2}(\boldsymbol{Y},\boldsymbol{r})$ are calculated;

(4) steps (2) and (3) are repeated 10 000 times;

(5) related empirical quantiles $\widehat{t}_n(1-\alpha)$, $\widetilde{t}_n(1-\alpha)$, $\widehat{t}_{n,1}(1-\alpha)$, $\widetilde{t}_{n,1}$ $(1-\alpha)$, $\widehat{t}_{n,2}(1-\alpha)$ and $\widetilde{t}_{n,2}(1-\alpha)$ are computed.

Empirical quantiles $\widehat{t}_n(1-\alpha)$, $\widetilde{t}_n(1-\alpha)$, $\widehat{t}_{n,1}(1-\alpha)$, $\widetilde{t}_{n,1}(1-\alpha)$, $\widehat{t}_{n,2}(1-\alpha)$ and $\widetilde{t}_{n,2}(1-\alpha)$ are approximations for $\widehat{t}_n(\boldsymbol{Y}, 1-\alpha)$, $\widetilde{t}_n(\boldsymbol{Y}, 1-\alpha)$, $\widehat{t}_{n,1}(\boldsymbol{Y}, 1-\alpha)$, $\widetilde{t}_{n,1}(\boldsymbol{Y}, 1-\alpha)$ and $\widehat{t}_{n,2}(\boldsymbol{Y}, 1-\alpha)$, $\widetilde{t}_{n,2}(\boldsymbol{Y}, 1-\alpha)$. Quantiles (90%, 95%) calculated under H_0 and H_1 ($\gamma = 1$) are summarized in Tables 1 and 2.

Table 1. Empirical quantiles of $\widehat{t}_n(1-\alpha)$, $\widetilde{t}_n(1-\alpha)$, $\widehat{t}_{n,1}(1-\alpha)$, $\widetilde{t}_{n,1}(1-\alpha)$, $\widehat{t}_{n,2}(1-\alpha)$ and $\widetilde{t}_{n,2}(1-\alpha)$ for three different design matrices and given $e_1, \dots, e_n$ under H_0 ($\gamma = 0$) and H_1 ($\gamma = 1$) simulated from normal distribution.

quantile	matrix 1		matrix 2		matrix 3	
	90%	95%	90%	95%	90%	95%
$n = 100 \quad \gamma = 0$						
$\widehat{t}_n(1-\alpha)$	2.647	3.731	2.710	3.817	2.750	4.002
$\widetilde{t}_n(1-\alpha)$	2.617	3.691	2.670	3.758	2.756	3.939
$\widehat{t}_{n,1}(1-\alpha)$	1.544	2.135	1.207	1.677	1.329	1.892
$\widetilde{t}_{n,1}(1-\alpha)$	1.395	1.977	1.050	1.546	1.329	1.896
$\widehat{t}_{n,2}(1-\alpha)$	2.726	3.792	2.697	3.828	2.737	3.894
$\widetilde{t}_{n,2}(1-\alpha)$	2.696	3.703	2.656	3.828	2.692	3.863
$n = 100 \quad \gamma = 1$						
$\widehat{t}_n(1-\alpha)$	2.653	3.750	2.675	3.830	2.729	3.855
$\widetilde{t}_n(1-\alpha)$	2.617	3.691	2.670	3.758	2.756	3.939
$\widehat{t}_{n,1}(1-\alpha)$	14.564	19.562	1.423	1.966	1310.621	1855.426
$\widetilde{t}_{n,1}(1-\alpha)$	1.395	1.977	1.050	1.546	1.329	1.896
$\widehat{t}_{n,2}(1-\alpha)$	2.646	3.719	2.697	3.828	2.784	3.916
$\widetilde{t}_{n,2}(1-\alpha)$	2.696	3.703	2.656	3.828	2.692	3.863
$n = 500 \quad \gamma = 0$						
$\widehat{t}_n(1-\alpha)$	2.642	3.696	2.760	3.940	2.809	3.839
$\widetilde{t}_n(1-\alpha)$	2.641	3.663	2.778	3.939	2.794	3.834
$\widehat{t}_{n,1}(1-\alpha)$	2.179	3.012	2.030	2.822	2.112	2.909
$\widetilde{t}_{n,1}(1-\alpha)$	2.195	3.025	1.948	2.764	2.195	2.959
$\widehat{t}_{n,2}(1-\alpha)$	2.685	3.746	2.725	3.863	2.714	3.782
$\widetilde{t}_{n,2}(1-\alpha)$	2.693	3.770	2.670	3.799	2.661	3.767
$n = 500 \quad \gamma = 1$						
$\widehat{t}_n(1-\alpha)$	2.801	3.882	2.736	3.921	2.709	3.846
$\widetilde{t}_n(1-\alpha)$	2.641	3.663	2.778	3.939	2.794	3.834
$\widehat{t}_{n,1}(1-\alpha)$	11.573	16.694	1.942	2.769	1812.366	2609.936
$\widetilde{t}_{n,1}(1-\alpha)$	2.195	3.025	1.949	2.764	2.195	2.959
$\widehat{t}_{n,2}(1-\alpha)$	2.769	3.924	2.724	3.907	2.721	3.932
$\widetilde{t}_{n,2}(1-\alpha)$	2.693	3.770	2.670	3.799	2.661	3.767

Notice that sample sizes $n = 100$ and $n = 500$ give similar results.

Table 2. Empirical quantiles of $\widehat{t}_n(1-\alpha)$, $\widetilde{t}_n(1-\alpha)$, $\widehat{t}_{n,1}(1-\alpha)$, $\widetilde{t}_{n,1}(1-\alpha)$, $\widehat{t}_{n,2}(1-\alpha)$ and $\widetilde{t}_{n,2}(1-\alpha)$ for three different generated samples $e_1, \dots, e_n$, the matrix is fixed (matrix 1).

quantile	normal		Laplace		Cauchy	
	90%	95%	90%	95%	90%	95%
$n = 100 \quad \gamma = 0$						
$\widehat{t}_n(1-\alpha)$	2.647	3.731	2.651	3.641	2.688	3.700
$\widetilde{t}_n(1-\alpha)$	2.617	3.691	2.618	3.663	2.652	3.635
$\widehat{t}_{n,1}(1-\alpha)$	1.544	2.135	1.393	2.009	0.443	0.615
$\widetilde{t}_{n,1}(1-\alpha)$	1.395	1.977	1.156	1.689	0.428	0.590
$\widehat{t}_{n,2}(1-\alpha)$	2.726	3.792	2.662	3.674	2.798	3.892
$\widetilde{t}_{n,2}(1-\alpha)$	2.696	3.703	2.607	3.678	2.711	3.854
$n = 100 \quad \gamma = 1$						
$\widehat{t}_n(1-\alpha)$	2.653	3.750	2.603	3.738	2.720	3.735
$\widetilde{t}_n(1-\alpha)$	2.617	3.691	2.618	3.663	2.652	3.635
$\widehat{t}_{n,1}(1-\alpha)$	14.564	19.562	8.874	12.241	2.732	3.821
$\widetilde{t}_{n,1}(1-\alpha)$	1.395	1.977	1.156	1.689	0.428	0.590
$\widehat{t}_{n,2}(1-\alpha)$	2.646	3.719	2.641	3.649	2.694	3.806
$\widetilde{t}_{n,2}(1-\alpha)$	2.696	3.703	2.607	3.678	2.711	3.854

The results in Tables 1 and 2 on $\widehat{t}_n(1-\alpha)$, $\widetilde{t}_n(1-\alpha)$, $\widehat{t}_{n,2}(1-\alpha)$ and $\widetilde{t}_{n,2}(1-\alpha)$ are in accordance with the asymptotic results presented in Theorems A, B and C. Notice that the empirical quantiles based on the reduced and full models do not show any difference. There is a different situation for $\widehat{t}_{n,1}(1-\alpha)$ and $\widetilde{t}_{n,1}(1-\alpha)$. Moreover, Tables 1 and 2 indicate for these quantiles an influence of the particular design matrix for given error terms $e_1, \dots, e_n$ and vice versa, a certain influence the distribution of $e_1, \dots, e_n$ for the same design matrix.

According to Tables 1 and 2 it is recommended to apply the empirical quantiles based on the full model.

To avoid the influence of the error terms we 1000 times generated the sample $e_1, \dots, e_n$ from normal distribution and for every case we apply the permutation procedure. Table 3 shows the mean of obtained selected empirical quantiles under H_0 ($\gamma = 0$).

Table 3. The mean of empirical quantiles for $\widehat{t}_n(1-\alpha)$, $\widetilde{t}_n(1-\alpha)$, $\widehat{t}_{n,1}(1-\alpha)$, $\widetilde{t}_{n,1}(1-\alpha)$, $\widehat{t}_{n,2}(1-\alpha)$ and $\widetilde{t}_{n,2}(1-\alpha)$ for 1000 replications of $e_1, \dots, e_n$ under H_0.

quantile means	$n = 100$			$n = 500$		
	90%	95%	99%	90%	95%	99%
$\widehat{t}_n(1-\alpha)$	2.685	3.783	6.376	2.700	3.820	6.534
$\widetilde{t}_n(1-\alpha)$	2.656	3.746	6.308	2.692	3.815	6.519
$\widehat{t}_{n,1}(1-\alpha)$	1.952	2.693	4.406	2.030	2.845	4.772
$\widetilde{t}_{n,1}(1-\alpha)$	1.824	2.540	4.198	1.971	2.775	4.674
$\widehat{t}_{n,2}(1-\alpha)$	2.676	3.777	6.401	2.689	3.816	6.510
$\widetilde{t}_{n,2}(1-\alpha)$	2.647	3.744	6.350	2.687	3.810	6.497

Table 3 shows that the approximations $\widetilde{t}_n(1-\alpha), \dots$, $\widetilde{t}_{n,1}(1-\alpha)$, $\widetilde{t}_{n,2}(1-\alpha)$ based on residuals of the full models perform better than $\widehat{t}_n(1-\alpha)$, $\widehat{t}_{n,1}(1-\alpha)$, $\widehat{t}_{n,2}(1-\alpha)$. In other words, the approximations $\widetilde{t}_n(1-\alpha), \dots$, $\widetilde{t}_{n,1}(1-\alpha)$, $\widetilde{t}_{n,2}(1-\alpha)$ are preferred.

We are also interested in the testing problem on the hypothesis H_0 : $\gamma = 0$. We used the test statistic T_n (3), T_n (7), $T_{n,1}$ (9) and $T_{n,2}$ (14) and the critical values were taken from Table 3. We generated $e_1, \dots, e_n$ 1 000 times from normal and Laplace distributions and we applied the test for a fixed design matrix (matrix 1). Tables 4 and 5 give the number of rejection of H_0 when $\gamma = 0$ and $\gamma = 0.05$. Figures 1–4 illustrate the simulated power of the test.

Figures 1–4 below present graphs of the power for the considered tests under various setups. They clearly indicate expected results, i.e., for the normally distributed errors the test statistic T_n has a higher power while for errors with Laplace distribution $T_{n,1}$ and $T_{n,2}$ are preferred to T_n.

Table 4. The number of rejection H_0 among 1000 cases on level α=0.1, 0.05, 0.01 for the normal distribution of errors.

estimator	model					
	reduced			full		
	90%	95%	99%	90%	95%	99%
$n = 100 \quad \gamma = 0$						
T_n	101	64	15	104	64	15
$T_{n,1}$	116	64	17	136	73	18
$T_{n,2}$	106	63	8	108	63	8
$n = 100 \quad \gamma = 0.05$						
T_n	889	811	611	889	812	620
$T_{n,1}$	768	666	459	780	683	481
$T_{n,2}$	719	610	360	720	613	365
$n = 500 \quad \gamma = 0$						
T_n	98	48	9	98	48	9
$T_{n,1}$	96	46	8	101	48	9
$T_{n,2}$	88	49	6	88	50	6
$n = 500 \quad \gamma = 0.05$						
T_n	1000	1000	1000	1000	1000	1000
$T_{n,1}$	1000	1000	998	1000	1000	998
$T_{n,2}$	1000	1000	991	1000	1000	991

Table 5. The number of rejection H_0 among 1000 cases on level α=0.1, 0.05, 0.01 for the Laplace distribution of errors.

estimator	model					
	reduced			full		
	90%	95%	99%	90%	95%	99%
$n = 100 \quad \gamma = 0$						
T_n	115	56	14	117	58	15
$T_{n,1}$	54	26	7	65	33	8
$T_{n,2}$	107	47	11	110	49	11
$n = 100 \quad \gamma = 0.05$						
T_n	671	556	331	674	558	336
$T_{n,1}$	754	635	395	763	666	420
$T_{n,2}$	762	668	421	766	674	426
$n = 500 \quad \gamma = 0$						
T_n	96	47	12	96	47	12
$T_{n,1}$	39	16	5	42	20	5
$T_{n,2}$	87	42	8	88	43	9
$n = 500 \quad \gamma = 0.05$						
T_n	998	996	973	998	996	973
$T_{n,1}$	1000	999	997	1000	999	998
$T_{n,2}$	999	999	998	999	999	998

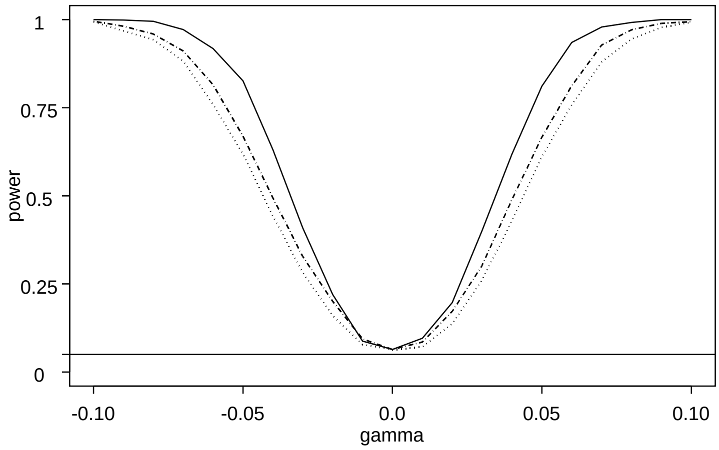

Figure 1. Simulated Type I Error Probability – relative number of rejection of H_0 for T_n (solid), $T_{n,1}$ (dashed), $T_{n,2}$ (dotted) at level $\alpha = 0.05$, sample size $n = 100$, normal distribution of errors.

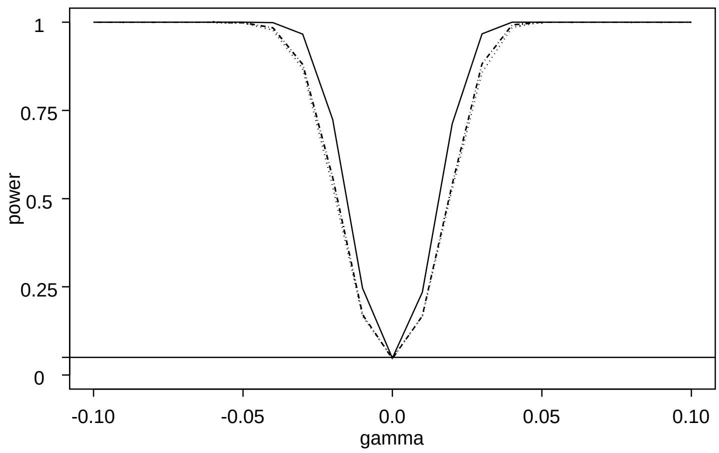

Figure 2. Simulated Type I Error Probability – relative number of rejection of H_0 for T_n (solid), $T_{n,1}$ (dashed), $T_{n,2}$ (dotted) at level $\alpha = 0.05$, sample size $n = 500$, normal distribution of errors.

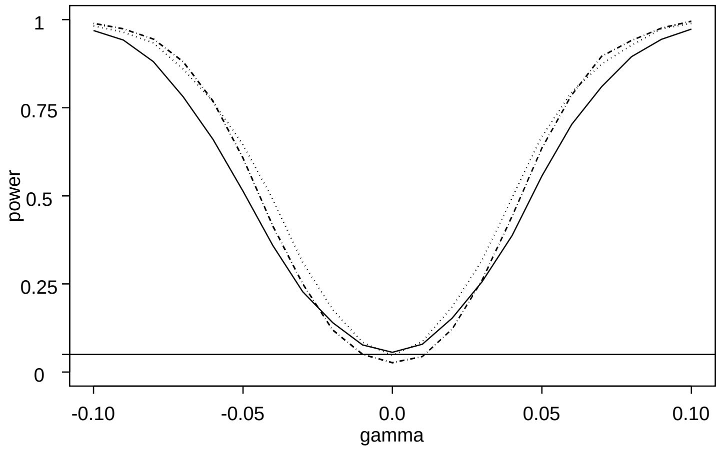

Figure 3. Simulated Type I Error Probability – relative number of rejection of H_0 for T_n (solid), $T_{n,1}$ (dashed), $T_{n,2}$ (dotted) at level $\alpha = 0.05$, sample size $n = 100$, Laplace distribution of errors.

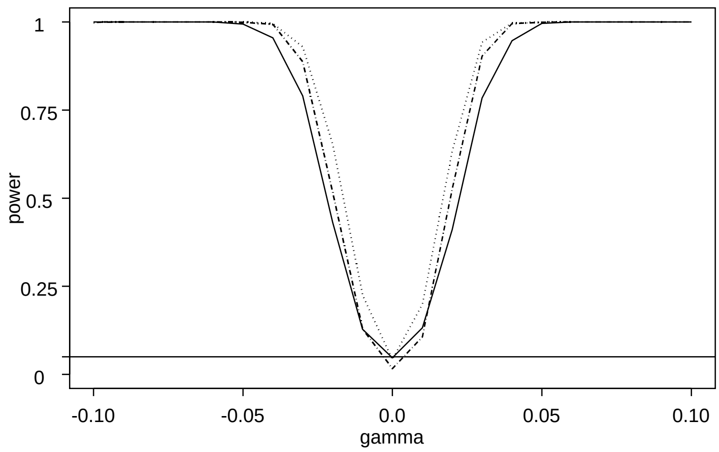

Figure 4. Simulated Type I Error Probability – relative number of rejection of H_0 for T_n (solid), $T_{n,1}$ (dashed), $T_{n,2}$ (dotted) at level $\alpha = 0.05$, sample size $n = 500$, Laplace distribution of errors.

Acknowledgement

The authors thank the referee for his valuable comments and suggestions, which led to an improved version of the text. The work of the first author was supported by grants GAČR 201/03/0945, GAUK 279/2003 and MSM 113200008. The other author's research was supported by grants GAČR 201/02/0049 and MSM 245100303.

References

[1] J. Antoch and M. Hušková, *Detection of structural changes in regression*, Tatra Mountains Publications **26** (2003), 201–215.

[2] B.M. Brown and J.S. Maritz, *Distribution-free methods in regression*, Australian Journal of Statistics **24** (1982), 318–331.

[3] B.S. Cade and J.D. Richards, *Permutation tests for least absolute deviation regression*, Biometrics **52** (1996), 886–902.

[4] D. De Angelis, P. Hall and G.A. Young, *Analytical and bootstrap approximations to estimators in l_1 regression*, Journal of the American Statistical Association **88** (1993), 1310–1316.

[5] D. Edwards, *Exact simulation-based inference: A survey, with additions*, Journal of Statistical Computation and Simulation **22** (1985), 307–326.

[6] B, Efron and R.J. Tibshirani, *An Introduction to the Bootstrap*, Chapman and Hall, New York 1993.

[7] P. Good, *Permutation Tests*, Springer Verlag, New York 1994.

[8] J. Hájek, *Some extensions of the Wald-Wolfowitz-Noether theorem*, Annals of Mathematical Statistics **32** (1961), 506–523.

[9] M. Hušková, *The rate of convergence of simple linear rank statistics*, Annals of Statistics **5** (1977), 658–670.

[10] M. Hušková, *Permutation principle and bootstrap in change point analysis*, in: Asymptotic Methods in Stochastics. Fields Institute Communications, eds. L. Horváth and B. Szyszkowicz **44** pp. 273–292.

[11] M. Hušková and J. Picek, *M-tests for detection of structural changes in regression*, in: Statistical Data Analysis Based on the L_1-Norm and Related Methods, ed. Y. Dodge, Birhäuser, Basel (2002), 213–229.

[12] P.E. Kennedy, *Randomization tests in econometrics*, Journal of business and Economic Statistics **13** (1995), 85–94.

[13] R. Koenker and G. Basset, *Tests of linear hypotheses and* L_1*- regression*, Econometrica **50** (1982), 1577–1583.

[14] E.L. Lehmann, *Theory of Point Estimation*, Wadworth & Brooks/Cole, California 1991.

[15] B.F.J. Manly, *Randomization and Monte-Carlo Methods in Biology*, London: Chapman and Hall 1991.

[16] D.N. Politis, J.P. Romano and M. Wolf, *Subsampling*, Springer Verlag, New York 1999.

[17] C.J.F. ter Braak, *Permutations versus bootstrap significance in multiple regression and ANOVA*, in: Bootstrapping and Related Techniques, eds. K.H. Jockel, G. Rotheand W. Sendler, Berlin-Springer-Verlag (1992), 79–86.

[18] Shao Jun and Tu Dongsheng, *The Jackknife and Bootstrap*, Springer Verlag, New York 1995.

[19] W.J. Welsh, *Construction of permutation tests*, Journal of the American Statistical Association **85** (1990), 693–689.

Received 3 December 2003