Working Paper 95-29
Statistics and Econometrics Series 08
February 1995

Departamento de Estadística y Econometría
Universidad Carlos III de Madrid
Calle Madrid, 126
28903 Getafe (Spain)
Fax (341) 624-9849

BREAKDOWN AND ASYMPTOTIC PROPERTIES OF RESAMPLED τ-ESTIMATES

Jorge Adrover and Ana M. Bianco*

Abstract

In this paper we study the breakdown and asymptotic properties of resampled τ-estimates. We find that they retain the finite breakdown point of the exact estimator. We also study their consistency and their order of convergence under nonstandard assumptions on the loss functions.

Key Words

Robustness; resampling; regression; breakdown point; order of convergence.

*Adrover, Departamento de Estadística y Econometría, Universidad Carlos III de Madrid; Bianco, Universidad de Buenos Aires, Argentina. AMS: 1991 subject classification 62F35.

1 Introduction

High breakdown point of regression estimates have been widely considered in the last ten years. We may mention the least median of squares (LMS) estimate, the S–estimate and the τ–estimate among them. All these estimators are defined by the minimization of a nonconvex function of the residuals which may have several local minima. This fact makes very difficult the search of the absolute minimum. In practice, the procedures more frequently used are based in resampling methods, which yield an approximation of the exact estimator. These approximate estimates are based on the resampling scheme introduced by Rousseeuw (1984) to compute LMS–estimate. Basically, the idea is to look for the minimum among a number of planes that fit subsamples of size p chosen at random. In the case of S–estimates and τ–estimates these resampled estimates may be used as an initial estimator to start an iterative weighted least squares algorithm.

Since the resampling methods give just an approximation of the exact estimator, it is natural to wonder if they retain the same good properties of the exact estimators. Some efforts have been done in this direction. Rousseeuw and Basset (1991) studied the case of the resampled estimates for least trimmed of squares (LTS) and LMS–estimates when all possible p subsets are considered. They proved that the p-subset algorithm shares the equivariance and breakdown properties of the exact estimator. Hawkins (1993) gave some arguments which suggest the consistency of the resampled LMS- and LST–estimates, but he did not give a complete proof of this fact. On the other hand, Stromberg (1993) showed that in the case of LMS–estimates it is possible to compute the exact estimator by searching over the Chebyshev fits of all $(p+1)$–subsets.

The aim of this paper is to study the properties of resampled τ–estimates. We concern about the breakdown point, the consistency and the order of convergence of these approximate estimators.

We assume that the target model is the linear model:

$$y = \boldsymbol{\theta}_o'\mathbf{x} + \epsilon, \tag{1}$$

where $\mathbf{x} = (x_1, \ldots, x_p)'$ is a random vector in $\Re^p$ with distribution G_o, $(\mathbf{x}', y)'$ has distribution H_o, $\boldsymbol{\theta}_o = (\theta_{o1}, \ldots, \theta_{op})'$ is the vector of the true regression parameters and the error ϵ is a random variable with distribution $F_o(u/\sigma)$, independent of $\mathbf{x}$. Since we are considering only equivariant estimators, without loss of generality, we can suppose that $\boldsymbol{\theta}_o = 0$ and $\sigma = 1$.

The τ–estimators were introduced by Yohai and Zamar (1988). They depend on two loss functions ρ_1 and ρ_2. Given a sample $\mathbf{u} = (u_1, \ldots, u_n)'$ denote by F_n the empirical distribution. Huber (1981) defined the M–estimate of scale by:

$$s_n(\mathbf{u}) = \min\left\{s : E_{F_n}\rho_1\left(\frac{u}{s}\right) \leq b\right\}, \tag{2}$$

where E_F denotes the mean under the distribution F and $0 \leq b \leq 1$. A τ-scale of $\mathbf{u}$ satisfies:

$$\tau_n^2(\mathbf{u}) = s_n^2(\mathbf{u})E_{F_n}\rho_2\left(\frac{u}{s_n(\mathbf{u})}\right). \tag{3}$$

In the regression setting, given a sample of random vectors $\mathbf{z}_i = (\mathbf{x}_i', y_i)'$, $i = 1, \ldots, n$, for any $\boldsymbol{\theta} \in \Re^p$ we define the residual vector based on $\boldsymbol{\theta}$ as $\mathbf{r}(\boldsymbol{\theta}) = (r_1(\boldsymbol{\theta}), \ldots, r_n(\boldsymbol{\theta}))'$ where $r_i(\boldsymbol{\theta}) = y_i - \mathbf{x}_i'\boldsymbol{\theta}$. Then, the τ–estimator is defined as the argument $\hat{\boldsymbol{\theta}}_n$ which minimizes the τ–scale of the residuals, that is

$$\hat{\boldsymbol{\theta}}_n = \arg\min_{\boldsymbol{\theta}} \tau_n^2(\mathbf{r}(\boldsymbol{\theta})).$$

Yohai and Zamar (1988) showed that these estimators combine high breakdown point and high efficiency under normal errors if ρ_1 and ρ_2 are properly chosen.

We will consider the family of ρ functions that satisfy the following conditions:

A1: $\rho(0) = 0$.

A2: $\rho(-t) = \rho(t)$.

A3: ρ is a non–decreasing, bounded and left–continuous function. Furthermore, it is continuous at 0.

A4: There exists a real number c such that $\rho(t) = \sup \rho = 1$ if $|t| \geq c$.

In the following we will assume that the τ–estimates are defined using functions ρ_1 and ρ_2 satisfying (**A1**)–(**A4**). It is interesting to note that under regular conditions on the loss functions stated by Zamar and Yohai (1988), if $\rho_1 = \rho_2$ we get an S–estimator. By taking ρ functions as the jump function defined by

$$\rho_J(t) = \begin{cases} 0 & \text{if} \quad |t| \leq 1 \\ 1 & \text{if} \quad |t| > 1 \end{cases}$$

we obtain the least median of squares estimator by minimizing (2) if $b = 1/2$. From now on, we will take ρ_1 such that $E_{F_o}\rho_1(r(\mathbf{0})) = b$ in order to assure the Fisher–consistency of the estimators.

In this paper we consider approximate τ-estimates of the following form. Given $\mathbf{z}_i = (\mathbf{x}_i', y_i)', i = 1, \ldots, n$, a sample of size n, the resampling version of the τ–estimate is defined as follows:

Generate at random N subsamples of size p from the sample. For the k–th subsample ($k = 1, \ldots, N$) fit a hyperplane ($y = \boldsymbol{\alpha}_k'\mathbf{x}$) containing the p points. Define $D_n^N = \{\boldsymbol{\alpha}_k : k = 1, \ldots, N\}$. Then,

$$\mathbf{T}_n = \arg \min_{\boldsymbol{\theta} \in D_n^N} \tau_n^2(\mathbf{r}(\boldsymbol{\theta})). \tag{4}$$

We establish here some notation which will be used along the paper. For the sake of simplicity, we will denote $\tau_n^2(\boldsymbol{\theta}) = \tau_n^2(\mathbf{r}(\boldsymbol{\theta}))$ and $s_n(\boldsymbol{\theta}) = s_n(\mathbf{r}(\boldsymbol{\theta}))$ from now on. Given $\boldsymbol{\lambda} \in \Re^r$, $N_{\boldsymbol{\lambda}^\perp}$ will denote the set $\{\mathbf{x} \in \Re^r : \mathbf{x}'\boldsymbol{\lambda} = 0\}$ and c_2 is defined as $\max\{t : \rho_2(t) = 0\}$.

In Section 2 we find the breakdown point of the resampled τ–estimates and in Section 3 we state their consistency and their order of convergence under some different assumptions. Some technical lemmas and auxilliary results are proved in the Appendix.

2 Breakdown Point

The finite sample breakdown point was defined by Donoho and Huber (1983). Let $\mathbf{Z}_n = \{\mathbf{z}_1, ..., \mathbf{z}_n\} \subset \Re^{p+1}$ with $\mathbf{z}_i = (\mathbf{x}_i', y_i)'$ and $\mathbf{Z}_{n,m}$ be any contaminated sample of size n obtained by replacing m observations of the original sample $\mathbf{Z}_n$ by arbitrary outliers. Let $\mathcal{Z}$ be the set of all these possible $\mathbf{Z}_{n,m}$ and $\mathbf{T}_n = \mathbf{T}_n(\mathbf{Z}_n)$ be the estimate defined by the minimization of the τ–scale over D_n^N. The regression estimate $\mathbf{T}_n$ defined for samples of size n is said to break down at $\mathbf{Z}_n$ for a given m if $\sup \|\mathbf{T}_n(\mathbf{Z}_{n,m})\| = \infty$ where the supremum is taken over all $\mathbf{Z}_{n,m} \subset \mathcal{Z}$. Let m_o be the minimum m that $\mathbf{T}_n$ breaks down. Then, the finite breakdown point of $\mathbf{T}_n$ at $\mathbf{Z}_n$ is $\epsilon^*(\mathbf{T}_n, \mathbf{Z}_n) = m_o/n$.

For the sake of simplicity, we will consider the case where b in equation (2) equals 1/2. Yohai and Zamar (1988) proved that when the observations are in general position, the breakdown point of the τ–estimate is $\epsilon^* \geq (1-2c_n)/2(1-c_n)$, where $c_n = (p-1)/n$. Observe that when $n \to \infty$ this yields $\epsilon^* = 1/2$.

Since the resampled version of the τ–estimator described above contains a randomization, its value for a fixed sample will be random and hence, we have to modify the breakdown definition accordingly. Maronna and Yohai (1993) introduced a suitable notion of breakdown point for this case.

Call $\Pi(n,m,p,N)$ the probability that at least one subsample of size p is contained in $\mathbf{Z}_n$. Then, $\Pi(n,m,p,N) = 1-(1-\beta)^N$, where $\beta = \binom{n-m}{p} / \binom{n}{p} \cong (1-m/n)^p$ (Maronna and Yohai (1993)). This probability can be taken as close to 1 as desired by choosing N large enough.

The result of the next theorem can be interpreted as stating that the resampled τ–estimator does not break down for $m < [n/2]-(p-1)$ with probability at least $\Pi(n,m,p,N)$ which may be taken arbitrarily close to 1. This means that the proposed approximate estimators retain the finite breakdown point of the original estimators with arbitrarily large probability.

Theorem 1 *Assume that ρ_1 and ρ_2 satisfy (***A1***)-(***A4***) and that $c_2 = 0$. Let $\mathbf{Z}_n$ be such that every subset $\{\mathbf{x}_{i_j} : j = 1,...,p\}$ is linearly independent. Let D_n^N be defined by the resampling procedure and $m < [n/2]-(p-1)$. Then, there exists a positive constant K (depending only on $\mathbf{Z}_n$) such that for all $\mathbf{Z}_{n,m} \in \mathcal{Z}$, $\|\mathbf{T}_n(\mathbf{Z}_{n,m})\| \leq K$ with probability larger than $\Pi(n,m,p,N)$.*

Proof: Initially, we will show that there exists a compact set $\mathcal{K} \subset (0,\infty)$ such that $\sup_{\mathbf{Z}_{n,m}\in\mathcal{Z}} s_n(\mathbf{T}_n(\mathbf{Z}_{n,m}))\in\mathcal{K}$. Suppose that there exists a sequence of sets $\{\mathbf{Z}^j_{n,m}\}$ such that

$$\limsup_j s_n(\mathbf{T}_n(\mathbf{Z}^j_{n,m})) = \infty.$$

At the same time, let us assume that

$$\liminf_j E_{H_n}\rho_2\left(\frac{|r_i(\mathbf{T}_n(\mathbf{Z}^j_{n,m}))|}{s_n(\mathbf{T}_n(\mathbf{Z}^j_{n,m}))}\right) = 0.$$

Then $\liminf_j E_{H_n}\rho_1\left(\frac{|r_i(\mathbf{T}_n(\mathbf{Z}^j_{n,m}))|}{s_n(\mathbf{T}_n(\mathbf{Z}^j_{n,m}))}\right) = 0$. By taking a subsequence if necessary, we can choose a positive sequence ϵ_j such that $\lim_{j\to\infty}\epsilon_j = 0$ and such that

$$\liminf_j E_{H_n}\rho_1\left(\frac{|r_i(\mathbf{T}_n(\mathbf{Z}^j_{n,m}))|}{s_n(\mathbf{T}_n(\mathbf{Z}^j_{n,m}))-\epsilon_j}\right) = 0.$$

This contradicts the fact $s_n(\mathbf{T}_n(\mathbf{Z}^j_{n,m}))$ is a minimum for $j \geq j_o$ and we conclude that

$$\liminf_j E_{H_n}\rho_2\left(\frac{r(\mathbf{T}_n(\mathbf{Z}^j_{n,m}))}{s_n(\mathbf{T}_n(\mathbf{Z}^j_{n,m}))}\right) > 0.$$

Therefore, we obtain that

$$\limsup_j \tau_n^2(\mathbf{T}_n(\mathbf{Z}^j_{n,m})) = \infty.$$

However, for every $\mathbf{Z}_{n,m}^j$ there always exists a vector in D_n^N, $\mathbf{T}_n^{j,G}$, which fits exactly to p points in $\mathbf{Z}_n$. Since at least $[n/2]+p$ points belong to $\mathbf{Z}_n$, equation (2) implies that $\limsup_j s_n(\mathbf{T}_n^{j,G}) < \infty$. Therefore, $\tau_n^2(\mathbf{T}_n^{j,G}) < \tau_n^2(\mathbf{T}_n(\mathbf{Z}_{n,m}^j))$ for $j \geq j_o$ which contradicts the definition of τ–estimators.

Now assume that $\limsup_j \|\mathbf{T}_n(\mathbf{Z}_{n,m}^j)\| = \infty$. Then, at least $[n/2]+1$ residuals should tend to infinity. To prove this, notice that without loss of generality we may suppose that

$$\lim_j \frac{\mathbf{T}_n(\mathbf{Z}_{n,m}^j)}{\|\mathbf{T}_n(\mathbf{Z}_{n,m}^j)\|} = \mathbf{T}_o.$$

Let us take $N_{\mathbf{T}_o^\perp} \subset \Re^p$. If $(\mathbf{x}_i', y_i) \in \mathbf{Z}_n$, $r_i(\mathbf{T}_n) \neq 0$ and $\mathbf{x}_i \notin N_{\mathbf{T}_o^\perp}$, then

$$\lim_j \left(y_i - \|\mathbf{T}_n(\mathbf{Z}_{n,m}^j)\| \mathbf{x}_i' \frac{\mathbf{T}_n(\mathbf{Z}_{n,m}^j)}{\|\mathbf{T}_n(\mathbf{Z}_{n,m}^j)\|} \right) = \infty.$$

There exist at most $p-1$ points such that in $(\mathbf{x}_i', y_i) \in \mathbf{Z}_n$ and $\mathbf{x}_i \in N_{\mathbf{T}_o^\perp}$. Thus we can find at least $[n/2]+1$ residuals which tend to infinity. Therefore, by (2) $\limsup_j s_n(\mathbf{T}_n(\mathbf{Z}_{n,m}^j)) = \infty$ which contradicts the first part of the proof. □

Remark 1 *The case $b \neq 1/2$ follows similarly to the one treated above with the corresponding breakdown point.*

Remark 2 *The case of resampled least median of squares estimator can be proved in a similar way, obtaining high probabilistic breakdown point.*

3 Asymptotic Theory

In order to obtain the asymptotic properties of resampled τ–estimates some technical results are needed. Given a closed set A in an Euclidean space, define $L^\infty(A)$ as the set of all bounded and measurable real functions on A metrized with the supremum norm. Let $\mathcal{S}_q = \{\boldsymbol{\lambda} \in \Re^q : \|\boldsymbol{\lambda}\| = 1\}$ and $\mathbf{z}_1, \ldots, \mathbf{z}_n$ be i.i.d. random vectors in $\Re^q$ with distribution P and call H_n the corresponding empirical distribution. Define for $\boldsymbol{\gamma} \in \mathcal{S}_q$ the process $U_n(\boldsymbol{\gamma}) = n^{1/2}\left(H_n(\boldsymbol{\gamma}'\mathbf{z} \leq 0) - P(\boldsymbol{\gamma}'\mathbf{z} \leq 0)\right)$. Proposition 1 of Section 4 of Beran and Millar (1986) states that U_n converges weakly as a random element of $L^\infty(\mathcal{S}_q)$ to a Gaussian process. A generalization of Glivenko-Cantelli theorem due to Wolfowitz establishes that $V_n(\boldsymbol{\gamma}) = H_n(\boldsymbol{\gamma}'\mathbf{z} \leq 0) - P(\boldsymbol{\gamma}'\mathbf{z} \leq 0)$ converges almost surely uniformly in $\boldsymbol{\gamma}$.

The following lemma, which is proved in the Appendix, is a generalization of these results.

We will denote by $\rho^*(t) = 1 - \rho(t)$ (we have $\rho_1^*(t) = 1 - \rho_1(t)$ and $\rho_2^*(t) = 1 - \rho_2(t)$ as particular cases). Obviously, this does not affect the definition of the M-scale estimator if one replaces b by $1-b$ in (2).

Lemma 1 *Assume that ρ satisfies conditions* (**A1**)*-*(**A4**) *and that $(\mathbf{x}_i', y_i), i = 1, \ldots, n$ are i.i.d. random vectors in $\Re^{p+1}$. If $r(\boldsymbol{\theta}) = y - \mathbf{x}'\boldsymbol{\theta}$ where $\boldsymbol{\theta} \in \Re^p$ and H_n denotes the corresponding empirical function, then*

(a) $$n^{1/2} \sup_{\boldsymbol{\theta}, s} \left| E_{H_n}\rho^*\left(\frac{r(\boldsymbol{\theta})}{s}\right) - E\rho^*\left(\frac{r(\boldsymbol{\theta})}{s}\right) \right| = O_p(1).$$

(b) $P\left(\lim_{n\to\infty}\sup_{\boldsymbol{\theta},s}\left|E_{H_n}\rho^*\left(\frac{r(\boldsymbol{\theta})}{s}\right) - E\rho^*\left(\frac{r(\boldsymbol{\theta})}{s}\right)\right| = 0\right) = 1.$

Let f_o a density of the distribution function F_o. Some of the following assumptions will be required along the paper:

F1. f_o is a bounded, non–increasing and continuous function.

F2. f_o and ρ_1^* have a common point of decrease.

F3. $\int_0^c f_o(t)t d\rho_1^*(t) < 0.$

F4. $\int_0^c f_o'(t) d\rho_1^*(t) > 0.$

F5. $\int_0^c f_o'(t) d\rho_2^*(t) > 0.$

F6. $f_o'(t)$ is bounded and continuous in $\Re$.

F7. $\rho_1(c_2) < b.$

F8. $\sup_{\boldsymbol{\theta}\geq m_1>0} P_{G_o}(|\mathbf{x}'\boldsymbol{\theta}| \leq c_2) < 1$, where $m_1 = \inf\{\|\boldsymbol{\theta}\| : E_{G_o}\rho_1^{(s)}(\mathbf{x}'\boldsymbol{\theta}) \geq b\}$ if $\rho_1^{(s)}(t) = \lim_{x\to t^+}\rho_1(x)$.

Define $R(\boldsymbol{\theta}, s) = E_{H_o}\rho^*\left(\frac{y-\mathbf{x}'\boldsymbol{\theta}}{s}\right)$. Davies (1990, page 1655) states the following result which will be useful here.

Lemma 2 *Assume that* (**F1**)*,*(**F2**) *and* (**A1**)*–*(**A4**) *hold. Then*

(i) $R : \Re^p \times \Re^+ \to [0,1]$ *is continuous.*

(ii) $R(\boldsymbol{\theta}, r) \leq R(\mathbf{0}, r)$ *for* $\boldsymbol{\theta} \in \Re^p$, $r \in \Re^+$.

(iii) $\sup_{\|\boldsymbol{\theta}\|>\eta} R(\boldsymbol{\theta}, 1) < R(\mathbf{0}, 1)$ *for all* $\eta > 0$.

(iv) $R(\boldsymbol{\theta}, r)$ *is strictly increasing in* r *for fixed* $\boldsymbol{\theta}$.

Let $s(\boldsymbol{\theta})$ be defined through the equation

$$E_{H_o}\rho_1\left(\frac{r(\boldsymbol{\theta})}{s(\boldsymbol{\theta})}\right) = b. \tag{5}$$

Then, we get the following result which was proved by Yohai and Zamar (1988) under slightly stronger conditions.

Corollary 1 *Assume* (**F1**)*,*(**F2**) *and* (**A1**)*–*(**A4**) *hold. Let* $s(\boldsymbol{\theta})$ *be defined by equation (5). Then,* $s(\boldsymbol{\theta})$ *has a unique minimum at* $\boldsymbol{\theta} = \mathbf{0}$.

Proof: By Lemma 2 (ii) and (iv) $s(\mathbf{0}) \le s(\boldsymbol{\theta})$ for all $\boldsymbol{\theta} \in \Re^p$. Now assume that there exists $\boldsymbol{\theta}_1$ such that $s(\boldsymbol{\theta}_1) = s(\mathbf{0})$. By Lemma 2 (iii), we get that

$$b = R(\boldsymbol{\theta}_1, s(\mathbf{0})) < \sup_{\|\boldsymbol{\theta}\| > \|\boldsymbol{\theta}_1\|/2} R(\boldsymbol{\theta}, s(\mathbf{0})) < R(\mathbf{0}, s(\mathbf{0})) = b$$

since this is a contradiction, the statement follows. □

The proof of the following theorem is in the Appendix.

Theorem 2 *Assume that ρ_1 satisfies* (**A1**)- (**A4**) *and* (**F1**)-(**F3**) *hold. Then*

a) There exist positive constants K_1, K_2 and δ such that

i) $E_{H_o}\rho_1^*\left(\frac{r(\boldsymbol{\theta})}{s}\right) > b + K_1(s-1)$ *if* $s > 1$ *and* $\|\boldsymbol{\theta}\| < \delta$.

ii) $\sup_{\boldsymbol{\theta} \in \Re^p} E_{H_o}\rho_1^*\left(\frac{r(\boldsymbol{\theta})}{s}\right) < b - K_2(1-s)$ *if* $s < 1$.

b) If $\{\boldsymbol{\theta}_n\}$ denotes a sequence of random vectors satisfying $\|\boldsymbol{\theta}_n\| = o_p(1)$, then $n^{1/2}(s_n(\boldsymbol{\theta}_n) - 1) = O_p(1)$.

To get consistency and weak rates of convergence we will require some further assumptions concerning to the functions ρ:

A5: ρ is absolutely continuous with a bounded Lebesgue density ψ.

A6: $2\rho(t) - \psi(t)t \ge 0$ a.e. t.

A7: ρ is a continuous function.

The following lemma, proved by Yohai and Zamar (1988) under slightly stronger conditions, states the Fisher–consistency of the τ–estimates.

Lemma 3 *Assume that ρ_1 satisfies* (**A1**)-(**A4**), *ρ_2 satisfies* (**A1**)-(**A6**) *and f_o verifies* (**F1**) *and* (**F2**). *Let us suppose that* $\sup_{\boldsymbol{\lambda} \in \Re^p - \{\mathbf{0}\}} P(\boldsymbol{\lambda}'\mathbf{x} = 0) < 1$. *Define $\tau(\boldsymbol{\theta})$ by*

$$\tau^2(\boldsymbol{\theta}) = s^2(\boldsymbol{\theta}) E_{H_o}\rho_2\left(\frac{r(\boldsymbol{\theta})}{s(\boldsymbol{\theta})}\right),$$

where $s(\boldsymbol{\theta})$ is defined by (5) using as ρ the function ρ_1. Then, $\tau(\boldsymbol{\theta})$ has a unique minimum at $\boldsymbol{\theta} = \mathbf{0}$.

The following lemma will be useful in order to show that the resampled τ–estimator is weakly consistent. The proof can be found in the Appendix.

Lemma 4 *Under* (**A1**)-(**A4**), (**F1**) *and* (**F2**),

(i) $\liminf_{m_1 \to \infty, m_2 \to \infty} \liminf_{n \to \infty} \inf_{\|\boldsymbol{\theta}\| > m_1 s, \|\boldsymbol{\theta}\| > m_2} E_{H_n}\rho_1\left(\frac{r(\boldsymbol{\theta})}{s}\right) > b$ *a.e.*

(ii) *Given* $m > 0$, $\liminf_{s_o \to 0} \liminf_{n\to\infty} \inf_{\|\boldsymbol{\theta}\| \le m, s < s_o} E_{H_n} \rho_1 \left(\frac{r(\boldsymbol{\theta})}{s} \right) > b$ *a.e.*

(iii) *Given* $m > 0$, $\limsup_{s_o \to \infty} \limsup_{n\to\infty} \sup_{s > s_o, \|\boldsymbol{\theta}\| < m} E_{H_n} \rho_1 \left(\frac{r(\boldsymbol{\theta})}{s} \right) = 0$ *a.e.*

(iv) $\limsup_{s_o \to 0, m\to\infty} \limsup_{n\to\infty} \sup_{\|\boldsymbol{\theta}\| > m, \|\boldsymbol{\theta}\| < s s_o} E_{H_n} \rho_1 \left(\frac{r(\boldsymbol{\theta})}{s} \right) = 0$ *a.e.*

The consistency of the resampled τ–estimators will be obtained if one of the following three conditions is valid:

(I) ρ_1 satisfies **(A1)**-**(A4)** and **(A7)**, ρ_2 satisfies **(A1)**-**(A6)** and **(F7)** holds,

(II) $\rho_1(t) = \rho_2(t) = \rho_J(t)$ for all $t \in \Re$

(III) $\rho_1 = \rho_J$, ρ_2 satisfies **(A1)**-**(A6)**, **(F7)** and **(F8)** hold.

Then, we get the following result whose proof may be found in the Appendix.

Lemma 5 *Assume* **(F1)**,**(F2)**. *Let us suppose that* $\sup_{\boldsymbol{\lambda} \in \Re^p - \{\mathbf{0}\}} P(\boldsymbol{\lambda}'\mathbf{x} = 0) < 1 - b$. *Let* $s(\boldsymbol{\theta})$ *be defined by equation (5). If any of the assumptions (I), (II) or (III) is valid, then*

(a) *Given a compact set* $C \subset \Re^p$,

(6) $$\lim_{n\to\infty} \sup_{\boldsymbol{\theta} \in C} |s_n(\boldsymbol{\theta}) - s(\boldsymbol{\theta})| = 0 \text{ almost surely}$$

(7) $$\lim_{n\to\infty} \sup_{\boldsymbol{\theta} \in C} |\tau_n(\boldsymbol{\theta}) - \tau(\boldsymbol{\theta})| = 0 \text{ almost surely}$$

(b) $\inf_{\boldsymbol{\theta} \in \Re^p} E_{H_o} \rho_2 \left(\frac{r(\boldsymbol{\theta})}{s(\boldsymbol{\theta})} \right) > 0.$

(c) $\liminf_{n\to\infty} \inf_{\boldsymbol{\theta} \in \Re^p} E_{H_n} \rho_2 \left(\frac{r(\boldsymbol{\theta})}{s_n(\boldsymbol{\theta})} \right) > 0$ *a.e.*.

Consistency is shown in the following theorem.

Theorem 3 *Assume that* $(\mathbf{x}_1', y_1)', \ldots, (\mathbf{x}_n', y_n)'$ *are (p+1)-dimensional random vectors i.i.d. which satisfy model (1),* $\mathbf{T}_n$ *is the approximate estimate defined by (4) and* $D_n^N = \{\boldsymbol{\alpha}_1, \ldots, \boldsymbol{\alpha}_N\}$ *is an arbitrary set in* $\Re^p$ *satisfying that* $\min_{\boldsymbol{\alpha} \in D_n^N} \|\boldsymbol{\alpha}\| = o_p(1)$. *Suppose also that* $\sup_{\boldsymbol{\lambda} \in \Re^p - \{\mathbf{0}\}} P_{G_o}(\boldsymbol{\lambda}'\mathbf{x} = 0) < 1 - b$, *that* **(F1)**-**(F3)** *are satisfied. If any of the assumptions (I), (II) or (III) is valid, then*

(a) $\|\mathbf{T}_n\| = O_p(1)$

(b) $\lim_n \|\mathbf{T}_n\| = 0$ *in probability.*

Proof:

(a) Let us take the set $A_1 = \{\liminf_{n\to\infty} s_n(\mathbf{T}_n) > 0\}$. Lemma 4 (i) and (ii) entail that $P_{H_o}(A_1) = 1$. Let us take $\boldsymbol{\alpha}_n = \arg\min_{\boldsymbol{\alpha}\in D_n^N} \|\boldsymbol{\alpha}\|$. By hipothesis $\boldsymbol{\alpha}_n = o_p(1)$ and $|s_n(\boldsymbol{\alpha}_n) - 1| = o_p(1)$ by Theorem 2 (b). By Lemma 1 (a) and Lemma 2 (i) $1 - E_{H_n}\rho_2\left(\frac{r(\boldsymbol{\alpha}_n)}{s_n(\boldsymbol{\alpha}_n)}\right) = o_p(1)$. By the definition of $\mathbf{T}_n$, $\tau^2(\mathbf{T}_n) \le \tau^2(\boldsymbol{\alpha}_n)$. Thus, using Lemma 5 (c) we obtain

$$(8) \qquad s_n^2(\mathbf{T}_n) \le \left(E_{H_n}\rho_2\left(\frac{r(\mathbf{T}_n)}{s_n(\mathbf{T}_n)}\right)\right)^{-1} s_n^2(\boldsymbol{\alpha}_n) E_{H_n}\rho_2\left(\frac{r(\boldsymbol{\alpha}_n)}{s_n(\boldsymbol{\alpha}_n)}\right) = O_p(1).$$

Therefore, $s_n^2(\mathbf{T}_n) = O_p(1)$. Given $K > 0$

$$\{\|\mathbf{T}_n\| > K\} \subset \{\|\mathbf{T}_n\| > K \cap \|\mathbf{T}_n\|/s_n(\mathbf{T}_n) > \sqrt{K}\} \cup \{s_n(\mathbf{T}_n) > \sqrt{K}\}$$

By taking K large enough and using Lemma 4 (i) and (8) we get that $\|\mathbf{T}_n\| = O_p(1)$.

(b) Since $\min_{\boldsymbol{\alpha}\in D_n^N} \|\boldsymbol{\alpha}\| = o_p(1)$ by Theorem 2 (b) we have that $|s_n(\boldsymbol{\alpha}_n) - 1| = o_p(1)$. By (a), given $\epsilon > 0$ there exists $K > 0$ such that $P(\|\mathbf{T}_n\| \le K) > 1 - \epsilon$ if $n > n_o(\epsilon, K)$. Let us note that $\lim_{\|\boldsymbol{\theta}\|\to\infty} \tau^2(\boldsymbol{\theta}) = \infty$. Then, by Lemma 3 and continuity of τ, given $\epsilon_o > 0$ there exists $\delta > 0$ such that

$$\inf_{\epsilon_o \le \|\boldsymbol{\theta}\| \le K} \tau^2(\boldsymbol{\theta}) > \tau^2(\mathbf{0}) + \delta.$$

Given $\epsilon_1/2 > 0$ and $\delta/2$ it holds that $P(E_n) = P(\{\sup_{\epsilon_o \le \|\boldsymbol{\theta}\| \le K} |\tau^2(\boldsymbol{\theta}) - \tau_n^2(\boldsymbol{\theta})| \le \delta/2\}) > 1 - \epsilon_1/2$ and $P(\tau_n^2(\boldsymbol{\alpha}_n) - \tau^2(\mathbf{0}) > \delta/2) \le \epsilon_1/2$ for n large enough . Therefore

$$1 - \epsilon_1/2 < P(E_n) = P\left(\left\{\inf_{\epsilon_o \le \|\boldsymbol{\theta}\| \le K} \tau^2(\boldsymbol{\theta}) - \tau_n^2(\boldsymbol{\theta}) + \tau_n^2(\boldsymbol{\theta}) > \tau^2(\mathbf{0}) + \delta\right\} \cap E_n\right) \le$$
$$P\left(\inf_{\epsilon_o \le \|\boldsymbol{\theta}\| \le K} \delta/2 + \tau_n^2(\boldsymbol{\theta}) - \tau^2(\mathbf{0}) > \delta\right) \le P(\inf_{\epsilon_o \le \|\boldsymbol{\theta}\| \le K} \tau_n^2(\boldsymbol{\theta}) - \tau^2(\mathbf{0}) > \delta/2)$$

Consequently,

$$P(\|\mathbf{T}_n\| \ge \epsilon_o) \le P(\tau_n^2(\mathbf{T}_n) > \tau^2(\mathbf{0}) + \delta/2) + \epsilon_1/2 \le P(\tau_n^2(\boldsymbol{\alpha}_n) > \tau^2(\mathbf{0}) + \delta/2) + \epsilon_1/2 \le \epsilon_1$$

for n large enough. Since the result holds for every ϵ_o the conclusion of (b) follows. □

Corollary 2 *Under the same assumptions as in Theorem 3, we have that* $n^{1/2}(s_n(\mathbf{T}_n) - 1) = O_p(1)$.

Theorem 3 requires the existence of a consistent sequence of vectors in the resampling set. The next lemma, which is proved in the Appendix, shows that this is guaranteed by the resampling scheme.

Lemma 6 *Assume that $D_n^{N(n)}$ is obtained by the resampling scheme and that $\lim_{n\to\infty} N(n) = \infty$, then $\min_{\boldsymbol{\alpha}\in D_n^{N(n)}} \|\boldsymbol{\alpha}\| = o_p(1)$.*

In order to get the rate of convergence we follow the techniques introduced by Kim and Pollard (1990) which allows them to get cube root asymptotics. They derived limit theorems for several statistics defined by maximization or constrained minimization of processes derived from the empirical measure, such as the LMS estimator with rate $n^{-1/3}$. We require a similar result to Lemma 4.1 of Kim and Pollard (1990) which was generalized by Davies (1990) in Theorem 7.

Lemma 7 *Assume that $E(\|\mathbf{x}\|) < \infty$ and $\rho^*(t) = 1 - \rho(t)$. Then, for each $\eta > 0$, there exist random variables $\{M_n\}_1^\infty$ of order $O_p(1)$ such that*

(a) if $\rho = \rho_J$, then

$$\left| E_{H_n}\left(\rho^*\left(\frac{r(\boldsymbol{\theta})}{s}\right) - \rho^*\left(\frac{r(\mathbf{0})}{s}\right)\right) - E_{H_o}\left(\rho^*\left(\frac{r(\boldsymbol{\theta})}{s}\right) - \rho^*\left(\frac{r(\mathbf{0})}{s}\right)\right)\right| \leq \eta(\|\boldsymbol{\theta}\|^2 + |s-1|^2) + n^{-2/3}M_n^2$$

for all $(\boldsymbol{\theta}, s)$ satisfying $\|\boldsymbol{\theta}\| < 1$, $|s-1| < 1$.

*(b) if ρ verifies conditions (**A1**)-(**A5**), then*

$$\left| E_{H_n}\left(\rho^*\left(\frac{r(\boldsymbol{\theta})}{s}\right) - \rho^*\left(\frac{r(\mathbf{0})}{s}\right)\right) - E_{H_o}\left(\rho^*\left(\frac{r(\boldsymbol{\theta})}{s}\right) - \rho^*\left(\frac{r(\mathbf{0})}{s}\right)\right)\right| \leq \eta(\|\boldsymbol{\theta}\|^2 + |s-1|^2) + n^{-1}M_n^2 \quad .$$

for all $(\boldsymbol{\theta}, s)$ satisfying $\|\boldsymbol{\theta}\| < 1$, $|s-1| < 1$.

*(c) if ρ verifies conditions (**A1**)-(**A5**), then*

$$\left| E_{H_n}\left(\rho^*\left(\frac{r(\mathbf{0})}{s}\right) - \rho^*(r(\mathbf{0}))\right) - E_{H_o}\left(\rho^*\left(\frac{r(\boldsymbol{\theta})}{s}\right) - \rho^*(r(\mathbf{0}))\right)\right| \leq \eta|s-1|^2 + n^{-1}M_n^2$$

if $|s-1| < \delta < 1$.

Proof: The proof of (a) is given in Lemma 4 of Davies (1990). The proofs of (b) and (c) follow analogously with some slight modifications. □

The following theorem states the rate of convergence of the resampled τ–estimates under different assumptions on the ρ's functions.

Theorem 4 *Suppose that $(\mathbf{x}_1', y_1), \ldots, (\mathbf{x}_n', y_n)$ are $(p+1)$-dimensional random vectors i.i.d. such that satisfy model (1) and $\mathbf{T}_n = \mathbf{T}_n(\mathbf{Z}_n)$. Assume that ρ_1 and ρ_2 satisfy (**A1**)-(**A4**), (**F1**)-(**F7**) hold, $E(\|\mathbf{x}\|^2) < \infty$ and that the conclusions of Theorem 3 remain valid. Furthermore, we suppose that $n^{1/2} \min_{\boldsymbol{\alpha} \in D_n^{N(n)}} \|\boldsymbol{\alpha}\| = O_p(1)$.*

*(a) If ρ_1 satisfies (**A5**) and ρ_2 (**A5**) and (**A6**), then*
$$n^{1/2}\|\mathbf{T}_n\| = O_p(1).$$

(b) If $\rho_1 = \rho_2 = \rho_J$, then
$$n^{1/3}\|\mathbf{T}_n\| = O_p(1).$$

*(c) If $\rho_1 = \rho_J$ and ρ_2 satisfies (**A5**) and (**A6**), then*
$$n^{1/3}\|\mathbf{T}_n\| = O_p(1).$$

Proof: (a) Let $\boldsymbol{\alpha}_n = \arg \min_{\boldsymbol{\alpha} \in D_n^{N(n)}} \|\boldsymbol{\alpha}\|$. It holds that

$$0 \le \tau^2(\boldsymbol{\alpha}_n) - \tau^2(\mathbf{T}_n) = s_n^2(\boldsymbol{\alpha}_n) E_{H_n}\rho_2\left(\frac{r(\boldsymbol{\alpha}_n)}{s_n(\boldsymbol{\alpha}_n)}\right) - s_n^2(\mathbf{T}_n) E_{H_n}\rho_2\left(\frac{r(\mathbf{T}_n)}{s_n(\mathbf{T}_n)}\right). \tag{9}$$

For any $\boldsymbol{\theta} \in \Re^p$ we have that

$$\begin{aligned}
E_{H_n}\rho_2\left(\frac{r(\boldsymbol{\theta})}{s_n(\boldsymbol{\theta})}\right) = &\left[\left(E_{H_n}\rho_2\left(\frac{r(\boldsymbol{\theta})}{s_n(\boldsymbol{\theta})}\right) - E_{H_n}\rho_2\left(\frac{r(\mathbf{0})}{s_n(\boldsymbol{\theta})}\right)\right) - \right.\\
&\left.\left(E_{H_o}\rho_2\left(\frac{r(\boldsymbol{\theta})}{s_n(\boldsymbol{\theta})}\right) - E_{H_o}\rho_2\left(\frac{r(\mathbf{0})}{s_n(\boldsymbol{\theta})}\right)\right)\right] +\\
&\left[\left(E_{H_n}\rho_2\left(\frac{r(\mathbf{0})}{s_n(\boldsymbol{\theta})}\right) - E_{H_n}\rho_2(r(\mathbf{0}))\right) - \left(E_{H_o}\rho_2\left(\frac{r(\mathbf{0})}{s_n(\boldsymbol{\theta})}\right) - E_{H_o}\rho_2(r(\mathbf{0}))\right)\right] +\\
&\left[\left(E_{H_o}\rho_2\left(\frac{r(\boldsymbol{\theta})}{s_n(\boldsymbol{\theta})}\right) - E_{H_o}\rho_2\left(\frac{r(\mathbf{0})}{s_n(\boldsymbol{\theta})}\right)\right)\right] + \left[\left(E_{H_n}\rho_2(r(\mathbf{0})) - E_{H_o}\rho_2(r(\mathbf{0}))\right)\right] +\\
&E_{H_o}\rho_2\left(\frac{r(\mathbf{0})}{s_n(\boldsymbol{\theta})}\right).
\end{aligned} \tag{10}$$

From (9) and (10) for $\rho^*(t) = 1 - \rho(t)$ we get that

$$\begin{aligned}
&0 \le \tau^2(\boldsymbol{\alpha}_n) - \tau^2(\mathbf{T}_n)\\
&\le s_n^2(\boldsymbol{\alpha}_n) E_{H_n}\rho_2\left(\frac{r(\boldsymbol{\alpha}_n)}{s_n(\boldsymbol{\alpha}_n)}\right) - s_n^2(\mathbf{T}_n) E_{H_n}\rho_2\left(\frac{r(\mathbf{T}_n)}{s_n(\mathbf{T}_n)}\right) =\\
&s_n^2(\boldsymbol{\alpha}_n)\left\{-\left[\left(E_{H_n}\rho_2^*\left(\frac{r(\boldsymbol{\alpha}_n)}{s_n(\boldsymbol{\alpha}_n)}\right) - E_{H_n}\rho_2^*\left(\frac{r(\mathbf{0})}{s_n(\boldsymbol{\alpha}_n)}\right)\right) - \right.\right.\\
&\left.\left(E_{H_o}\rho_2^*\left(\frac{r(\boldsymbol{\alpha}_n)}{s_n(\boldsymbol{\alpha}_n)}\right) - E_{H_o}\rho_2^*\left(\frac{r(\mathbf{0})}{s_n(\boldsymbol{\alpha}_n)}\right)\right)\right] -\\
&\left[\left(E_{H_n}\rho_2^*\left(\frac{r(\mathbf{0})}{s_n(\boldsymbol{\alpha}_n)}\right) - E_{H_n}\rho_2^*(r(\mathbf{0}))\right) - \left(E_{H_o}\rho_2^*\left(\frac{r(\mathbf{0})}{s_n(\boldsymbol{\alpha}_n)}\right) - E_{H_o}\rho_2^*(r(\mathbf{0}))\right)\right] -\\
&\left.\left(E_{H_o}\rho_2^*\left(\frac{r(\boldsymbol{\alpha}_n)}{s_n(\boldsymbol{\alpha}_n)}\right) - E_{H_o}\rho_2^*\left(\frac{r(\mathbf{0})}{s_n(\boldsymbol{\alpha}_n)}\right)\right) - \left(E_{H_n}\rho_2^*(r(\mathbf{0})) - E_{H_o}\rho_2^*(r(\mathbf{0}))\right)\right\} +\\
&s_n^2(\boldsymbol{\alpha}_n) E_{H_o}\rho_2\left(\frac{r(\mathbf{0})}{s_n(\boldsymbol{\alpha}_n)}\right) - s_n^2(\mathbf{T}_n)\left\{-\left[\left(E_{H_n}\rho_2^*\left(\frac{r(\mathbf{T}_n)}{s_n(\mathbf{T}_n)}\right) - E_{H_n}\rho_2^*\left(\frac{r(\mathbf{0})}{s_n(\mathbf{T}_n)}\right)\right) - \right.\right.\\
&\left.\left(E_{H_o}\rho_2^*\left(\frac{r(\mathbf{T}_n)}{s_n(\mathbf{T}_n)}\right) - E_{H_o}\rho_2^*\left(\frac{r(\mathbf{0})}{s_n(\mathbf{T}_n)}\right)\right)\right] -\\
&\left[\left(E_{H_n}\rho_2^*\left(\frac{r(\mathbf{0})}{s_n(\mathbf{T}_n)}\right) - E_{H_n}\rho_2^*(r(\mathbf{0}))\right) - \left(E_{H_o}\rho_2^*\left(\frac{r(\mathbf{0})}{s_n(\mathbf{T}_n)}\right) - E_{H_o}\rho_2^*(r(\mathbf{0}))\right)\right] -\\
&\left.\left(E_{H_o}\rho_2^*\left(\frac{r(\mathbf{T}_n)}{s_n(\mathbf{T}_n)}\right) - E_{H_o}\rho_2^*\left(\frac{r(\mathbf{0})}{s_n(\mathbf{T}_n)}\right)\right) - \left(E_{H_n}\rho_2^*(r(\mathbf{0})) - E_{H_o}\rho_2^*(r(\mathbf{0}))\right)\right\} -\\
&s_n^2(\mathbf{T}_n) E_{H_o}\rho_2\left(\frac{r(\mathbf{0})}{s_n(\mathbf{T}_n)}\right)
\end{aligned}$$

Define the set $A_n = \{s_n(\boldsymbol{\alpha}_n) \le s_n(\mathbf{T}_n)\}$. By **(A5)** and **(A6)** we have that $s^2 E_{H_o}\rho_2(r(\mathbf{0})/s)$ is non–decreasing in s, so in A_n

$$s_n^2(\boldsymbol{\alpha}_n) E_{H_o}\rho_2\left(\frac{r(\mathbf{0})}{s_n(\boldsymbol{\alpha}_n)}\right) - s_n^2(\mathbf{T}_n) E_{H_o}\rho_2\left(\frac{r(\mathbf{0})}{s_n(\mathbf{T}_n)}\right) \le 0.$$

On the other hand,

$$-s_n^2(\alpha_n)(E_{H_n}\rho_2^*(r(0)) - E_{H_o}\rho_2^*(r(0))) + s_n^2(\mathbf{T}_n)(E_{H_n}\rho_2^*(r(0)) - E_{H_o}\rho_2^*(r(0))) =$$
$$(s_n^2(\mathbf{T}_n) - s_n^2(\alpha_n))(E_{H_n}\rho_2^*(r(0)) - E_{H_o}\rho_2^*(r(0))) = n^{-1}O_p(1)$$

Hence, by Lemma 7 (b) and (c) in A_n we obtain the following bound

$$\begin{aligned}
&0 \le \tau^2(\alpha_n) - \tau^2(\mathbf{T}_n) \\
&\le s_n(\alpha_n)^2[\eta_1(\|\alpha_n\|^2 + |s_n(\alpha_n) - 1|^2) + n^{-1}M_{1n} + \eta_2|s_n(\alpha_n) - 1|^2) + n^{-1}M_{2n}] \\
&-s_n(\alpha_n)^2\left[E_{H_o}\rho_2^*\left(\frac{r(\alpha_n)}{s_n(\alpha_n)}\right) - E_{H_o}\rho_2^*\left(\frac{r(0)}{s_n(\alpha_n)}\right)\right] \\
&+s_n(\mathbf{T}_n)^2[\eta_3(\|\mathbf{T}_n\|^2 + |s_n(\mathbf{T}_n) - 1|^2) + n^{-1}M_{3n} + \eta_4|s_n(\mathbf{T}_n) - 1|^2) + n^{-1}M_{4n}] \\
&+s_n(\mathbf{T}_n)^2\left[E_{H_o}\rho_2^*\left(\frac{r(\mathbf{T}_n)}{s_n(\mathbf{T}_n)}\right) - E_{H_o}\rho_2^*\left(\frac{r(0)}{s_n(\mathbf{T}_n)}\right)\right] + n^{-1}O_p(1)
\end{aligned} \tag{11}$$

On the other hand, if $s_n(\alpha_n) > s_n(\mathbf{T}_n)$, i.e. in A_n^c, we have that

$$0 \le E_{H_n}\rho_1\left(\frac{r(\alpha_n)}{s_n(\mathbf{T}_n)}\right) - E_{H_n}\rho_1\left(\frac{r(\mathbf{T}_n)}{s_n(\mathbf{T}_n)}\right), \tag{12}$$

hence in A_n^c we get that

$$\begin{aligned}
&0 \le -\left[E_{H_n}\rho_1^*\left(\frac{r(\alpha_n)}{s_n(\mathbf{T}_n)}\right) - E_{H_n}\rho_1^*\left(\frac{r(0)}{s_n(\mathbf{T}_n)}\right)\right] - \left[E_{H_n}\rho_1^*\left(\frac{r(\mathbf{T}_n)}{s_n(\mathbf{T}_n)}\right) - E_{H_n}\rho_1^*\left(\frac{r(0)}{s_n(\mathbf{T}_n)}\right)\right] = \\
&-\left[\left(E_{H_n}\rho_1^*\left(\frac{r(\alpha_n)}{s_n(\mathbf{T}_n)}\right) - E_{H_n}\rho_1^*\left(\frac{r(0)}{s_n(\mathbf{T}_n)}\right)\right) - \left(E_{H_o}\rho_1^*\left(\frac{r(\alpha_n)}{s_n(\mathbf{T}_n)}\right) - E_{H_o}\rho_1^*\left(\frac{r(0)}{s_n(\mathbf{T}_n)}\right)\right)\right] \\
&+\left[\left(E_{H_n}\rho_1^*\left(\frac{r(\mathbf{T}_n)}{s_n(\mathbf{T}_n)}\right) - E_{H_n}\rho_1^*\left(\frac{r(0)}{s_n(\mathbf{T}_n)}\right)\right) - \left(E_{H_o}\rho_1^*\left(\frac{r(\mathbf{T}_n)}{s_n(\mathbf{T}_n)}\right) - E_{H_o}\rho_1^*\left(\frac{r(0)}{s_n(\mathbf{T}_n)}\right)\right)\right] \\
&-\left[E_{H_o}\rho_1^*\left(\frac{r(\alpha_n)}{s_n(\mathbf{T}_n)}\right) - E_{H_o}\rho_1^*\left(\frac{r(0)}{s_n(\mathbf{T}_n)}\right)\right] + \left[E_{H_o}\rho_1^*\left(\frac{r(\mathbf{T}_n)}{s_n(\mathbf{T}_n)}\right) - E_{H_o}\rho_1^*\left(\frac{r(0)}{s_n(\mathbf{T}_n)}\right)\right].
\end{aligned}$$

Therefore, in A_n^c, from Lemma 7 we get that

$$\begin{aligned}
&0 \le E_{H_n}\rho_1\left(\frac{r(\alpha_n)}{s_n(\mathbf{T}_n)}\right) - E_{H_n}\rho_1\left(\frac{r(\mathbf{T}_n)}{s_n(\mathbf{T}_n)}\right) \le \\
&\eta_5(\|\alpha_n\|^2 + |s_n(\mathbf{T}_n) - 1|^2) + n^{-1}M_{5n} + \eta_6(\|\mathbf{T}_n\|^2 + |s_n(\mathbf{T}_n) - 1|^2) + n^{-1}M_{6n} \\
&-\left[E_{H_o}\rho_1^*\left(\frac{r(\alpha_n)}{s_n(\mathbf{T}_n)}\right) - E_{H_o}\rho_1^*\left(\frac{r(0)}{s_n(\mathbf{T}_n)}\right)\right] + \left[E_{H_o}\rho_1^*\left(\frac{r(\mathbf{T}_n)}{s_n(\mathbf{T}_n)}\right) - E_{H_o}\rho_1^*\left(\frac{r(0)}{s_n(\mathbf{T}_n)}\right)\right]
\end{aligned} \tag{13}$$

Let $\xi_1 = \max(\eta_1, \eta_2, \eta_5)$ and $\xi_2 = \max(\eta_3, \eta_4, \eta_6)$. From (11) and (13), we obtain that

$$\begin{aligned}
0 \le\ &\Big\{s_n(\alpha_n)^2[\xi_1\|\alpha_n\|^2 + 2\xi_1|s_n(\alpha_n) - 1|^2 + n^{-1}(M_{1n} + M_{2n})] + \\
&s_n(\mathbf{T}_n)^2[\xi_2\|\mathbf{T}_n\|^2 + 2\xi_2|s_n(\mathbf{T}_n) - 1|^2 + n^{-1}(M_{3n} + M_{4n})] - \\
&s_n(\alpha_n)^2\left[E_{H_o}\rho_2^*\left(\frac{r(\alpha_n)}{s_n(\alpha_n)}\right) - E_{H_o}\rho_2^*\left(\frac{r(0)}{s_n(\alpha_n)}\right)\right] + \\
&s_n(\mathbf{T}_n)^2\left[E_{H_o}\rho_2^*\left(\frac{r(\mathbf{T}_n)}{s_n(\mathbf{T}_n)}\right) - E_{H_o}\rho_2^*\left(\frac{r(0)}{s_n(\mathbf{T}_n)}\right)\right] + n^{-1}O_p(1)\Big\} I_{A_n} + \\
&\{\ \xi_1\|\alpha_n\|^2 + \xi_1|s_n(\mathbf{T}_n) - 1|^2 + \xi_2\|\mathbf{T}_n\|^2 + \xi_2|s_n(\mathbf{T}_n) - 1|^2 + n^{-1}(M_{5n} + M_{6n})] - \\
&\left[E_{H_o}\rho_1^*\left(\frac{r(\alpha_n)}{s_n(\alpha_n)}\right) - E_{H_o}\rho_1^*\left(\frac{r(0)}{s_n(\alpha_n)}\right)\right] + \\
&\left[E_{H_o}\rho_1^*\left(\frac{r(\mathbf{T}_n)}{s_n(\mathbf{T}_n)}\right) - E_{H_o}\rho_1^*\left(\frac{r(0)}{s_n(\mathbf{T}_n)}\right)\right]\Big\} I_{A_n^c}.
\end{aligned} \tag{14}$$

Using a Taylor expansion we get that

$$\left[E_{H_o}\rho_2^*\left(\frac{r(\boldsymbol{\alpha}_n)}{s_n(\boldsymbol{\alpha}_n)}\right) - E_{H_o}\rho_2^*\left(\frac{r(\mathbf{0})}{s_n(\boldsymbol{\alpha}_n)}\right)\right] = -\int_0^c f_o'(t)d\rho_2^*(t)\boldsymbol{\alpha}_n'\mathbf{Q}\boldsymbol{\alpha}_n + o_p(\|\boldsymbol{\alpha}_n\|^2),$$

where $\mathbf{Q} = E(\mathbf{xx}')$. For the last three terms in (14), a similar approximation holds. Therefore, we get the following bound

$$\begin{aligned}
&0 \le \|\boldsymbol{\alpha}_n\|^2\xi_1(s_n^2(\boldsymbol{\alpha}_n)I_{A_n} + I_{A_n^c}) + \xi_1|s_n(\boldsymbol{\alpha}_n) - 1|^2 2s_n^2(\boldsymbol{\alpha}_n)I_{A_n} + \\
&\|\mathbf{T}_n\|^2\xi_2(s_n^2(\mathbf{T}_n)I_{A_n} + I_{A_n^c}) + |s_n(\mathbf{T}_n) - 1|^2(2s_n^2(\mathbf{T}_n)\xi_2 I_{A_n} + (\xi_1+\xi_2)I_{A_n^c}) + \\
&n^{-1}\{[s_n^2(\boldsymbol{\alpha}_n)(M_{1n} + M_{2n}) + s_n^2(\mathbf{T}_n)(M_{3n} + M_{4n})]I_{A_n} + (M_{5n} + M_{6n})I_{A_n^c}\} \\
(15)\qquad &+ \left[s_n^2(\boldsymbol{\alpha}_n)\int_0^c f_o'(t)d\rho_2^*(t)I_{A_n} + \int_0^c f_o'(t)d\rho_1^*(t)I_{A_n^c}\right]\boldsymbol{\alpha}_n'\mathbf{Q}\boldsymbol{\alpha}_n \\
&- \left[s_n^2(\mathbf{T}_n)\int_0^c f_o'(t)d\rho_2^*(t)I_{A_n} + \int_0^c f_o'(t)d\rho_1^*(t)I_{A_n^c}\right]\mathbf{T}_n'\mathbf{Q}\mathbf{T}_n \\
&+ o_p(\|\boldsymbol{\alpha}_n\|^2) + o_p(\|\mathbf{T}_n\|^2) + n^{-1}O_p(1).
\end{aligned}$$

Define $m_1 = \min(-\int_0^c f_o'(t)d\rho_2^*(t), -\int_0^c f_o'(t)d\rho_1^*(t))$, $M_1 = \max(-\int_0^c f_o'(t)d\rho_2^*(t), -\int_0^c f_o'(t)d\rho_1^*(t))$. Using the fact that $\|\mathbf{x}\|_Q = \mathbf{x}'Q\mathbf{x}$ is equivalent to the Euclidean norm and (15) we may choose positive constants l and L such that

$$\begin{aligned}
&\|T_n\|^2(s_n^2(\mathbf{T}_n)I_{A_n} + I_{A_n^c})(-\xi_2 - M_1 L) \le \\
&\|\boldsymbol{\alpha}_n\|^2\xi_1(s_n^2(\boldsymbol{\alpha}_n)I_{A_n} + I_{A_n^c}) + 2\xi_1|s_n(\boldsymbol{\alpha}_n) - 1|^2 s_n^2(\boldsymbol{\alpha}_n)I_{A_n} \\
&+|s_n(\mathbf{T}_n) - 1|^2(2\xi_2 s_n^2(\boldsymbol{\alpha}_n)I_{A_n} + (\xi_1+\xi_2)I_{A_n^c}) \\
&n^{-1}\{[s_n^2(\boldsymbol{\alpha}_n)(M_{1n} + M_{2n}) + s_n^2(\mathbf{T}_n)(M_{3n} + M_{4n})]I_{A_n} + (M_{5n} + M_{6n})I_{A_n^c}\} \\
&-m_1 l(s_n^2(\boldsymbol{\alpha}_n)I_{A_n} + I_{A_n^c})\|\boldsymbol{\alpha}_n\|^2 + o_p(\|\boldsymbol{\alpha}_n\|^2) + o_p(\|\mathbf{T}_n\|^2) + n^{-1}O_p(1)
\end{aligned}$$

Since $-M_1$ is positive, by taking ξ_2 small enough, part (a) of the theorem follows.

(b) It follows from inequality (12) using Lemma 7 a).

(c) Parts (a) and (b) can be proved similarly using Lemma 7 a) and b). □

Remark 3 *We conjecture that when $\rho_1 = \rho_J$ and ρ_2 satisfies* (**A5**) *and* (**A6**) *the order of the convergence of $\|\mathbf{T}_n\|$ may be higher to the one established by Theorem 4.*

Remark 4 *The previous theorem assumes that $n^{1/2}\min_{\boldsymbol{\alpha}\in D_n^{N(n)}}\|\boldsymbol{\alpha}\| = O_p(1)$. This may be guaranteed by a slight modification in the resampling scheme, that is by including in the set D_n^N an estimator of order $n^{-1/2}$.*

4 Appendix

Proof of Lemma 1:

(a) The result will follow by an approximating argument. First, assume that $\rho^*(t)$ is a step function in $[-b,b]$,that is

$$\rho^*(|t|) = \rho_l^*(|t|) = \sum_{i=1}^{l} a_i I_{E_i}(|t|) = \sum_{i=1}^{l-1}(a_i - a_{i+1})I_{[0,b_i]}(|t|) + a_l, \qquad (16)$$

where $0 = b_o < b_1 < \ldots < b_l = b$ and $E_i = (b_{i-1}, b_i]$ para $i = 2, \ldots, l$ and $E_1 = [b_o, b_1]$. From Beran and Millar (1986) it follows that $\sup_{\|\gamma\|=1} U_n(\gamma) = O_p(1)$. Then, we have that

$$
\begin{aligned}
& n^{1/2} \sup_{\theta,s,b} \left| E_{H_n} I_{[-b,b]}\left(\frac{r(\theta)}{s}\right) - E_{H_o} I_{[-b,b]}\left(\frac{r(\theta)}{s}\right) \right| = \\
& n^{1/2} \sup_{\theta,s,b} \left| E_{H_n} I_{[-b,b]}\left(\frac{r(\theta)}{s}\right) - E_{H_o} I_{[-b,b]}\left(\frac{r(\theta)}{s}\right) \right| = \\
& n^{1/2} \sup_{\theta,s,b} \left| P_{H_n}\left(|(y,\mathbf{x})'(1/s,\theta/s)| \le b\right) - P\left(|(y,\mathbf{x})'(1/s,\theta/s)| \le b\right) \right| = \\
& n^{1/2} \sup_{\theta,s,b} \left| P_{H_n}\left((y,\mathbf{x})'(1/s,\theta/s) \le b, -(y,\mathbf{x})'(1/s,\theta/s) \le b\right) - \right. \\
& \left. P\left((y,\mathbf{x})'(1/s,\theta/s) \le b, -(y,\mathbf{x})'(1/s,\theta/s) \le b\right) \right| = \\
& n^{1/2} \sup_{\theta,s,b} \left| P_{H_n}\left((y,\mathbf{x},1)'\frac{(1/s,\theta/s,-b)}{\|(1/s,\theta/s,-b)\|} \le 0, (y,\mathbf{x},1)'\frac{(-1/s,-\theta/s,-b)}{\|(-1/s,-\theta/s,-b)\|} \le 0\right) - \right. \\
& \left. P\left((y,\mathbf{x},1)'\frac{(1/s,\theta/s,-b)}{\|(1/s,\theta/s,-b)\|} \le 0, (y,\mathbf{x},1)'\frac{(-1/s,-\theta/s,-b)}{\|(-1/s,-\theta/s,-b)\|} \le 0\right) \right| = O_p(1)
\end{aligned}
$$

Then,

$$
\begin{aligned}
& n^{1/2} \sup_{\theta,s} \left| E_{H_n}\left(\sum_{i=1}^{l-1}(a_i - a_{i+1}) I_{[0,b_i]}\left(\frac{|r(\theta)|}{s}\right)\right) - E_{H_o}\left(\sum_{i=1}^{l-1}(a_i - a_{i+1}) I_{[0,b_i]}\left(\frac{|r(\theta)|}{s}\right)\right) \right| \le \\
& n^{1/2} \sup_{\theta,s} \sum_{i=1}^{l-1} |(a_i - a_{i+1})| \left| E_{H_n}\left(I_{[0,b_i]}\left(\frac{|r(\theta)|}{s}\right)\right) - E_{H_o}\left(I_{[0,b_i]}\left(\frac{|r(\theta)|}{s}\right)\right) \right| \le \\
& \sup_{\theta,s} \sum_{i=1}^{l-1} |(a_i - a_{i+1})| \left\{ n^{1/2} \sup_{\theta,s,b\in[0,c]} \left| E_{H_n}\left(I_{[0,b]}\left(\frac{|r(\theta)|}{s}\right)\right) - E_{H_o}\left(I_{[0,b]}\left(\frac{|r(\theta)|}{s}\right)\right) \right| \right\} = \\
& \sum_{i=1}^{l-1} |(a_i - a_{i+1})| O_p(1) = O_p(1),
\end{aligned}
$$

where $O_p(1)$ is independent of l.

Let ρ be a function satisfying (**A1**)-(**A4**). Consider a sequence of step functions $\{\rho_l^*\}$ such that uniformly converges to ρ^*. Let l be such that $|\rho^*(|t|) - \rho_l^*(|t|)| < n^{-1/2}\epsilon$, then

$$
\begin{aligned}
& n^{1/2} \sup_{\theta,s,b} \left| E_{H_n}\rho^*\left(\frac{r(\theta)}{s}\right) - E_{H_o}\rho^*\left(\frac{r(\theta)}{s}\right) \right| = \\
& n^{1/2} \sup_{\theta,s,b} \left| E_{H_n}\left(\rho^*\left(\frac{r(\theta)}{s}\right) - \rho_l^*\left(\frac{r(\theta)}{s}\right)\right) - E_{H_o}\left(\rho^*\left(\frac{r(\theta)}{s}\right) - \rho_l^*\left(\frac{r(\theta)}{s}\right)\right) + \right. \\
& \left. E_{H_n}\rho_l^*\left(\frac{r(\theta)}{s}\right) - E_{H_o}\rho_l^*\left(\frac{r(\theta)}{s}\right) \right| \le \\
& n^{1/2} \left\{ \sup_{\theta,s} 2n^{-1/2}\epsilon + \left| E_{H_n}\rho_l^*\left(\frac{r(\theta)}{s}\right) - E_{H_o}\rho_l^*\left(\frac{r(\theta)}{s}\right) \right| \right\} = O_p(1)
\end{aligned}
$$

and the statement follows.

(b) It follows equally than (a) by using an approximation argument. □

Proof of Theorem 2 First, we will show that

$$\int \rho_1^*(y)(sf_o(sy) - f_o(y))dy > K_1(s-1) \text{ for } s > 1 \tag{17}$$

$$\int \rho_1^*(y)(sf_o(sy) - f_o(y))dy < K_2(s-1) \text{ for } s < 1 \tag{18}$$

For $s > 1$, inequality (17) is equivalent to $\int \rho_1^*(y)\left(\frac{sf_o(sy) - f_o(y)}{s-1}\right) dy > K_1$.

Let ρ_l^* be as in (16). If we take $a_i = \rho_1^*(b_i)$, we have that $\rho_l^*(t) \leq \rho_1^*(t)$. Thus, there exists $\delta > 0$ such that for any $l > l_o$ we get that

$$\begin{aligned}
&\int_{-c}^{c} \rho_1^*(y)\left(\frac{sf_o(sy) - f_o(y)}{s-1}\right) dy = \int_{-c}^{c} (\rho_1^*(y) - \rho_l^*(y))\left(\frac{sf_o(sy) - f_o(y)}{s-1}\right) dy + \\
&\int_{-c}^{c} \rho_l^*(y)\left(\frac{sf_o(sy) - f_o(y)}{s-1}\right) dy = \sum_{i=1}^{l-1}(a_i - a_{i+1})f_o(b_i)b_i + \int_{-c}^{c}(\rho_1^*(y) - \rho_l^*(y))f_o(sy)dy - \\
&2\int_0^c (\rho_1^*(y) - \rho_l^*(y))\left(\frac{f_o(y) - f_o(sy)}{s-1}\right) dy > \\
&\sum_{i=1}^{l-1}(a_i - a_{i+1})f_o(b_i)b_i - 2\epsilon\int_0^c \left(\frac{f_o(y) - f_o(sy)}{s-1}\right) dy = \\
&\sum_{i=1}^{l-1}(a_i - a_{i+1})f_o(b_i)b_i - \epsilon\left[\frac{1}{s-1}\left((F_o(c) - F_o(-c)) - \frac{1}{s}(F_o(sc) - F_o(-cs))\right)\right] = \\
&\sum_{i=1}^{l-1}(a_i - a_{i+1})f_o(b_i)b_i - \epsilon\left[\frac{1}{s-1}\left(\left(\frac{1}{s}F_o(-cs) - F_o(-c))\right) - \left(\frac{1}{s}F_o(sc) - F_o(c)\right)\right)\right] > \\
&-\epsilon(-2cf_o(c) + F_o(c) - F_o(-c)) - \int_0^c f_o(t)t d\rho_1^*(t) - \delta > K_1
\end{aligned}$$

and (17) follows choosing a suitable positive constant K_1. Inequality (18) entails analogously.

(a) (i) Since

$$\begin{aligned}
&E_{H_o}\rho_1^*\left(\frac{r(0)}{s}\right) = \int \rho_1^*\left(\frac{y}{s}\right) f_o(y)dy = \\
&\int \rho_1^*(y)sf_o(sy)dy - \int \rho_1^*(y)f_o(y)dy + \int \rho_1^*(y)f_o(y)dy = \\
&b + \int \rho_1^*(y)(sf_o(sy) - f_o(y))dy,
\end{aligned}$$

the statement holds by (17) and Lemma 2 (i).

(ii) By Lemma 2 (ii) we know that $\sup_{\theta\in\Re^p} E_{H_o}\rho_1^*\left(\frac{r(\theta)}{s}\right) = E_{H_o}\rho_1^*\left(\frac{r(0)}{s}\right)$. Then, it is enough to prove that $E_{H_o}\rho_1^*\left(\frac{r(0)}{s}\right) < b - K_1(1-s)$. Then, we proceed as in (a) (i).

(b) By Lemma 1 we have that

$$h_n = \sup_{\theta\in\Re^p, s\in\Re^+}\left|E_{H_n}\rho_1^*\left(\frac{r(\theta)}{s}\right) - E_{H_o}\rho_1^*\left(\frac{r(\theta)}{s}\right)\right| = O_p(n^{-1/2})$$

Then,

$$E_{H_n}\rho_1^*\left(\frac{r(\theta_n)}{s}\right) = E_{H_n}\rho_1^*\left(\frac{r(\theta_n)}{s}\right) - E_{H_o}\rho_1^*\left(\frac{r(\theta_n)}{s}\right) + E_{H_o}\rho_1^*\left(\frac{r(\theta_n)}{s}\right) \le$$
$$\sup_{\theta\in\Re^p, s\in\Re^+}\left|E_{H_n}\rho_1^*\left(\frac{r(\theta)}{s}\right) - E_{H_o}\rho_1^*\left(\frac{r(\theta)}{s}\right)\right| + E_{H_o}\rho_1^*\left(\frac{r(\theta_n)}{s}\right) = h_n + E_{H_o}\rho_1^*\left(\frac{r(\theta_n)}{s}\right)$$

Let $\tilde{s}_n = 1 - h_n/K_2$. Then

$$E_{H_n}\rho_1^*\left(\frac{r(\theta_n)}{\tilde{s}_n}\right) \le h_n + E_{H_o}\rho_1^*\left(\frac{r(\theta_n)}{\tilde{s}_n}\right) < b$$

by (a) (ii). Hence, $s_n(\theta_n) > 1 - h_n/K_2$. On the other hand, let us take $\tilde{s}_n = 1 + h_n/K_1$. Given $\epsilon > 0$ and $\delta > 0$, the set $D_n = \{\|\theta_n\| > \delta\}$ verifies that $P(D_n) < \epsilon$ for n large enough. Then using item (a) (i) it holds in D_n that

$$E_{H_n}\rho_1^*\left(\frac{r(\theta_n)}{\tilde{s}_n}\right) \ge b + h_n \ge b$$

and then, $s_n(\theta_n) < 1 + h_n/K_1$. Finally, it entails that, in D_n

$$1 - \frac{h_n}{K_2} \le s_n(\theta_n) \le 1 + \frac{h_n}{K_1}$$

and b) follows. □

Proof of Lemma 4

(i) There exists $h_o > 0$ such that for all $\lambda \in \mathcal{S}_p$ it holds that $P_{G_o}(N_{\lambda^\perp}) \le 1/2 - h_o$ where $N_{\lambda^\perp}$ was defined above. Therefore, $P_{G_o}((N_{\lambda^\perp})^c) \ge b + h_o$. Let $\epsilon > 0$ be such that $(1-\epsilon)(b + h_o/2) > b + h_o/4$. We take $A_M = \{y \in [0, M]\}$ satisfying $P_{H_o}(A_M) > 1 - \epsilon$. Given $\epsilon_2 > 0$ there exists a compact set $K_\lambda \subset (N_{\lambda^\perp})^c \subset \Re^p$ such that $|\mathbf{x}'\lambda| \ge 3\epsilon_2$ and $P_{G_o}(K_\lambda) \ge b + h_o/2$. Hence we can choose $\delta_2 = \delta_2(\lambda, \epsilon_2)$ such that $\|\theta/\|\theta\| - \lambda\| < \delta_2$ and $\mathbf{x} \in K_\lambda$ imply that $|\mathbf{x}'\theta/\|\theta\|| \ge 2\epsilon_2$. Let $m_2 > 0$ be such that $|y|/m_2 < \epsilon_2$ if $y \in A_M$. Denote by $V(\lambda, \delta) = \{\|\theta\| > m_1 s, \|\theta/\|\theta\| - \lambda\| \le \delta, \|\theta\| > m_2\}$. If $y \in A_M$, $\mathbf{x} \in K_\lambda$ and $\theta \in V((\lambda, \delta))$, then

$$0 < \epsilon < \left|\mathbf{x}'\frac{\theta}{\|\theta\|}\right| - \frac{|y|}{\|\theta\|} < \left|\mathbf{x}'\frac{\theta}{\|\theta\|} - \frac{y}{\|\theta\|}\right|.$$

Since ρ_1 is an even function we obtain

$$\inf_{V(\lambda,\delta_2)} E_{H_n}\rho_1\left(\frac{r(\theta)}{s}\right) = \inf_{V(\lambda,\delta_2)} E_{H_n}\rho_1\left(\left|\frac{r(\theta)}{s}\right|\right) \ge$$
$$E_{H_n}\inf_{V(\lambda,\delta_2)}\rho_1\left(\frac{\|\theta\|}{s}\left|\frac{y}{\|\theta\|} - \mathbf{x}'\frac{\theta}{\|\theta\|}\right|\right) \ge$$
$$E_{H_n}\inf_{V(\lambda,\delta_2)}\rho_1\left(m_1\left|\frac{y}{\|\theta\|} - \mathbf{x}'\frac{\theta}{\|\theta\|}\right|\right) I_{A_M}(y) I_{K_\lambda}(\mathbf{x}) \ge$$
$$\rho_1(m_1\epsilon_2)\, E_{H_n} I_{A_M}(y) I_{K_\lambda}(\mathbf{x}).$$

Consequently, it holds almost everywhere that

$$\liminf_{m_1\to\infty}\liminf_{n\to\infty}\rho_1(m_1\epsilon_2)\,E_{H_n}I_{A_M}(y)I_{K_\lambda}(\mathbf{x}) =$$
$$\liminf_{m_1\to\infty}\rho_1(m_1\epsilon_2)\,E_{H_o}I_{A_M}(y)I_{K_{\boldsymbol{\lambda}}}(\mathbf{x}) \geq (b+h_o/2)(1-\epsilon) > b+h_o/4$$

and

$$\liminf_{m_1\to\infty, m_2\to\infty}\ \liminf_{n\to\infty}\ \inf_{V(\boldsymbol{\lambda},\delta_2)} E_{H_n}\rho_1\left(\frac{r(\boldsymbol{\theta})}{s}\right) \geq b+h_o/4.$$

For all $\boldsymbol{\lambda}\in S_p$ there exists $\delta_2(\boldsymbol{\lambda},\epsilon_2)$ verifying the last inequality. Since $\mathcal{S}_p$ is a compact set we consider a finite number of vectors in $\mathcal{S}_p$, $\boldsymbol{\theta}_1,\ldots,\boldsymbol{\theta}_q$ and radius $\delta_2(\boldsymbol{\theta}_1,\epsilon_2),\ldots,\delta_2(\boldsymbol{\theta}_q,\epsilon_2)$ such that

$$\inf_{\|\boldsymbol{\theta}\|>m_1 s, \|\boldsymbol{\theta}\|>m_2} E_{H_n}\rho_1\left(\frac{r(\boldsymbol{\theta})}{s}\right) = \min_{1\leq j\leq q}\ \inf_{V(\boldsymbol{\theta}_j,\delta_2(\boldsymbol{\theta}_j,\epsilon_2))} E_{H_n}\rho_1\left(\frac{r(\boldsymbol{\theta})}{s}\right)$$

and the statement follows.

(ii) Let $\boldsymbol{\lambda}\in B_m=\{\mathbf{z}\in\Re^p:\|\mathbf{z}\|\leq m\}$. Given $\epsilon>0$ we can choose $\delta=\delta(\boldsymbol{\lambda},\epsilon)$ and $c>0$ verifying that $P_{H_o}(A_\delta)=P_{H_o}(|y-\mathbf{x}'\boldsymbol{\lambda}|\leq\delta)\leq\epsilon$ and $P_{G_o}(A_c)=P_{G_o}(\|\mathbf{x}\|\leq c)>1-\epsilon$. If $(\mathbf{x}',y)'\in A_\delta^c$, $\mathbf{x}\in A_c$ and $\|\boldsymbol{\theta}-\boldsymbol{\lambda}\|<\delta/2c$ we get that

$$|y-\mathbf{x}'\boldsymbol{\theta}|\geq|y-\mathbf{x}'\boldsymbol{\lambda}|-|\mathbf{x}'(\boldsymbol{\lambda}-\boldsymbol{\theta})|\geq\delta-\delta/2.$$

Then, since ρ_1 is an even function it holds that

$$\inf_{\|\boldsymbol{\theta}-\boldsymbol{\lambda}\|\leq\delta/2c, s<s_o} E_{H_n}\rho_1\left(\frac{r(\boldsymbol{\theta})}{s}\right)\geq E_{H_n}\inf_{\|\boldsymbol{\theta}-\boldsymbol{\lambda}\|\leq\delta/2c}\rho_1\left(\frac{r(\boldsymbol{\theta})}{s_o}\right)\geq$$
$$E_{H_n}\inf_{\|\boldsymbol{\theta}-\boldsymbol{\lambda}\|\leq\delta/2c}\rho_1\left(\frac{r(\boldsymbol{\theta})}{s_o}\right)I_{A_\delta^c}(y,\mathbf{x})I_{A_c}(\mathbf{x})\geq$$
$$\rho_1\left(\frac{\delta}{2s_o}\right)E_{H_n}I_{A_\delta^c}(y,\mathbf{x})I_{A_c}(\mathbf{x}).$$

Consequently, it holds almost everywhere that

$$\liminf_{s_o\to 0}\liminf_{n\to\infty}\rho_1\left(\frac{\delta}{2s_o}\right)E_{H_n}I_{A_\delta^c}(y,\mathbf{x})I_{A_c}(\mathbf{x})=$$
$$\liminf_{s_o\to 0}\rho_1\left(\frac{\delta}{2s_o}\right)E_{H_o}I_{A_\delta^c}(y,\mathbf{x})I_{A_c}(\mathbf{x})\geq 1-2\epsilon$$

and

$$\liminf_{s_o\to 0}\liminf_{n\to\infty}\inf_{\|\boldsymbol{\theta}-\boldsymbol{\lambda}\|\leq\delta/2c, s>s_o} E_{H_n}\rho_1\left(\frac{r(\boldsymbol{\theta})}{s}\right)=1.$$

A compactness argument similar to that of (i) for the set B_m allows us to conclude that the claim of (ii) holds.

(iii) Since

$$\sup_{s>s_o,\|\boldsymbol{\theta}\|<m} E_{H_n}\rho_1\left(\frac{r(\boldsymbol{\theta})}{s}\right)\leq E_{H_n}\sup_{\|\boldsymbol{\theta}\|<m}\rho_1\left(\frac{r(\boldsymbol{\theta})}{s_o}\right),$$

we have almost everywhere that

$$\limsup_{n\to\infty} \sup_{s>s_o,\|\theta\|<m} E_{H_n}\rho_1\left(\frac{r(\theta)}{s}\right) \le E_{H_o} \sup_{\|\theta\|<m} \rho_1\left(\frac{r(\theta)}{s_o}\right).$$

Given $\epsilon > 0$, we take A_M and A_c as above. Therefore,

$$E_{H_o} \sup_{\|\theta\|<m} \rho_1\left(\frac{r(\theta)}{s_o}\right) \le E_{H_o} \sup_{\|\theta\|<m} \rho_1\left(\frac{r(\theta)}{s_o}\right) I_{A_M}(y) I_{A_c}(\mathbf{x}) + \epsilon.$$

Given $\epsilon_1 > 0$ we can choose $\theta_{y,\mathbf{x},s_o,\epsilon_1} \in B_m$ such that

$$E_{H_o} \sup_{\|\theta\|<m} \rho_1\left(\frac{r(\theta)}{s_o}\right) \le E_{H_o}\rho_1\left(\frac{r(\theta_{y,\mathbf{x},s_o,\epsilon_1})}{s_o}\right) I_{A_M}(y) I_{A_c}(\mathbf{x}) + \epsilon_1 + \epsilon.$$

Thus, by using the Dominated Convergence Theorem we obtain that $\lim_{s_o\to\infty} E_{H_o} \sup_{\|\theta\|<m} \rho_1\left(\frac{r(\theta)}{s_o}\right) = 0$, which entails (iii).

(iv) Given $\epsilon > 0$ and taking A_M and A_c as above, we get that

$$E_{H_n}\rho_1\left(\frac{r(\theta)}{s}\right) \le E_{H_n}\rho_1\left(\frac{r(\theta)}{s}\right) I_{A_M}(y) I_{A_c}(\mathbf{x}) + E_{H_n}\rho_1\left(\frac{r(\theta)}{s}\right) I_{(A_M\times A_c)^c}(y,\mathbf{x}) \le E_{H_n}\rho_1\left(\frac{s_o M}{m} + c s_o\right) + 2\epsilon$$

By taking limit for $n \to \infty$ in both sides of the previous inequality, (iv) follows.

Proof of Lemma 5

(a) The proof follows completely analogous to that of Lemma 4.5 of Yohai and Zamar (1986).

(b) Let us call $R_1(\theta, s) = E_{H_o}\rho_1\left(\frac{r(\theta)}{s}\right)$ and $R_2(\theta, s) = E_{H_o}\rho_2\left(\frac{r(\theta)}{s}\right)$. For each θ, $R_2(\theta, s(\theta)) > 0$ by equation (5) and condition (**F7**). Next, let us suppose that $\lim_{n\to\infty} \|\theta_n\| = \infty$ and $\lim_{n\to\infty} R_2(\theta_n, s(\theta_n)) = 0$. Without loss of generality we can say that $\lim_{n\to\infty} \theta_n/\|\theta_n\| = \theta_1$. If

(19) $$\limsup_{n\to\infty} \|\theta_n\|/s(\theta_n) = \infty,$$

then, by taking a subsequence if it were necessary, we obtain that $P_{G_o}(N^c_{\theta_1^\perp}) = 0$, but this contradicts that $P_{G_o}(N_{\theta_1^\perp}) < 1 - b$. In case that

(20) $$\limsup_{n\to\infty} \|\theta_n\|/s(\theta_n) < \infty$$

we may assume that $\lim_{n\to\infty} \|\theta_n\|/s(\theta_n) = a$. Then $E_{G_o}\rho_2(a\theta_1'\mathbf{x}) = 0$, or equivalently,

(21) $$P_{G_o}(|a\theta_1'\mathbf{x}| \le c_2) = 1.$$

Let us suppose that ρ_1 is continuous . By the Dominated Convergence Theorem we have that $b = E_{G_o}\rho_1(a\theta_1'\mathbf{x})$, but this contradicts (**F7**) and (21). If $\rho_1 = \rho_2 = \rho_J$ equation (5) and (21) together yield a contradiction. Hence, $\inf_{\boldsymbol{\theta}} R_2(\boldsymbol{\theta}, s(\boldsymbol{\theta})) > 0$. Under (III), $\rho_1 = \rho_J$. Hence, let us consider the function

$$\rho_1^r(t) = \begin{cases} 0 \text{ if } |t| \leq 1 \\ |t-1|/r \text{ if } 1 < |t| < 1+r \\ 1 \text{ if } |t| \geq 1+r \end{cases}$$

Let us call $s^r(\boldsymbol{\theta}_n)$ the scale obtained using $\rho_1^r(t)$ in equation (5). It holds that $|s^r(\boldsymbol{\theta}_n) - s(\boldsymbol{\theta}_n)| \leq s(\boldsymbol{\theta}_n) r/(1+r)$, or equivalently

$$\left| \frac{s^r(\boldsymbol{\theta}_n)}{s(\boldsymbol{\theta}_n)} - 1 \right| \leq r/(1+r).$$

Consequently

$$b = E_{H_o} \rho_1^r \left(\frac{r(\boldsymbol{\theta}_n)}{s^r(\boldsymbol{\theta}_n)} \right) \leq E_{H_o} \rho_1^r \left((r+1) \frac{r(\boldsymbol{\theta}_n)}{s(\boldsymbol{\theta}_n)} \right).$$

After taking limit we get that

$$b \leq E_{H_o} \rho_1^r \left((r+1) a \boldsymbol{\theta}_1' \mathbf{x} \right) \leq E_{H_o} \rho_1 \left((r+1) a \boldsymbol{\theta}_1' \mathbf{x} \right).$$

Therefore $P_{G_o}((r+1)|a\boldsymbol{\theta}_1'\mathbf{x}| > 1) \geq b$ for every $r > 0$ which, let us infer that

$$P_{G_o}(|a\boldsymbol{\theta}_1'\mathbf{x}| \geq 1) \geq b. \tag{22}$$

From (**F7**) we get that $\rho_1(c_2) = 0$. Then either $c_2 < 1$, which contradicts (21), or $c_2 = 1$, but (**F8**), (21) and (22) yield a contradiction. Therefore the claim of the item holds.

(c) Let us first assume that ρ_1 is continuous. Suppose that there exists a sequence $\{\boldsymbol{\theta}_n\}$ such that $\lim_{n\to\infty} \|\boldsymbol{\theta}_n\| = \infty$, $\lim_{n\to\infty} \boldsymbol{\theta}_n/\|\boldsymbol{\theta}_n\| = \boldsymbol{\theta}_1$. Using $s_n(\boldsymbol{\theta}_n)$ instead of $s(\boldsymbol{\theta}_n)$ in (19) we yield a contradiction. By assuming (20) with $s_n(\boldsymbol{\theta}_n)$ we proceed as follows. Let us call

$$Y_n = \sup_{\boldsymbol{\theta} \in \Re^p, s \in \Re^+} \left| E_{H_n} \rho_1^* \left(\frac{r(\boldsymbol{\theta})}{s} \right) - E_{H_o} \rho_1^* \left(\frac{r(\boldsymbol{\theta})}{s} \right) \right|$$

and

$$Z_n = \sup_{\boldsymbol{\theta} \in \Re^p, s \in \Re^+} \left| E_{H_n} \rho_2 \left(\frac{r(\boldsymbol{\theta})}{s} \right) - E_{H_o} \rho_2 \left(\frac{r(\boldsymbol{\theta})}{s} \right) \right|.$$

Y_n and Z_n converge almost surely to 0 by Lemma 1 (b). By definition of s_n it holds that

$$\begin{aligned} 1 - b &\leq E_{H_n} \rho_1^* \left(\frac{r(\boldsymbol{\theta}_n)}{s_n(\boldsymbol{\theta}_n)} \right) - E_{H_o} \rho_1^* \left(\frac{r(\boldsymbol{\theta}_n)}{s_n(\boldsymbol{\theta}_n)} \right) + E_{H_o} \rho_1^* \left(\frac{r(\boldsymbol{\theta}_n)}{s_n(\boldsymbol{\theta}_n)} \right) \leq \\ & Y_n + E_{H_o} \rho_1^* \left(\frac{r(\boldsymbol{\theta}_n)}{s_n(\boldsymbol{\theta}_n)} \right) \end{aligned} \tag{23}$$

$$\begin{aligned} 1 - b \geq\ & E_{H_n} \rho_1^* \left(\frac{r(\boldsymbol{\theta}_n)}{(1-\epsilon_n) s_n(\boldsymbol{\theta}_n)} \right) - E_{H_o} \rho_1^* \left(\frac{r(\boldsymbol{\theta}_n)}{(1-\epsilon_n) s_n(\boldsymbol{\theta}_n)} \right) + \\ & E_{H_o} \rho_1^* \left(\frac{r(\boldsymbol{\theta}_n)}{(1-\epsilon_n) s_n(\boldsymbol{\theta}_n)} \right) \geq -Y_n + E_{H_o} \rho_1^* \left(\frac{r(\boldsymbol{\theta}_n)}{(1-\epsilon_n) s_n(\boldsymbol{\theta}_n)} \right) \end{aligned} \tag{24}$$

where $\{\epsilon_n\}$ is a positive sequence converging to 0. Therefore, after taking limit in (23) and (24) it follows that $E_{G_o}\rho_1(a\boldsymbol{\theta}_1'\mathbf{x}) = b$. Since

$$(25)\quad \begin{array}{l} E_{H_n}\rho_2\left(\frac{r(\boldsymbol{\theta}_n)}{s_n(\boldsymbol{\theta}_n)}\right) - E_{H_o}\rho_2\left(\frac{r(\boldsymbol{\theta}_n)}{s_n(\boldsymbol{\theta}_n)}\right) + E_{H_o}\rho_2\left(\frac{r(\boldsymbol{\theta}_n)}{s_n(\boldsymbol{\theta}_n)}\right) \geq \\ -Z_n + E_{H_o}\rho_2\left(\frac{r(\boldsymbol{\theta}_n)}{s_n(\boldsymbol{\theta}_n)}\right), \end{array}$$

$\lim_{n\to\infty} E_{H_o}\rho_2\left(\frac{r(\boldsymbol{\theta}_n)}{s_n(\boldsymbol{\theta}_n)}\right) = 0$ implies (21) again, but this contradicts **(F7)**. Hence, there exists a compact set C in $\Re^p$ such that $\inf_{\boldsymbol{\theta}\in C^c} E_{H_n}\rho_2\left(\frac{r(\boldsymbol{\theta})}{s_n(\boldsymbol{\theta})}\right) > 0$. On the other hand

$$(26)\quad \begin{array}{l} \inf_{\boldsymbol{\theta}\in C} E_{H_n}\rho_2\left(\frac{r(\boldsymbol{\theta})}{s_n(\boldsymbol{\theta})}\right) - E_{H_o}\rho_2\left(\frac{r(\boldsymbol{\theta})}{s_n(\boldsymbol{\theta})}\right) + E_{H_o}\rho_2\left(\frac{r(\boldsymbol{\theta})}{s_n(\boldsymbol{\theta})}\right) - E_{H_o}\rho_2\left(\frac{r(\boldsymbol{\theta})}{s(\boldsymbol{\theta})}\right) + \\ E_{H_o}\rho_2\left(\frac{r(\boldsymbol{\theta})}{s(\boldsymbol{\theta})}\right) \geq -Z_n - \sup_{\boldsymbol{\theta}\in C}\left|E_{H_o}\rho_2\left(\frac{r(\boldsymbol{\theta})}{s_n(\boldsymbol{\theta})}\right) - E_{H_o}\rho_2\left(\frac{r(\boldsymbol{\theta})}{s(\boldsymbol{\theta})}\right)\right| + \\ \inf_{\boldsymbol{\theta}} E_{H_o}\rho_2\left(\frac{r(\boldsymbol{\theta})}{s(\boldsymbol{\theta})}\right) \end{array}$$

Using Lemma 1 (b) and items (a) and the uniform continuity of $R_2(\boldsymbol{\theta}, s)$ on compact sets, $-Z_n - \sup_{\boldsymbol{\theta}\in C}\left|E_{H_o}\rho_2\left(\frac{r(\boldsymbol{\theta})}{s_n(\boldsymbol{\theta})}\right) - E_{H_o}\rho_2\left(\frac{r(\boldsymbol{\theta})}{s(\boldsymbol{\theta})}\right)\right|$ converges to 0 almost surely. Using (b), the statement follows. Under assumptions (II) or (III) the proof follows closely to that of item (b) under such conditions.

Proof of Lemma 6 Let

$$C_n = \{f : \{1,\ldots,p\} \to \{1,\ldots,n\} \text{ such that } f(i) \neq f(j) \text{ if } i \neq j\} \text{ and}$$

$$T^{(n)} = \{t_i = \{f_i(1),\ldots,f_i(p)\} \text{ where } f_i \in C_n\}.$$

$\#T^{(n)} = \binom{n}{p} = c_n$. If $T_n = (t_1,\ldots,t_N)$ denotes a vector of N randomly chosen elements of $T^{(n)}$ then

$$P(T_n = (t_1^o,\ldots,t_N^o)) = 1/(c_n c_{n-1}\ldots(c_n - (N-1)))$$

for a given $(t_1^o,\ldots,t_N^o)$. Let $\mathbf{Y}_{t_i} \in \Re^p$, $\mathbf{X}_{t_i} \in \Re^{p\times p}$ and $\mathbf{A}_{t_i} \in \Re^{p\times p}$ be given by $\mathbf{Y}_{t_i} = (Y_{f_i(1)},\ldots,Y_{f_i(p)})'$, $\mathbf{X}_{t_i} = (\mathbf{x}_{f_i(1)},\ldots,\mathbf{x}_{f_i(p)})^T$ and and $\mathbf{A}_{t_i} = (\mathbf{X}_{t_i}\mathbf{X}_{t_i}')^{-1}$. $(\mathbf{Y}'\mathbf{A}\mathbf{Y})_{T_n} > \epsilon^2$ will denote the set $\{\mathbf{Y}_{t_1}'\mathbf{A}_{t_1}\mathbf{Y}_{t_1} > \epsilon^2,\ldots,\mathbf{Y}_{t_N}'\mathbf{A}_{t_N}\mathbf{Y}_{t_N} > \epsilon^2\}$. Define the set

$$I_h = \{\{t_{i_1},\ldots,t_{i_h}\} \text{ such that } t_{i_j} \in T^{(n)} \text{ and } t_{i_k} \cap t_{i_j} = \emptyset \text{ if } 1 \leq k \neq j \leq h\}$$

and $N'(T_n) = \max_{I_h} \#I_h$. Therefore, given $\epsilon > 0$ and $M > 0$ we get that

$$\begin{array}{l} P(\min_{\alpha\in D_n} \|\alpha\| > \epsilon) = P(\min_{\alpha\in D_n} \|\alpha\|_n > \epsilon \cap N'(T_n) > M) + \\ P(\min_{\alpha\in D_n} \|\alpha\| > \epsilon \cap N'(T_n) \leq M) = A_1 + B_1, \end{array}$$

where A_1 and B_1 denotes the first and second term respectively. We have that

$$\begin{aligned} A_1 = \; & E_{T_n} P(\min_{\boldsymbol{\alpha} \in D_n} \|\boldsymbol{\alpha}\| > \epsilon \cap N'(T_n) > M | T_n) = \\ & \sum_{\overline{T}_n} P(T_n = \overline{T}_n) P((\mathbf{Y}'\mathbf{A}\mathbf{Y})_{T_n} > \epsilon^2 \cap N'(T_n) > M | T_n = \overline{T}_n) = \\ & \sum_{\overline{T}_n} P(T_n = \overline{T}_n) P((\mathbf{Y}'\mathbf{A}\mathbf{Y})_{\overline{T}_n} > \epsilon^2) I_{(M,\infty)}(N'(\overline{T}_n)) \end{aligned}$$

where the last equality follows since $(\mathbf{x}_i', y_i)'$ and T_n are independent. Thus,

$$\begin{aligned} & \sum_{\overline{T}_n} P(T_n = \overline{T}_n) P((\mathbf{Y}'\mathbf{A}\mathbf{Y})_{\overline{T}_n} > \epsilon^2) I_{(M,\infty)}(N'(\overline{T}_n)) \leq \\ & \sum_{\overline{T}_n} P(T_n = \overline{T}_n) P(\mathbf{Y}_{\bar{t}_1}' \mathbf{A}_{\bar{t}_1} \mathbf{Y}_{\bar{t}_1} > \epsilon^2)^M I_{(M,\infty)}(N'(\overline{T}_n)) \leq P(\mathbf{Y}_1'\mathbf{A}_1\mathbf{Y}_1 > \epsilon)^M \end{aligned}$$

where $\mathbf{Y}_1$ and $\mathbf{A}_1$ are based on the first p vectors $(\mathbf{x}_1', y_1)', \ldots, (\mathbf{x}_p', y_p)'$.

On the other hand,

$$\begin{aligned} B_1 = \; & E_{T_n} P(\min_{\boldsymbol{\alpha} \in D_n} \|\boldsymbol{\alpha}\| > \epsilon \cap N'(T_n) < M | T_n) = \\ & \sum_{\overline{T}_n} P(T_n = \overline{T}_n) P((\mathbf{Y}'\mathbf{A}\mathbf{Y})_{T_n} > \epsilon^2 \cap N'(T_n) < M | T_n = \overline{T}_n) = \\ & \sum_{\overline{T}_n} P(T_n = \overline{T}_n) P((\mathbf{Y}'\mathbf{A}\mathbf{Y})_{\overline{T}_n} > \epsilon^2) I_{[0,M]}(N'(\overline{T}_n)) \\ & \sum_{\overline{T}_n} P(T_n = \overline{T}_n) I_{[0,M]}(N'(\overline{T}_n)). \end{aligned}$$

We will prove that

(27) $$\lim_{n\to\infty} \sum_{\overline{T}_n} P(T_n = \overline{T}_n) I_{[0,M]}(N'(\overline{T}_n)) = 0.$$

Take a subset of indexes J_{N_0} such that $\# J_{N_0} = N_0$. We suppose that $N_0 = N_0(n)$ and $\Delta = \limsup_{n\to\infty} N_0(n) < \infty$. Now, we define the set

$$A_{J_{N_0}} = \{\overline{T}_n = (t_1, \ldots, t_N), t_i \in T^{(n)} \text{ and } t_i \cap J_{N_0} \neq 0 \text{ if } 1 \leq i \leq N\}$$

and we denote

$$b_{n,N_0} = \sum_{k=1}^{p} \binom{N_0}{k} \binom{n - N_0}{p - k} = \binom{n}{p} - \binom{n - N_0}{p}.$$

Hence,

$$\frac{b_{n,N_0}}{c_n} = 1 - \prod_{j=1}^{p} \frac{n - N_0 - p + j}{n - p + j},$$

$$P(T_n \in A_{J_{N_0}}) = \prod_{j=1}^{N} \frac{b_{n,N_0} - j + 1}{c_n - j + 1} = \left(\frac{b_{n,N_0}}{c_n}\right)^N \prod_{j=1}^{N-1} \frac{1 - (j-1)/b_{n,N_0}}{1 - (j-1)/c_n} \text{ and}$$

$$P(T_n \in A_{J_{N_0}} \text{ for some } J_{N_0}) \leq \binom{n}{N_0} P(T_n \in A_{J_{N_0}}).$$

We want to prove that $\lim_{n\to\infty}\binom{n}{N_0}P(T_n \in A_{J_{N_0}}) = 0$. To get this result we should note that

(i) $\prod_{j=1}^{N-1}\frac{1-j/b_{n,N_0}}{1-j/c_n} \leq 1$, since each factor is less than 1.

(ii) $\binom{n}{N_0}\left(\frac{b_{n,N_0}}{c_n}\right)^N \leq M_0\left(1-\prod_{j=1}^{p}\frac{1-(N_0+p-j)/n}{1-(p-j)/n}\right)^{N-N_0}$ for some constant $M_0 > 0$.

$$\binom{n}{N_0}\left(\frac{b_{n,N_0}}{c_n}\right)^N =$$

$$\frac{1}{N_0!}\prod_{j=1}^{N_0}\left(1-\frac{(N_0-j)}{n}\right)\left(1-\frac{N_0-j}{n}\right)\left(1-\prod_{j=1}^{p}\frac{1-(N_0+p-j)/n}{1-(p-j)/n}\right)^{N_0} n^{N_0}$$
$$\times\left(1-\prod_{j=1}^{p}\frac{1-(N_0+p-j)/n}{1-(p-j)/n}\right)^{N-N_0} \leq M_0\left(1-\prod_{j=1}^{p}\frac{1-(N_0+p-j)/n}{1-(p-j)/n}\right)^{N-N_0}$$

for some positive constant M_0. This follows from the fact that

$$n\left(1-\prod_{j=1}^{p}\frac{1-(N_0+p-j)/n}{1-(p-j)/n}\right) = n\left(\frac{1+O(1/n)-1+O(1/n)}{\prod_{j=1}^{p}1-(p-j)/n}\right) = O(1).$$

Consequently,

$$\lim_{n\to\infty}\binom{n}{N_0}\left(\frac{b_{n,N_0}}{c_n}\right)^N\prod_{j=1}^{N-1}\frac{1-(j-1)/b_{n,N_0}}{1-(j-1)/c_n} = 0,$$

$$\lim_{n\to\infty}\sum_{i=1}^{\Delta}\binom{n}{i}\left(\frac{b_{n,i}}{c_n}\right)^N\prod_{j=1}^{N-1}\frac{1-(j-1)/b_{n,i}}{1-(j-1)/c_n} = 0.$$

Let $I_{N_n'(\overline{T}_n)}$ be defined as above and $s_n \in I_{N_n'(\overline{T}_n)}$. Put $E_{\Delta_n} = \cup_{\bar{t}_i\in s_n}\bar{t}_i$ where $\Delta_n = \#E_{\Delta_n} \leq pM$ $\forall$ n. It follows that $I_{[0,M)}(N_n'(\overline{T}_n)) = 1$ implies that $\overline{T}_n \in A_{E_{\Delta_n}}$. Hence,

$$\lim_{n\to\infty}\sum_{\overline{T}_n}P(T_n=\overline{T}_n)I_{[0,M]}(N'(\overline{T}_n)) \leq$$
$$\lim_{n\to\infty}\sum_{i=1}^{pM}P(T_n\in A_{E_{\Delta_n}} \text{ for some } E_{\Delta_n} \text{ if } \Delta_n = i) =$$
$$\lim_{n\to\infty}\sum_{i=1}^{pM}\binom{n}{i}\left(\frac{b_{n,i}}{c_n}\right)^N\prod_{j=1}^{N-1}\frac{1-j/b_{n,i}}{1-j/c_n} = 0$$

and (27) is valid.

Thus,

$$\lim_{n\to\infty}P(\min_{\alpha\in D_n}\|\alpha\| > \epsilon) \leq P(\mathbf{Y}_1'\mathbf{A}_1\mathbf{Y}_1 > \epsilon^2)^M$$

Since M can be taken arbitrarily large, we can conclude that $\min_{\alpha\in D_n}\|\alpha\| = o_p(1)$. □

Acknowledgements: The authors would like to thank to Victor J. Yohai for having suggested the problem and for his generous guidance which improved substantially the quality of this paper and to Graciela L. Boente for her comments on a first version of this paper which led to improvements in readability and correctness.

References

[1] Beran, R. and Millar, P. W. (1986), *Confidence sets for multivariate distributions* , Ann. Statist. 14, 431–443.

[2] Davies, L. (1990), *The asymptotics of S–estimators in the linear regression model* , Ann. Statist. 18, 1651–1675.

[3] Donoho, D. L. and Huber, P. C. (1983), *The notion of breakdown point* . In Festschrift for Erich L. Lehmann, P.J. Bickel, K.L. Docksum and J. L. Hodges Jr. Eds Wadsworth, Belmont, California, 157–184.

[4] Hawkins, D. M. (1993), *The accuracy of elemental set approximations for regression* , J. Amer. Statist. Assoc. 88, 580–589.

[5] Huber, P. (1981), *Robust Statistics* (Wiley, New York).

[6] Kim, J. and Pollard, D. (1990), *Cube roots asymptotics* , Ann. Statist. 18, 191–219.

[7] Maronna, R. and Yohai, V. J. (1993), *Bias–robust estimates of regression based on projections* , Ann. Statist. 21,965–990.

[8] Rousseuw, P. (1984), *Least median of squares regression* , J. Amer. Statist. Assoc. 79, 871–880.

[9] Rousseeuw, P. and Basset, G. (1991), *Robustness of the p–subset algorithm for regression with high breakdown point* , W.Stahel and S. Weisberg eds., Directions in Robust Statistics and Diagnostics, Part II, IMA vols. in Math. and its Appl. No. 34 (Springer, New York), 185–194.

[10] Stromberg, A. J. (1993), *Computing the exact least median of squares estimate and stability diagnostics in multiple linear regression* , SIAM J. Sci. Comput. 14, 1289–1299.

[11] Yohai, V. J. and Zamar, R. (1988), *High breakdown point estimates of regression by means of the minimization of an efficient scale* , J. Amer. Statist. Assoc. 83, 406–413.

[12] Yohai, V. J. and Zamar, R. (1986), *High breakdown point estimates of regression by means of the minimization of an efficient scale* , Technical Report 84, University of Washington, Department of Statistics.