Jan Ámos Víšek

Consistency of the least weighted squares under heteroscedasticity

Kybernetika, Vol. 47 (2011), No. 2, 179--206

Persistent URL: `http://dml.cz/dmlcz/141567`

KYBERNETIKA — VOLUME 47 (2011), NUMBER 2, PAGES 179–206

CONSISTENCY OF THE LEAST WEIGHTED SQUARES UNDER HETEROSCEDASTICITY

Jan Ámos Víšek

A robust version of the Ordinary Least Squares accommodating the idea of weighting the order statistics of the squared residuals (rather than directly the squares of residuals) is recalled and its properties are studied. The existence of solution of the corresponding extremal problem and the consistency under heteroscedasticity is proved.

Keywords: robustness, weighting the order statistics of the squared residuals, consistency of the least weighted squares under heteroscedasticity

Classification: 62J02, 62F35

1. BASIC FRAMEWORK AND WEIGHTING THE ORDER STATISTICS

Let $\mathcal{N}$ denote the set of all positive integers, R the real line and R^p the p-dimensional Euclidean space. All vectors will be assumed to be the column ones and throughout the paper, we assume that all r.v.'s are defined on a basic probability space $(\Omega, \mathcal{A}, P)$. For a sequence of $(p+1)$-dimensional random variables $\{(X_i^{'}, e_i)^{'}\}_{i=1}^{\infty}$, any $n \in \mathcal{N}$ and $\beta^0 \in R^p$ the linear regression model given as

$$Y_i = X_i^{'}\beta^0 + e_i = \sum_{j=1}^{p} X_{ij}\beta_j^0 + e_i, \quad i = 1, 2, \ldots, n \tag{1}$$

will be considered. Further, for any $\beta \in R^p$ $\quad r_i(\beta) = Y_i - X_i^{'}\beta$ denotes the ith residual and $r_{(h)}^2(\beta)$ stays for the hth order statistic among the squared residuals, i. e. we have

$$r_{(1)}^2(\beta) \le r_{(2)}^2(\beta) \le \cdots \le r_{(n)}^2(\beta). \tag{2}$$

Without loss of generality we may assume that $\beta^0 = 0$ (otherwise we should write in what follows $\beta - \beta^0$ instead of β). For any matrix $A = \{a_{ij}\}_{i=1,j=1}^{n,m}$ denote by $\|A\|$ Frobenius norm, i. e. $\sqrt{\sum_{i=1}^{n}\sum_{j=1}^{m} a_{ij}^2}$. Finally, for any $n \in \mathcal{N}$ let $w_i \in [0, 1]$, $i = 1, 2, \ldots, n$ be weights.

We are going to give a proof of consistency of the robust estimator of the regression coefficients given in the next definition, see Víšek [19], under heteroscedasticity of error terms.

Definition 1.1. The solution of the extremal problem

$$\hat{\beta}^{(\mathrm{LWS},n,w)} = \underset{\beta \in R^p}{\arg\min} \sum_{i=1}^{n} w_i r_{(i)}^2(\beta) \tag{3}$$

is called the *Least Weighted Squares* estimator (LWS).

Although the consistency was already proved (under homoscedasticity) in Víšek [20, 21] and Mašíček [10], the proofs were very complicated (employing e. g. a sophisticated modification of Prokhorov metric). Present way was opened by establishing uniform convergence (uniform with respect to regression coefficients) of empirical distribution functions of residuals (generally heteroscedastic, see Lemma A.7) to the theoretical one and of similar result for regression combinations of explanatory variables (see Lemma A.6). These results are similar to the results which are usually established in the theory of empirical processes but here we need the only assumption, namely the independence of the r. v.'s in the sequence $\{(X_i^{'}, e_i)^{'}\}_{i=1}^{\infty}$. The present result allows to start the studies concerning robustified White test (especially its power) and proposals of White-type estimator of covariance matrix of the LWS-estimates of regression coefficients. Such estimator will be resistant against heteroscedasticity – similarly as the "classic" White estimate of covariance matrix for OLS-estimates – and so it will allow to evaluate properly the significance of explanatory variables (neglecting the influence of heteroscedasticity leads frequently to overestimation of significance of explanatory variables, consequently to an overfitted model and hence finally to generally (and unfortunately frequently) to less efficient estimates of regression coefficients). Moreover, although the estimators in the overfitted model are generally unbiased, for the datasets which are not very large, the etimators can attain quite misleading values.

First of all, let's show that (3) has a solution and then briefly remind the reasons for the definition.

Theorem 1.2. Let $\{(X_i^{'}, e_i)^{'}\}_{i=1}^{\infty}$ be a sequence of random variables. Then for any $n \in \mathcal{N}$ the solution of (3) always exists.

Proof. Fix an $\omega_0 \in \Omega$, $n_0 \in \mathcal{N}$ and put $W = \mathrm{diag}\{w_1, w_2, \ldots, w_{n_0}\}$. Then consider observations $\{(Y_i(\omega_0), X_i^{'}(\omega_0))^{'}\}_{i=1}^{n_0}$ with $Y_i(\omega_0) = X_i^{'}(\omega_0)\beta^0 + e_i(\omega_0)$ and define matrix $X(\omega_0) = (X_1(\omega_0), X_2(\omega_0), \ldots, X_{n_0}(\omega_0))^{'}$ and vector $Y(\omega_0) = (Y_1(\omega_0), Y_2(\omega_0), \ldots$ $\ldots, Y_{n_0}(\omega_0))^{'}$. For a given permutation π of indices $\{1, 2, \ldots, n_0\}$ denote $Y(\pi, \omega_0)$ and $X(\pi, \omega_0)$ the vector and the matrix obtained as corresponding permutation of coordinates of vector $Y(\omega_0)$ and of rows of matrix $X(\omega_0)$, respectively. For the data $\big(Y(\pi, \omega_0), X(\pi, \omega_0)\big)$ evaluate the *Weighted Least Squares* by (classical) formula

$$\hat{\beta}^{(\mathrm{WLS},n_0,W,\pi)}(\omega_0) = \left(X^{'}(\pi, \omega_0) \cdot W \cdot X(\pi, \omega_0)\right)^{-1} \cdot X^{'}(\pi, \omega_0) \cdot W \cdot Y(\pi, \omega_0)$$

(where we have assumed that $X^{'}(\pi, \omega_0) \cdot W \cdot X\,(\pi, \omega_0)$ is regular; if it doesn't hold we use pseudoinverze). Repeat it for all permutations. Then select that permutation, say $\pi_{\min} = \pi_{\min}(\omega_0)$, for which

$$\sum_{i=1}^{n_0} w_i \cdot \left(Y_i(\pi, \omega_0) - X_i^{'}(\pi, \omega_0) \hat{\beta}^{(\mathrm{WLS}, n_0, W, \pi)}(\omega_0) \right)^2 \tag{4}$$

is minimal. Then $\hat{\beta}^{(\mathrm{WLS}, n_0, W, \pi_{\min})}(\omega_0)$ is solution of (3) at the point ω_0 because for any other $\tilde{\pi}$

$$\begin{aligned}
& \sum_{i=1}^{n_0} w_i \cdot \left(Y_i(\pi_{\min}, \omega_0) - X_i^{'}(\pi_{\min}, \omega_0) \hat{\beta}^{(\mathrm{WLS}, n_0, W, \pi_{\min})}(\omega_0) \right)^2 \\
\leq \ & \sum_{i=1}^{n_0} w_i \cdot \left(Y_i(\tilde{\pi}, \omega_0) - X_i^{'}(\tilde{\pi}, \omega_0) \hat{\beta}^{(\mathrm{WLS}, n_0, W, \tilde{\pi})}(\omega_0) \right)^2 \\
= \ & \inf_{\beta \in R^p} \sum_{i=1}^{n_0} w_i \cdot \left(Y_i(\tilde{\pi}, \omega_0) - X_i^{'}(\tilde{\pi}, \omega_0) \beta \right)^2 .
\end{aligned}$$

It means that $\hat{\beta}^{(\mathrm{LWS}, n_0, w)}(\omega_0) = \hat{\beta}^{(\mathrm{WLS}, n_0, W, \pi_{\min})}(\omega_0)$.

Repeating this at first for all $\omega \in \Omega$ and secondly for all $n \in \mathcal{N}$, we conclude the proof. □

Remark 1.3. Let's return to the fact that $\hat{\beta}^{(\mathrm{LWS}, n, w)}(\omega) = \hat{\beta}^{(\mathrm{WLS}, n_0, W, \pi_{\min})}(\omega)$ (which we found at the end of proof of Theorem 1.2). Moreover, let's recall that the estimate by means of *Weighted Least Squares* $\hat{\beta}^{(\mathrm{WLS}, n_0, W, \pi_{\min})}(\omega)$ is one of the solutions of the normal equations

$$X^{'}(\pi_{\min}, \omega) \cdot W \cdot \left(Y(\pi_{\min}, \omega) - X(\pi_{\min}, (\omega)) \beta \right) = 0.$$

Then we conclude that $\hat{\beta}^{(\mathrm{LWS}, n, w)}(\omega)$ is one of solutions of the same normal equations, written usually without stressing dependence on ω as

$$X^{'}(\pi_{\min}) \cdot W \cdot \left(Y(\pi_{\min}) - X(\pi_{\min}) \beta \right) = 0. \tag{5}$$

Remark 1.4. Putting for any $n \in \mathcal{N}$ and for $h \in \{1, 2, \ldots, n\}$ $w_h = 1$ and $w_i = 0$ for $i \neq h$, (3) yields the *Least Median of Squares* (Rousseeuw [11])

$$\hat{\beta}^{(\mathrm{LMS}, n, h)} = \arg\min_{\beta \in R^p} r^2_{(h)}(\beta).$$

Similarly, $w_i = 1, \quad i \leq h$ and $w_i = 0$ for $i > h$ gives the *Least Trimmed Squares* (Hampel et al. [5])

$$\hat{\beta}^{(\mathrm{LTS}, n, h)} = \arg\min_{\beta \in R^p} \sum_{i=1}^{h} r^2_{(i)}(\beta).$$

Let's summarize pros and cons of $\hat{\beta}^{(\mathrm{LMS}, n, h)}$ and $\hat{\beta}^{(\mathrm{LTS}, n, h)}$. It will hint, what we should require to hold for the weights w_i's.

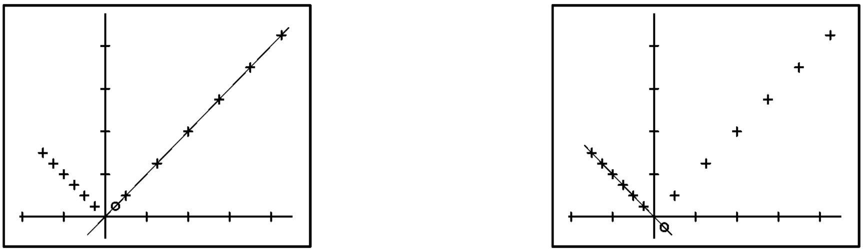

Fig. 1.

First of all, $\hat{\beta}^{(\mathrm{LMS},n,h)}$ and $\hat{\beta}^{(\mathrm{LTS},n,h)}$ are *scale* and *regression equivariant* and $\hat{\beta}^{(\mathrm{LWS},n,h)}$ shares this property with them[1].

Let's recall that for $h = \frac{n}{2} + \frac{p+1}{2}$ both $\hat{\beta}^{(\mathrm{LMS},n,h)}$ as well as $\hat{\beta}^{(\mathrm{LTS},n,h)}$ have asymptotically breakdown point equal to 0.5 (see Rousseeuw, Leroy [12]). Nevertheless, as the pictures (see Fig. 1) demonstrate the high breakdown point may cause high sensitivity to a small shift of observation (for real data exhibiting the same phenomenon see Hettmansperger, Sheather [6], together with Víšek [16]). The sensitivity is due to the fact that both estimators have the discontinuous "loss function", i. e. that the the weights w_i's are only either 0 or 1. Similarly, robust estimators with discontinuous "loss function" exhibit the (high) sensitivity with respect to the deletion of point(s), see e. g. Víšek [17, 18, 22, 23]. To remove it we should decrease the influence the influential observations in a less steep way.

Moreover, it is known that $\hat{\beta}^{(\mathrm{LMS},n,h)}$ is not $\sqrt{n}$-consistent while $\hat{\beta}^{(\mathrm{LTS},n,h)}$ possesses this property (Rousseeuw, Leroy [12]). It hints that probably the weights are to be nonzero for more than one observation and possibly nonincreasing.

Taking into account previous considerations and assuming that the weights are generated by a function w in the way $w_i = w\left(\frac{i-1}{n}\right)$, let's put:

Conditions $\mathcal{C}1$. The weight function $w(u)$ is continuous, nonincreasing, $w : [0,1] \to [0,1]$ with $w(0) = 1$.

The form of definition of LWS as given in (3) is not suitable for considerations on the consistency of the estimator. So, following Hájek and Šidák [4] for any

[1]Notice that many robust estimators as e. g. M-estimators, need not necessarily to posses it. Generally, to reach *scale* and *regression equivariance* for M-estimators, we have to studentize the residuals by *scale invariant* and *regression equivariant* estimate of scale of error terms, see Bickel [1] or Jurečková, Sen [8]. However, to establish such an estimator is not a simple task, see Croux, Rousseeuw [3], Jurečková, Sen [8] or Víšek [24]. Moreover, all of them are in fact based on a preliminary robust *scale*- and the *regression-equivariant* estimator of the regression coefficients. It implies that the (robust) estimators which need not require the studentization of residuals are preferable in the applications. $\hat{\beta}^{(\mathrm{LWS},n,h)}$ is one such possibility.

$i \in \{1, 2, \ldots, n\}$ let us put

$$\pi(\beta, i) = j \in \{1, 2, \ldots, n\} \;\Leftrightarrow\; r_i^2(\beta) = r_{(j)}^2(\beta) \tag{6}$$

(notice that again $\pi(\beta, i) = \pi(\beta, i, \omega)$, since it depends on $X_i(\omega)$'s and $e_i(\omega)$'s). Then we have from (3)

$$\hat{\beta}^{(\mathrm{LWS},n,w)} = \arg\min_{\beta \in R^p} \sum_{j=1}^{n} w\left(\frac{j-1}{n}\right) r_{(j)}^2(\beta) = \arg\min_{\beta \in R^p} \sum_{i=1}^{n} w\left(\frac{\pi(\beta, i) - 1}{n}\right) r_i^2(\beta). \tag{7}$$

Now, returning to (5) and employing (6), we obtain normal equations in the form

$$\sum_{i=1}^{n} w\left(\frac{\pi(\beta, i) - 1}{n}\right) \cdot X_i \cdot (Y_i - X_i'\beta) = 0. \tag{8}$$

Further, for any $\beta \in R^p$ and any $n \in \mathcal{N}$ the empirical distribution of the absolute value of residual will be denoted $F_\beta^{(n)}(r)$. It means that, denoting the indicator of a set A by $I\{A\}$, we have (remember we put $\beta^0 = 0$)

$$F_\beta^{(n)}(r) = \frac{1}{n}\sum_{j=1}^{n} I\{|r_j(\beta)| < r\} = \frac{1}{n}\sum_{j=1}^{n} I\left\{|e_j - X_j'\beta| < r\right\}. \tag{9}$$

Now, realize please, that having fixed $\beta \in R^p$ and denoting $|r_i(\beta)| = a_i(\beta)$, the order statistics $a_{(i)}(\beta)$'s and the order statistics of the squared residuals $r_{(i)}^2(\beta)$'s assign to given fix observation the same rank, i. e. if the squared residual of given fix observation is on the ℓth position (say) in the sequence

$$r_{(1)}^2(\beta) \le r_{(2)}^2(\beta) \le \ldots r_{(n)}^2(\beta), \tag{10}$$

then the absolute value of residual of the same observation is in the sequence

$$a_{(1)}(\beta) \le a_{(2)}(\beta) \le \ldots a_{(n)}(\beta) \tag{11}$$

also on the ℓth position. Now, let's realize that the empirical distribution function $F_\beta^{(n)}(r)$ has at point $a_{(\pi(\beta,i))}(\beta)$ its $\pi(\beta, i)$th jump and hence (notice the sharp inequality in our definition of the empirical distribution function, see (9))

$$F_\beta^{(n)}(a_{(\pi(\beta,i))}(\beta)) = F_\beta^{(n)}(|r_i(\beta)|) = \frac{\pi(\beta, i) - 1}{n} \tag{12}$$

(for $\pi(\beta, i)$ see (6)) and so (8) can be written as

$$I\!N E_{Y,X,n}(\beta) = \sum_{i=1}^{n} w\left(F_\beta^{(n)}(|r_i(\beta)|)\right) X_i \left(Y_i - X_i'\beta\right) = 0. \tag{13}$$

The main idea of proving consistency of $\hat{\beta}^{(\mathrm{LWS},n,w)}$ is to approximate $F_\beta^{(n)}(|r_i(\beta)|)$ by a continuous distribution function – as given in Lemma A.7. We shall need for it some assumptions.

Conditions $\mathcal{C}2$. The sequence $\left\{(X_i', e_i)'\right\}_{i=1}^{\infty}$ is sequence of independent $p+1$-dimensional random variables (r.v.'s) distributed according to distribution functions (d.f.) $F_{X,e_i}(x,r) = F_X(x) \cdot F_{e_i}(r)$ where $F_{e_i}(r) = F_e(r\sigma_i^{-1})$ with $I\!\!E e_i = 0$, $\operatorname{var}(e_i) = \sigma_i^2$ and $0 < \liminf_{i\to\infty} \sigma_i \le \limsup_{i\to\infty} \sigma_i < \infty$. Moreover, $F_e(r)$ is absolutely continuous with density $f_e(r)$ bounded by U_e. Finally, there is $q > 1$ so that $I\!\!E \left\|X_1\right\|^{2q} < \infty$ (as $F_X(x)$ doesn't depend on i, the sequence $\{X_i\}_{i=1}^{\infty}$ is sequence of independent and identically distributed (i.i.d.) r.v.'s).

Remark 1.5. The assumption that the d.f. $F_e(r)$ is continuous is not only technical assumption. Possibility that the error terms in regression model are discrete r.v.'s implies problems with treating response variable and it requires special considerations – see chapters on logit or probit models or limited response variables e. g. in Judge et. al. [7]. Absolute continuity is then a technical assumption. Without the density, even bounded density, we should assume that $F_e(r)$ is Lipschitz and it would bring a more complicated form of all what follows.

Remark 1.6. Notice that there are constants $0 < s_\sigma \le S_\sigma < \infty$ so that $s_\sigma \le \sigma_i \le S_\sigma$ for all i's. Moreover, as the density of e_i is given as $f_e(r \cdot \sigma_i^{-1}) \cdot \sigma_i^{-1}$, there is a constant $f_\sigma < \infty$ such that $\sup_{i\in\mathcal{N}} \sup_{r\in R} f_{e_i}(r) < f_\sigma$.

2. ALL SOLUTIONS OF NORMAL EQUATIONS ARE BOUNDED

First of all, we need some auxiliary lemma. Prior to proving it, we have to enlarge our notation. For any $\beta \in R^p$ the distribution of the product $\beta' X X' \beta = (X'\beta)^2$ will be denoted $F_{(X'\beta)^2}(u)$, i. e.

$$F_{(X'\beta)^2}(u) = P\left((X'\beta)^2 < u\right). \tag{14}$$

The empirical distribution of the sequence of i.i.d. r.v.'s $\left\{(X_j'\beta)^2\right\}_{j=1}^{\infty}$ will be denoted $F^{(n)}_{(X'\beta)^2}(u)$, so that

$$F^{(n)}_{(X'\beta)^2}(u) = \frac{1}{n}\sum_{j=1}^{n} I\left\{\left(X_j'\beta\right)^2 < u\right\}. \tag{15}$$

Finally, for any $\lambda \in R^+$ and any $a \in R$ put

$$\gamma_{\lambda,a} = \sup_{\|\beta\|=\lambda} F_{(X'\beta)^2}(a). \tag{16}$$

Notice please that due to the fact that the surface of the ball $\{\beta \in R^p, \|\beta\| = \lambda\}$ is compact, there is $\beta_{\gamma,a} \in \{\beta \in R^p, \|\beta\| = \lambda\}$ so that

$$\gamma_{\lambda,a} = F_{\left(X'\beta_{\gamma,a}\right)^2}(a). \tag{17}$$

Moreover, for any $\beta \in R^p$ denote $\tilde\beta = \beta \cdot \|\beta\|^{-1}$. Then we have

$$\begin{aligned} F_{(X'\beta)^2}(u) &= P((X_1'\beta)^2 < u) \\ &= P\Big(\frac{(X_1'\beta)^2}{\|\beta\|^2} < \frac{u}{\|\beta\|^2}\Big) = F_{\left(X'\tilde\beta\right)^2}\Big(\frac{u}{\|\beta\|^2}\Big). \end{aligned}$$

Then evidently

$$\gamma_{\lambda,a} = \gamma_{1,\frac{a}{\lambda^2}}.$$

It means that we may without any restriction of generality consider only $\gamma_{1,a}$. In what follows there are defined some constants inside the proofs of assertions, lemmas or theorems. They are assumed to be defined only inside the corresponding proof. Now we can prove:

Lemma 2.1. Under Conditions $\mathcal{C}1$ and $\mathcal{C}2$ there is $a > 0$ and $b \in (0, 1)$ so that

$$a \cdot (b - \gamma_{1,a}) \cdot w(b) > 0 \tag{18}$$

(for $\gamma_{1,a}$ see (16)).

Proof. Due to Condition $\mathcal{C}1$ there is $b \in (0, 1)$ such that $w(b) > 0$. Fix one such b. If for all $a \geq 0$ we have $\gamma_{1,a} \geq b$, we have

$$\liminf_{a \to 0_+} \gamma_{1,a} \geq b.$$

So, there is a sequence $\{a_k\}_{k=1}^{\infty}$ such that for all $k = 1, 2, \ldots, a_k > 0$ and

$$\lim_{k\to\infty} a_k = 0 \qquad \text{and} \qquad \liminf_{k\to\infty} \gamma_{1,a_k} \geq b.$$

Then, due to the fact that for each γ_{1,a_k} there is β_{γ,a_k} such that

$$\gamma_{1,a_k} = F_{\left(X'\beta_{\gamma,a_k}\right)^2}(a_k),$$

see (17), we have a sequence $\{\beta_{\gamma,a_k}\}_{k=1}^{\infty}$ such that

$$\liminf_{k\to\infty} F_{\left(X'\beta_{\gamma,a_k}\right)^2}(a_k) \geq b.$$

Applying (again) the argument about the compactness of unit ball, we find finally β^* and a subsequence $\{\beta_{\gamma,a_{k_j}}\}_{j=1}^{\infty}$ so that $\lim_{j\to\infty} \beta_{\gamma,a_{k_j}} = \beta^*$ coordinatewise and that

$$\liminf_{j\to\infty} F_{\left(X'\beta_{\gamma,a_{k_j}}\right)^2}(a_{k_j}) \geq b.$$

Applying Lema A.8 we conclude that

$$0 < b \leq F_{\left(X'\beta^*\right)^2}(0) = P\left((X^{'}\beta^*)^2 < 0\right)$$

which is a contradiction. $\square$

Lemma 2.2. Let Conditions $\mathcal{C}1$ and $\mathcal{C}2$ be fulfilled. Then for any $\varepsilon > 0$ there is $\theta > 0$, $\Delta > 0$ and $n_{\varepsilon,\Delta} \in \mathcal{N}$ such that for any $n > n_{\varepsilon,\Delta}$

$$P\left(\left\{\omega \in \Omega : \inf_{\|\beta\|\geq\theta} -\frac{1}{n}\beta^{'} I\!N E_{Y,X,n}(\beta) > \Delta\right\}\right) > 1 - \varepsilon.$$

In other words, any sequence $\{\hat{\beta}^{(\mathrm{LWS},n,w)}\}_{n=1}^{\infty}$ of the solutions of the sequence of normal equations $I\!N E_{Z,n}(\hat{\beta}^{(\mathrm{LWS},n,w)}) = 0, \quad n = 1, 2, \ldots$ (see (13)) is bounded in probability.

P r o o f. Let us multiply (13) from the left by the transposition of a $\beta \in R^p$ and write it then as

$$\frac{1}{n}\sum_{i=1}^{n} w\left(F_{\beta}^{(n)}(|r_i(\beta)|)\right)\beta^{'} X_i X_i^{'}\beta - \frac{1}{n}\sum_{i=1}^{n} w\left(F_{\beta}^{(n)}(|r_i(\beta)|)\right) e_i X_i^{'}\beta. \tag{19}$$

First of all, we shall pay attention to the quadratic part of (19), i. e. to

$$\frac{1}{n}\sum_{i=1}^{n} w\left(F_{\beta}^{(n)}(|r_i(\beta)|)\right)\beta^{'} X_i X_i^{'}\beta \tag{20}$$

and we'll find a positive definite quadratic form which uniformly for β outside the ball of diameter equal to 2 and with probability at least $1-\varepsilon$ is the lower bound of (20). This quadratic form is then for $\beta \in R^p$ with enough large norm, say larger then some $\theta > 0$, larger than the linear part of $-\frac{1}{n}\beta^{'} I\!\!N E_{Y,X,n}(\beta)$.

Fix $a > 0$ and $b \in (0,1)$, existence of which was shown in Lemma 2.1 and denote the set of all indices $i = 1, 2, \ldots, n$ by I_n. Further, for any $\beta \in R^p$ denote the set of indices for which $F_{\beta}^{(n)}(|r_i(\beta)|) < b$ by $I_b(\beta)$. Returning to (10) or (11), we easy verify that the empirical d.f. overcomes b not later than at its $[nb]+1$ jump, i. e. number of order statistics in (11) at which the empirical d.f. is less or equal to b is at least $[n \cdot b]$ (where $[\,\xi\,]$ denotes the integer part of ξ). It means that

$$\#I_b(\beta) \geq [n \cdot b] \tag{21}$$

where $\#A$ stays for the number of elements of the set A. Realize please that whenever index $i \in I_b(\beta)$, we have $F_{\beta}^{(n)}(|r_i(\beta)|) < b$ which implies that for $i \in I_b(\beta)$ we have (for any $\beta \in R^p$)

$$w\left(F_{\beta}^{(n)}(|r_i(\beta)|)\right) \geq w(b). \tag{22}$$

Now, let us denote $I_a(\beta)$ the set of indices (among $1, 2, \ldots, n$) for which $\beta' X_i X_i' \beta < a$. Finally, let us estimate $\#I_b(\beta)$ and $\#I_a(\beta)$ and take into account only those terms of (20) the indices of which are in $I_b(\beta) \setminus I_a(\beta)$. (There are some other positive terms of (20), contribution of which will be neglected, since their weights are smaller than $w(b)$ or $\beta' X_i X_i' \beta$ is smaller than a.) Note that for the set $I_b(\beta) \setminus I_a(\beta)$ we have

$$\#(I_b(\beta) \setminus I_a(\beta)) \geq \#I_b(\beta) - \#I_a(\beta). \tag{23}$$

Now, let us fix $\varepsilon > 0$, $\delta > 0$ and put

$$\kappa = \frac{a \cdot (b - \gamma_{1,a}) \cdot w(b)}{2}. \tag{24}$$

Then, according to Lemma 2.1, $\kappa > 0$. Employing Lemma A.6 find $n_1 \in \mathcal{N}$ so that for all $n > n_1$ we have

$$P\left(\left\{\omega \in \Omega : \sup_{\beta \in R^p} \sup_{u \in R} \left| F_{(X'\beta)^2}^{(n)}(u) - F_{(X'\beta)^2}(u)\right| \leq \frac{\kappa}{a \cdot w(b)}\right\}\right) > 1 - \frac{\varepsilon}{2} \tag{25}$$

and denote the corresponding set by $B_n^{(1)}$. Recalling that, due to the fact how the empirical distribution function is defined, we have

$$F^{(n)}_{(X'\beta)^2}(a) = \frac{\#\{i : \beta^{'} X_i X_i^{'} \beta < a\}}{n} = \frac{\# I_a(\beta)}{n}.$$

Then we conclude that (25) implies for any $n > n_1$ and $\omega \in B_n^{(1)}$

$$\# I_a(\beta) = n \cdot F^{(n)}_{(X'\beta)^2}(a) < n \cdot \left(F_{(X'\beta)^2}(a) + \frac{\kappa}{a \cdot w(b)} \right) \leq n \cdot \left(\gamma_{1,a} + \frac{\kappa}{a \cdot w(b)} \right) \quad (26)$$

(for $\gamma_{\lambda,a}$ see (16)). Notice that (26) holds only for $\{\beta \in R^p, \|\beta\| = 1\}$. Let us recall that we have denoted by $I_a(\beta)$ the number of indices (among $1, 2, \ldots, n$) for which $\beta' X_i X_i' \beta < a$. (26) then says that we have at most $n \cdot \left(\gamma_{1,a} + \frac{\kappa}{a \cdot w(b)} \right)$ such indices.

Consider $\omega \in B_n^{(1)}$ and $n > n_1$, and put

$$C_n(\beta) = \left\{ i \in I_n : F_\beta^{(n)}(|r_i(\beta)|) < b \ \text{ and } \ \beta' X_i X_i' \beta > a \right\} = I_b(\beta) \setminus I_a(\beta).$$

Then (21) and (26) imply that the number of indices of the set $C_n(\beta)$ is at least (see (23))

$$\# C_n(\beta) \geq \# I_b(\beta) - \# I_a(\beta) \geq n \cdot b - n \cdot \left(\gamma_{1,a} + \frac{\kappa}{a \cdot w(b)} \right) = n \cdot \left(b - \gamma_{1,a} - \frac{\kappa}{a \cdot w(b)} \right).$$

Now, we have for any $n > n_1$, any $\omega \in B_n^{(1)}$ and any $\|\beta\| = 1$

$$\frac{1}{n} \sum_{i=1}^{n} w\left(F_\beta^{(n)}(|r_i(\beta)|) \right) \beta^{'} X_i X_i^{'} \beta \geq \frac{1}{n} \sum_{i \in C_n(\beta)} w(b) \beta^{'} X_i X_i^{'} \beta$$

$$\geq a \cdot (b - \gamma_{1,a}) \cdot w(b) - \kappa > \kappa.$$

Consider now any $\beta \in R^p, \|\beta\| = \theta \geq 1$ and put $\tilde{\beta} = \theta^{-1} \cdot \beta$. Then

$$\beta^{'} X_i X_i^{'} \beta = \theta^2 \tilde{\beta}^{'} X_i X_i^{'} \tilde{\beta}. \quad (27)$$

We have proved that for any $n > n_1$, any $\omega \in B_n^{(1)}$ and any $\beta^* \in R^p, \|\beta^*\| = 1$

$$\# I_a(\beta^*) \leq n \cdot \left(\gamma_{1,a} + \frac{\kappa}{a \cdot w(b)} \right) \quad (28)$$

(see (26) and remember that $I_a(\beta^*)$ was defined as set of those indices from $\{1, 2, \ldots, n\}$ for which $\beta' X_i X_i' \beta < a$). Further, let's recall that $I_b(\beta)$ was defined so that $F_\beta^{(n)}(|r_i(\beta)|) < b$ and hence

$$\# I_b(\beta) \geq [b \cdot n] \quad (29)$$

and for any $i \in I_b(\beta)$

$$w\left(F_\beta^{(n)}(|r_i(\beta)|) \right) \geq w(b). \quad (30)$$

Now, we have from (27), (28), (29) and (30)

$$
\begin{aligned}
&\frac{1}{n}\sum_{i=1}^{n} w\left(F_{\beta}^{(n)}(|r_i(\beta)|)\right)\beta^{'}X_iX_i^{'}\beta \geq \frac{1}{n}\sum_{i\in I_b(\beta)} w(b)\beta^{'}X_iX_i^{'}\beta \\
= \;& \frac{1}{n}\,\theta^2 \sum_{i\in I_b(\beta)} w(b)\tilde{\beta}^{'}X_iX_i^{'}\tilde{\beta} \geq \frac{1}{n}\,\theta^2 \sum_{i\in I_b(\beta)-I_a(\tilde{\beta})} w(b)\tilde{\beta}^{'}X_iX_i^{'}\tilde{\beta} \\
\geq \;& \theta^2\left(a\cdot(b-\gamma_{1,a})\cdot w(b)-\kappa\right)\geq \theta^2\cdot\kappa. \qquad (31)
\end{aligned}
$$

So, we have proved that for any $n > n_1$ and any $\omega \in B_n^{(1)}$ and

$$\frac{1}{n}\sum_{i=1}^{n} w\left(F_{\beta}^{(n)}(|r_i(\beta)|)\right)\beta^{'}X_iX_i^{'}\beta \geq \theta^2\cdot\kappa = \|\beta\|\cdot\kappa.$$

Now, we shall consider the second term in (19). Let e be a r.v. distributed according to $F_e(u)$ and denote $I\!E\{|e|\cdot\|X_1\|\} = \tau$ and $\limsup_{i\to\infty}\sigma_i = \eta$. Then find $n_2 \in \mathcal{N}$ so that for any $n > n_2$ there is $B_n^{(2)}$ so that $P(B_n^{(2)}) > 1-\varepsilon/2$ and for any $\omega \in B_n^{(2)}$ we have (remember that $w(r) \in [0,1]$)

$$\frac{1}{n}\left|\sum_{i=1}^{n} w\left(F_{\beta}^{(n)}(|r_i(\beta)|)\right)e_iX_i^{'}\beta\right| \leq \frac{1}{n}\sum_{i=1}^{n}\left|e_iX_i^{'}\beta\right| \leq 2\tau\cdot\eta\cdot\|\beta\|. \qquad (32)$$

Consider $n > \max\{n_1, n_2\}$ and $\omega \in B_n = B_n^{(1)} \cap B_n^{(2)}$. It follows that $P(B_n) > 1-\varepsilon$ and (31) and (32) imply that for any $\beta \in R^p$ $\|\beta\| \geq 1$ and for κ we have defined in (24)

$$-\frac{1}{n}\beta^{'} I\!N E_{Y,X,n}(\beta) \geq \|\beta\|^2\cdot\kappa - 2\tau\cdot\eta\cdot\|\beta\|.$$

Then for any $\Delta > 0$ there is a $\theta \geq 1$ such that for any $\beta \in R^p, \|\beta\| > \theta$ with probability at least $1-\varepsilon$ we have

$$-\frac{1}{n}\beta^{'} I\!N E_{Y,X,n}(\beta) > \Delta.$$

□

Prior to deriving consistency of $\hat{\beta}^{(\mathrm{LWS},n,w)}$ we need some other results. For proving them we have to strengthen the assumptions.

Conditions $\mathcal{C}1'$. The weight function $w(u)$ is continuous nonincreasing, $w : [0,1] \to [0,1]$ with $w(0) = 1$. Moreover, w is Lipschitz in absolute value, i. e. there is L such that for any pair $u_1, u_2 \in [0,1]$ we have $|w(u_1) - w(u_2)| \leq L\cdot|u_1 - u_2|$.

Further let's put

$$\overline{F}_{n,\beta}(v) = \frac{1}{n}\sum_{i=1}^{n} F_{\beta,i}(v) \qquad (33)$$

where

$$F_{\beta,i}(v) = P\left(|Y_i - X_i'\beta| < v\right) = P\left(|e_i - X_i'\beta| < v\right) \tag{34}$$

(remember that e_i's have different variances σ_i^2 and that we have assumed that $\beta^0 = 0$).

Lemma 2.3. Let Conditions $\mathcal{C}1'$ and $\mathcal{C}2$ be fulfilled. Then for any $\varepsilon > 0$, $\delta \in (0,1)$ and $\zeta > 0$ there is $n_{\varepsilon,\delta,\zeta} \in \mathcal{N}$ so that for any $n > n_{\varepsilon,\delta,\zeta}$ we have

$$P\left(\left\{\omega \in \Omega \ : \ \sup_{\|\beta\| \leq \zeta} \left| \frac{1}{n} \sum_{i=1}^{n} \left\{ w\left(F_\beta^{(n)}(|r_i(\beta)|)\right) \beta' X_i \left(e_i - X_i'\beta\right) \right.\right.\right.\right.$$
$$\left.\left.\left.\left. -\beta' I\!\!E\left[w\left(\overline{F}_{n,\beta}(|r_i(\beta)|)\right) X_i \left(e_i - X_i'\beta\right)\right]\right\} \right| < \delta \right\}\right) > 1 - \varepsilon.$$

P r o o f. Throughout the proof please keep in mind that we have put

$$\overline{F}_{n,\beta}(v) = \frac{1}{n}\sum_{i=1}^{n} F_{\beta,i}(v).$$

Denoting $I\!\!E\,\|X_1\|^2 = \kappa$, let us fix a positive ε, $\delta \in (0,1)$ and $\zeta > 0$. Recalling that we have assumed that $\beta^0 = 0$, we shall consider for $\beta \in R^p$, $\|\beta\| \leq \zeta$ the normal equations (13)

$$I\!\!N E_{Y,X,n}(\beta) = \frac{1}{n}\sum_{i=1}^{n} w\left(F_\beta^{(n)}(|r_i(\beta)|)\right)\beta' X_i X_i'\beta - \frac{1}{n}\sum_{i=1}^{n} w\left(F_\beta^{(n)}(|r_i(\beta)|)\right) e_i X_i'\beta. \tag{35}$$

Let us start (again) with the first term in (35) and put $\tau^{(1)} = \delta/(20\kappa\zeta^2 \cdot L)$, for L see **Condition** $\mathcal{C}1'$. Due to Lemma A.7 we can find $n_1 \in \mathcal{N}$ so that for any $n > n_1$ there is a set $B_n^{(1)}$ such that $P(B_n^{(1)}) > 1 - \varepsilon/10$ and for any $\omega \in B_n^{(1)}$

$$\sup_{\beta \in R^p} \sup_{r \in R} \left| F_\beta^{(n)}(r) - \overline{F}_{n,\beta}(r) \right| \leq \tau^{(1)}. \tag{36}$$

Employing the law of large numbers, find $n_2 > n_1$ so that for any $n > n_2$ there is a set $B_n^{(2)}$ such that $P(B_n^{(2)}) > 1 - \varepsilon/10$ and for any $\omega \in B_n^{(2)}$

$$\frac{1}{n}\sum_{i=1}^{n} \|X_i\|^2 < 2\kappa. \tag{37}$$

Since then for any $n > n_2$ and any $\omega \in B_n^{(1)} \cap B_n^{(2)}$

$$\frac{1}{n}\sup_{\|\beta\| \leq \zeta} \left\| \sum_{i=1}^{n} \left\{ w\left(F_\beta^{(n)}(|r_i(\beta)|)\right) - w\left(F_\beta(|r_i(\beta)|)\right)\right\} X_i X_i' \right\|$$
$$\leq \ \frac{1}{n} L \cdot \tau^{(1)} \cdot \sum_{i=1}^{n} \|X_i\|^2 \leq L \cdot \tau^{(1)} \cdot 2\kappa = \frac{\delta}{10\zeta^2},$$

we have for any $n > n_2$ and any $\omega \in B_n^{(1)} \cap B_n^{(2)}$

$$\frac{1}{n} \sup_{\|\beta\| \le \zeta} \left| \sum_{i=1}^{n} \left\{ w\left(F_\beta^{(n)}(|r_i(\beta)|)\right) - w\left(\overline{F}_{n,\beta}(|r_i(\beta)|)\right) \right\} \beta' X_i X_i' \beta \right| \le \frac{\delta}{10}. \qquad (38)$$

Employing Lemma A.8, find for $\Delta = \frac{\delta}{20 \cdot L \cdot \kappa \zeta^2}$ such $\tau^{(2)} > 0$ that for

$$\mathcal{T}(\zeta, \tau^{(2)}) = \left\{ \left\|\beta^{(1)}\right\| \le \zeta, \left\|\beta^{(2)}\right\| \le \zeta, \left\|\beta^{(1)} - \beta^{(2)}\right\| < \tau^{(2)} \right\} \qquad (39)$$

we have

$$\sup_{\left(\beta^{(1)},\beta^{(2)}\right) \in \mathcal{T}(\zeta,\tau^{(2)})} \sup_{i \in \mathcal{N}} \sup_{r \in R} \left| F_{\beta^{(1)},i}(r) - F_{\beta^{(2)},i}(r) \right| < \Delta$$

and hence also

$$\sup_{\left(\beta^{(1)},\beta^{(2)}\right) \in \mathcal{T}(\zeta,\tau^{(2)})} \sup_{r \in R} \left| \overline{F}_{n,\beta^{(1)}}(r) - \overline{F}_{n,\beta^{(2)}}(r) \right|$$

$$= \sup_{\left(\beta^{(1)},\beta^{(2)}\right) \in \mathcal{T}(\zeta,\tau^{(2)})} \sup_{r \in R} \left| \frac{1}{n} \sum_{i=1}^{n} F_{\beta^{(1)},i}(r) - \frac{1}{n} \sum_{i=1}^{n} F_{\beta^{(2)},i}(r) \right| < \Delta \qquad (40)$$

Let's recall that we have restricted ourselves on $\|\beta\| \le \zeta$. Then due to (37), (39) and (40) for any $n > n_2$ and any $\omega \in B_n^{(1)} \cap B_n^{(2)}$

$$\begin{aligned} &\frac{1}{n} \sup_{(\beta^{(1)},\beta^{(2)}) \in \mathcal{T}(\zeta,\tau^{(2)})} \Big| \sum_{i=1}^{n} \Big\{ w\big(\overline{F}_{n,\beta^{(2)}}(|r_i(\beta^{(2)})|)\big) \\ &\qquad - w\big(\overline{F}_{n,\beta^{(1)}}(|r_i(\beta^{(2)})|)\big) \Big\} [\beta^{(2)}]' X_i X_i' \beta^{(2)} \Big| \\ \le\ & L \cdot \Delta \cdot \zeta^2 \cdot \frac{1}{n} \sum_{i=1}^{n} \|X_i\|^2 \le \frac{\delta}{10} \end{aligned} \qquad (41)$$

(notice that the in the previous inequality the subindices of the d.f.'s are $\beta^{(1)}$ and $\beta^{(2)}$ but the arguments are at the same point $\beta^{(2)}$). Further denote $\gamma^{(1)} = I\!\!E \|X_1\|^{2q}$, $\gamma^{(2)} = I\!\!E \|X_1\|$ and applying the law of large numbers find $n_3 > n_2$ so that for any $n > n_3$ there is a set $B_n^{(3)}$ such that $P(B^{(3)}) > 1 - \varepsilon/10$ and for any $\omega \in B_n^{(3)}$ we have

$$\frac{1}{n} \sum_{i=1}^{n} \|X_i\|^{2q} < 2\gamma^{(1)} \qquad \text{and} \qquad \frac{1}{n} \sum_{i=1}^{n} \|X_i\| < 2\gamma^{(2)}.$$

Finally, let us recall that $w(r) \in [0, 1]$, so that for any pair $r_1, r_2 \in R$ we have $|w(r_1) - w(r_2)| \le 1$ and hence for any $q' > 1$

$$|w(r_1) - w(r_2)|^{q'} \le |w(r_1) - w(r_2)|. \qquad (42)$$

Let q' be such that $\frac{1}{q'} + \frac{1}{q} = 1$ (for q see Conditions $\mathcal{C}2$). Then select some

$$\tau^{(3)} \in \left(0, \min\left\{\tau^{(2)}, \delta \cdot \left(2^{3q'+q} \cdot f_\sigma \cdot L \left[\gamma^{(1)}\right]^{\frac{q'}{q}} \cdot \gamma^{(2)} \cdot \zeta^{2q}\right)^{-1}\right\}\right)$$

(for f_σ see Remark 1.6, for L Conditions $\mathcal{C}1'$) and put

$$\mathcal{T}(\zeta,\tau^{(3)}) = \left\{ \left\|\beta^{(1)}\right\| \le \zeta, \left\|\beta^{(2)}\right\| \le \zeta, \left\|\beta^{(1)} - \beta^{(2)}\right\| < \tau^{(3)} \right\}.$$

Then (remember that $\sup_{i\in\mathcal{N}} \sup_{r\in R} f_{e_i}(r) < f_\sigma$, see Remark 1.6) for any $n > n_3$ and any $\omega \in B^{(1)} \cap B^{(2)} \cap B^{(3)}$

$$\sup_{\left(\beta^{(1)},\beta^{(2)}\right)\in\mathcal{T}(\zeta,\tau^{(3)})} \frac{1}{n}\left|\sum_{i=1}^{n} w\left(\overline{F}_{n,\beta^{(1)}}(|r_i(\beta^{(2)})|)\right) - w\left(\overline{F}_{n,\beta^{(1)}}(|r_i(\beta^{(1)})|)\right)\right| \le L\cdot f_\sigma \cdot \tau^{(3)} \cdot \|X_i\|. \tag{43}$$

(For a sake of space write in a few next lines $w_{n,\beta^{(1)}}(i,\beta^{(2)})$ instead of $w(\overline{F}_{n,\beta^{(1)}}(|r_i(\beta^{(2)})|))$.) Employing Hőlder's inequality we arrive at (again for any $n > n_3$ and any $\omega \in B^{(1)} \cap B^{(2)} \cap B^{(3)}$)

$$\begin{aligned}
&\sup_{(\beta^{(1)},\beta^{(2)})\in\mathcal{T}(\zeta,\tau^{(3)})} \frac{1}{n}\Big|\sum_{i=1}^{n}\{w_{n,\beta^{(1)}}(i,\beta^{(2)}) - w_{n,\beta^{(1)}}(i,\beta^{(1)})\}\,[\beta^{(2)}]' X_i X_i' \beta^{(2)}\Big| \\
\le\ & \sup_{(\beta^{(1)},\beta^{(2)})\in\mathcal{T}(\zeta,\tau^{(3)})} \left\{\left[\frac{1}{n}\sum_{i=1}^{n}\left|w_{n,\beta^{(1)}}(i,\beta^{(2)}) - w_{n,\beta^{(1)}}(i,\beta^{(1)})\right|^{q'}\right]^{\frac{1}{q'}} \cdot \right. \\
&\qquad \left. \cdot \left[\frac{1}{n}\sum_{i=1}^{n}|X_i'\cdot\beta^{(2)}|^{2q}\right]^{\frac{1}{q}}\right\} \\
\le\ & \sup_{(\beta^{(1)},\beta^{(2)})\in\mathcal{T}(\zeta,\tau^{(3)})} \left\{\left[\frac{1}{n}\sum_{i=1}^{n}\left|w_{n,\beta^{(1)}}(i,\beta^{(2)}) - w_{n,\beta^{(1)}}(i,\beta^{(1)})\right|\right]^{\frac{1}{q'}} \cdot \right. \\
&\qquad \left. \cdot \left[\frac{1}{n}\sum_{i=1}^{n}\|\beta^{(2)}\|^{2q}\cdot\|X_i\|^{2q}\right]^{\frac{1}{q}}\right\} \\
\le\ & \sup_{(\beta^{(1)},\beta^{(2)})\in\mathcal{T}(\zeta,\tau^{(3)})} \left\{\left[\frac{1}{n}\sum_{i=1}^{n}\left|w_{n,\beta^{(1)}}(i,\beta^{(2)}) - w_{n,\beta^{(1)}}(i,\beta^{(1)})\right|\right]^{\frac{1}{q'}} \cdot \right. \\
&\qquad \left. \cdot\, \zeta^{2q}\left[\frac{1}{n}\sum_{i=1}^{n}\|X_i\|^{2q}\right]^{\frac{1}{q}}\right\} \\
\le\ & \sup_{(\beta^{(1)},\beta^{(2)})\in\mathcal{T}(\zeta,\tau^{(3)})} \left\{ f_\sigma^{\frac{1}{q'}} \cdot L^{\frac{1}{q'}} \cdot [\tau^{(3)}]^{\frac{1}{q'}} \cdot \left[\frac{1}{n}\sum_{i=1}^{n}\|X_i\|\right]^{\frac{1}{q'}} \cdot \zeta^{2q}\left[\frac{1}{n}\sum_{i=1}^{n}\|X_i\|^{2q}\right]^{\frac{1}{q}}\right\} \\
\le\ & \zeta^2 \cdot f_\sigma^{\frac{1}{q'}} \cdot L^{\frac{1}{q'}} \cdot [\tau^{(3)}]^{\frac{1}{q'}} \cdot [2\gamma^{(2)}]^{\frac{1}{q'}} \cdot [2\gamma^{(1)}]^{\frac{1}{q}} \le \frac{\delta}{10}
\end{aligned} \tag{44}$$

where the step from the fourth to fifth line used (43). Along similar lines we derive

$$\sup_{\left(\beta^{(1)},\beta^{(2)}\right)\in\mathcal{T}(\zeta,\tau^{(3)})} \frac{1}{n}\left|\sum_{i=1}^{n} w\left(\overline{F}_{n,\beta^{(1)}}(|r_i(\beta^{(1)})|)\right)\left\{\left[\beta^{(2)}\right]' X_i X_i' \beta^{(2)} - \left[\beta^{(1)}\right]' X_i X_i' \beta^{(1)}\right\}\right| \le \frac{\delta}{10}. \tag{45}$$

Finally, utilizing Lemma A.9 find $\tau^{(4)} \in (0, \min\left\{\delta/10, \tau^{(3)}\right\})$ so that for any pair $\beta^{(1)},\ \beta^{(2)}\imath R^p,\ \|\beta^{(1)}\| \leq \zeta,\ \|\beta^{(2)}\| \leq \zeta,\ \|\beta^{(1)} - \beta^{(2)}\| \leq \tau^{(4)}$, we have uniformly in $i \in \mathcal{N}$ and uniformly in $n \in \mathcal{N}$

$$\Big|[\beta^{(1)}]I\!\!E\left[w_{n,\beta^{(1)}}(i,\beta^{(1)})X_i\left(e_i - X_i^{'}\beta^{(1)}\right)\right] \\ -[\beta^{(2)}]^{'}I\!\!E\left[w_{n,\beta^{(2)}}(i,\beta^{(2)})X_i\left(e_i - X_i^{'}\beta^{(2)}\right)\right]\Big| \leq \frac{\delta}{10}. \quad (46)$$

where again $w_{n,\beta^{(\ell)}}(i,\beta^{(\ell)})$ was written instead of $w\left(\overline{F}_{n,\beta^{(\ell)}}(|r_i(\beta^{(\ell)})|)\right)$. Now find a system of open balls of type $\mathcal{B}(\beta,\tau^{(4)})$ covering the p-dimensional ball with center at zero and radius ζ, i. e. covering $\mathcal{B}(\zeta) = \{\beta \in R^p : \|\beta\| \leq \zeta\}$. Due to the compactness of $\mathcal{B}(\zeta)$ there is a subsystem of balls covering $\mathcal{B}(\zeta)$ which has finite number of balls, say $K(\zeta)$, and denote this system by $\left\{\mathcal{B}(\beta^{(j)},\tau^{(4)})\right\}_{j=1}^{K(\zeta)}$. Utilizing the law of large numbers find for any $j \in \{1, 2, \ldots, K(\zeta)\}$ some $n_j^* \in \mathcal{N}$ so that for all $n > n_j^*$ the set

$$B_{nj}^{(4)} = \left\{\omega \in \Omega : \frac{1}{n}\left\|\sum_{i=1}^{n}\left\{w_{n,\beta^{(j)}}(i,\beta^{(j)})X_iX_i' \right.\right.\right. \\ \left.\left.\left. -I\!\!E\left[w_{n,\beta^{(j)}}(i,\beta^{(j)})X_iX_i'\right]\right\}\right\| < \frac{\delta}{10\zeta^2}\right\} \quad (47)$$

has probability at least $1 - \frac{\varepsilon}{10K(\zeta)}$. Finally put $n_{\varepsilon,\delta,\zeta}^{(1)} = \max\{n_3, n_1^*, n_2^*, \ldots, n_{K(\zeta)}^*\}$ and $B_n = B_n^{(1)} \cap B_n^{(2)} \cap B_n^{(3)} \cap_{j=1}^{K(\zeta)} B_{nj}^{(4)}$. We have $P(B_n) > 1 - \frac{\varepsilon}{2}$. Since for any $\beta \in R^p, \|\beta\| \leq \zeta$ there is $j \in \{1, 2, \ldots, K(\zeta)\}$ so that $\left\|\beta - \beta^{(j)}\right\| < \tau^{(4)}$, taking into account (38), (40), (41), (44), (45), (46) and (47) we have for any $\omega \in B_n$ and any $n > n_{\varepsilon,\delta,\zeta}^{(1)}$

$$\sup_{\|\beta\|\leq\zeta}\frac{1}{n}\left|\beta^{'}\sum_{i=1}^{n}\left\{w\left(F_\beta^{(n)}(|r_i(\beta)|)\right)X_iX_i^{'} - I\!\!E\left[w\left(\overline{F}_{n,\beta}(|r_i(\beta)|)\right)X_iX_i^{'}\right]\right\}\beta\right| < \frac{\delta}{2}. \quad (48)$$

Now, we shall consider the second term in (35). Along similar lines as in the first part of the proof, we can find $n_{\varepsilon,\delta,\zeta}^{(2)} \in \mathcal{N}$ so that for any $n > n_{\varepsilon,\delta,\zeta}^{(2)}$ there is C_n so that $P(C_n) > 1 - \varepsilon/2$ and for any $\omega \in C_n$ we have

$$\sup_{\|\beta\|\leq\zeta}\frac{1}{n}\left|\sum_{i=1}^{n}\left\{w\left(F_\beta^{(n)}(|r_i(\beta)|)\right)e_iX_i^{'}\beta - I\!\!E\left[w\left(\overline{F}_{n,\beta}(|r_i(\beta)|)\right)e_iX_i^{'}\beta\right]\right\}\right| < \frac{\delta}{2}. \quad (49)$$

Put $n_{\varepsilon,\delta,\zeta} = \max\left\{n_{\varepsilon,\delta,\zeta}^{(1)}, n_{\varepsilon,\delta,\zeta}^{(2)}\right\}$. Then for any $n > n_{\varepsilon,\delta,\zeta}$ we have $P(B_n \cap C_n) > 1 - \varepsilon$ and taking into account (48) and (49), we conclude the proof. □

Similarly as in other situations when estimating (identifying) parameters of a model we need some *identification condition.* Prior to give it, let us prove:

Lemma 2.4. Let Conditions $\mathcal{C}2$ hold and moreover $\frac{1}{n}\sum_{i=1}^{n}|1-\sigma_i| = 0$. Finally, let e be a r.v. distributed according to $F_e(v)$ and for any $\beta \in R^p$ denote $F_\beta(v) = P(|e - X_1'\beta| < v)$. Then for any $\lambda > 0$

$$\lim_{n\to\infty} \sup_{-\infty<v<\infty} \sup_{\|\beta\|\leq\lambda} \left|\overline{F}_{n,\beta}(v) - F_\beta(v)\right| = 0. \tag{50}$$

P r o o f. First of all, notice please that $P(e_i < v) = P(e\sigma_i < v)$. We have to show that

$$\forall(\varepsilon > 0)\ \exists(n_\varepsilon \in \mathcal{N})\ \forall(n > n_\varepsilon): \qquad \sup_{-\infty<v<\infty} \sup_{\|\beta\|\leq\lambda} \left|\overline{F}_{n,\beta}(v) - F_\beta(v)\right| < \varepsilon.$$

So, let's fix an $\varepsilon > 0$ and recall that

$$\overline{F}_{n,\beta}(v) = \frac{1}{n}\sum_{i=1}^{n} F_{\beta,i}(v),$$

$$F_{\beta,i}(v) = P\left(|e_i - X_i'\beta| < v\right) = \int_{\{-v<r-x'\beta<v\}} \mathrm{d}F_X(x) f_{e_i}(r)\mathrm{d}r$$

$$= \int_{x\in R^p}\left\{\int_{\{-v+x'\beta<r<v+x'\beta\}} f_e(r\sigma_i^{-1})\sigma_i^{-1}\mathrm{d}r\right\}\mathrm{d}F_X(x) \tag{51}$$

and

$$F_\beta(v) = \int_{\{-v<r-x'\beta<v\}} \mathrm{d}F_X(x) f_e(r)\mathrm{d}r$$

$$= \int_{x\in R^p}\left\{\int_{\{-v+x'\beta<r<v+x'\beta\}} f_e(r)\mathrm{d}r\right\}\mathrm{d}F_X(x). \tag{52}$$

Let us put for any $\sigma > 0$ $\quad F_{\beta,\sigma}(v) = P(|e\sigma - X_1'\beta| < v)$. Then due to absolute continuity of $F_e(v)$, we have

$$F_{\beta,\sigma}(v) = P\left(|e\sigma - X_1'\beta| < v\right) = \int_{x\in R^p}\left\{\int_{\{-v+x'\beta<r<v+x'\beta\}} f_e(r\sigma^{-1})\sigma^{-1}\mathrm{d}r\right\}\mathrm{d}F_X(x) \tag{53}$$

is continuous and hence, for any $\beta \in R^p$ and any $\sigma > 0$, there is $v_{\beta,\sigma} > 0$ so that

$$F_{\beta,\sigma}(v_{\beta,\sigma,\varepsilon}) = 1 - \frac{\varepsilon}{2}. \tag{54}$$

Put

$$v_{u,\varepsilon}^* = \sup_{\|\beta\|\leq\lambda} \sup_{s_\sigma\leq\sigma\leq S_\sigma} v_{\beta,\sigma,\varepsilon}. \tag{55}$$

Generally we can have $v_{u,\varepsilon}^* = \infty$. But, taking into account that $\{\|\beta\| \leq \lambda\} \times [s_\sigma, S_\sigma]$ is compact, using standard arguments we find $(\beta_u, \sigma_u) \in \{\|\beta\| \leq \lambda\} \times [s_\sigma, S_\sigma]$ so that

$$F_{\beta_u,\sigma_u}(v_{u,\varepsilon}^*) = 1 - \frac{\varepsilon}{2}. \tag{56}$$

Hence $0 \le v^*_{u,\varepsilon} < \infty$ and for any $\beta \in \{\|\beta\| \le \lambda\}$ and any $i = 1, 2, \ldots$

$$1 - \frac{\varepsilon}{2} = F_{\beta_u,\sigma_u}(v^*_{u,\varepsilon}) \le F_{\beta,i}(v^*_{u,\varepsilon}). \tag{57}$$

Finally, find v^*_ε do that $F_\beta(v^*_\varepsilon) = 1 - \frac{\varepsilon}{2}$, put $v_{u,\varepsilon} = \max\left\{v^*_{u,\varepsilon}, v^*_\varepsilon\right\}$ and keep in mind that $F_{\beta,i}(0) = 0$ for all $i = 1, 2, \ldots$ as well as $F_\beta(0) = 0$. Then for any $\beta \in \{\|\beta\| \le \lambda\}$, any $n = 1, 2, \ldots$ and any $v \in (-\infty, 0] \cup [v_{u,\varepsilon}, \infty)$

$$\left|\overline{F}_{n,\beta}(v) - F_\beta(v)\right| \le \varepsilon. \tag{58}$$

Now, employing substitution $y = r \cdot \sigma_i$, we obtain from (51)

$$F_{\beta,i}(v) = \int_{x\in R^p} \left\{ \int_{\{\frac{-v+x'\beta}{\sigma_i} < y < \frac{v+x'\beta}{\sigma_i}\}} f_e(y)\mathrm{d}y \right\} \mathrm{d}F_X(x).$$

Then

$$|P\left(|e - X_1'\beta| < v\right) - F_{\beta,i}(v)| \le f_\sigma \int_{x\in R^p} \left\{ \int_{a_i}^{b_i} \mathrm{d}r + \int_{c_i}^{d_i} \mathrm{d}r \right\} \mathrm{d}F_X(x)$$

where $a_i = \min\{\frac{-v+x'\beta}{\sigma_i}, -v+x'\beta\}$, $b_i = \max\{\frac{-v+x'\beta}{\sigma_i}, -v+x'\beta\}$, $c_i = \min\{\frac{v+x'\beta}{\sigma_i}, v+x'\beta\}$ and $d_i = \max\{\frac{v+x'\beta}{\sigma_i}, v+x'\beta\}$. It means that $\left|a_i - b_i\right| \le |v + x'\beta| \cdot \left|\frac{1}{\sigma_i} - 1\right| \le \frac{\left|v+x'\beta\right|}{s_\sigma} \cdot |1 - \sigma_i|$. It gives

$$|P\left(|e - X_1'\beta| < v\right) - F_{\beta,i}(v)| \le 2{\cdot}f_\sigma\,|v + \mathbb{E}X_1'\beta|{\cdot}\left|\frac{1-\sigma_i}{\sigma_i}\right| \le 2{\cdot}f_\sigma\,|v + \mathbb{E}X_1'\beta|{\cdot}\frac{|1-\sigma_i|}{s_\sigma}.$$

Then

$$\sup_{-\infty<v<\infty}\ \sup_{\|\beta\|\le\lambda} |\overline{F}_{n,\beta}(v) - F_\beta(v)| \le 2 \cdot f_\sigma \frac{[v_{u,\varepsilon} + \lambda\mathbb{E}\|X_1\|]}{s_\sigma} \cdot \frac{1}{n}\sum_{i=1}^n |1 - \sigma_i|$$

and the proof follows. □

Lemma 2.5. Let Conditions $\mathcal{C}1'$ and $\mathcal{C}2$ be fulfilled. Let again e be a r.v. distributed according to $F_e(v)$ and denote for any $\beta \in R^p$ $F_\beta(v) = P(|e - X_1'\beta| < v)$ and $r(\beta) = e - X_1'\beta$. Finally, let $\lim_{n\to\infty} \frac{1}{n}\sum_{i=1}^n \sigma_i = 1$. Then for any $\lambda > 0$

$$\lim_{n\to\infty}\ \sup_{\|\beta\|\le\lambda} \beta' \left\{ \frac{1}{n}\sum_{i=1}^n \mathbb{E}\left[w\left(\overline{F}_{n,\beta}(|r_i(\beta)|)\right) X_i\left(e_i - X_i'\beta\right)\right]\right.$$
$$\left.\mathbb{E}\left[w\left(F_\beta(|r(\beta)|)\right) X_1\left(e - X_1'\beta\right)\right]\right\} = 0.$$

P r o o f employs Lemma 2.4 and similar technical steps as the proof of Lemma 2.3. □

Corollary 2.6. Let Conditions $\mathcal{C}1'$ and $\mathcal{C}2$ be fulfilled. Moreover, let $\lim_{i\to\infty} \frac{1}{n}\sum_{i=1}^{n}\sigma_i = 1$. Then for any $\varepsilon > 0$, $\delta \in (0,1)$ and $\zeta > 0$ there is $n_{\varepsilon,\delta,\zeta} \in \mathcal{N}$ so that for any $n > n_{\varepsilon,\delta,\zeta}$ we have

$$P\left(\left\{\omega \in \Omega \ : \ \sup_{\|\beta\|\le\zeta} \left|\frac{1}{n}\sum_{i=1}^{n} w\left(F_\beta^{(n)}(|r_i(\beta)|)\right)\beta' X_i\left(e_i - X_i'\beta\right)\right.\right.\right.$$
$$\left.\left.\left. -\beta' I\!\!E\left[w\left(F_\beta(|r(\beta)|)\right)X_1\left(e - X_1'\beta\right)\right]\right| < \delta\right\}\right) > 1-\varepsilon.$$

P r o o f follows from Lemma 2.3 and 2.5. □

Conditions $\mathcal{C}3$. There is the only solution of

$$I\!\!E\left[w\left(F_\beta(|r(\beta)|)\right)X_1\left(e - X_1'\beta\right)\right] = 0 \tag{59}$$

namely $\beta^0 = 0$ (the equation (59) is assumed as a vector equation in $\beta \in R^p$). Moreover $\lim_{n\to\infty}\frac{1}{n}\sum_{i=1}^{n}\sigma_i = 1$.

Remark 2.7. For $w(u) \equiv 1$, i. e. for the (Ordinary) Least Squares, (59) is fulfilled as the normal equations have the only solution, namely the orthogonal projection of $Y = (Y_1, Y_2, \ldots, Y_n)'$ into the linear envelope of the columns of matrix $X = (X_1, X_2, \ldots, X_n)'$.

Theorem 2.8. Let Conditions $\mathcal{C}1'$, $\mathcal{C}2$ and $\mathcal{C}3$ be fulfilled. Then any sequence $\{\hat\beta^{(\mathrm{LWS},n,w)}\}_{n=1}^{\infty}$ of the solutions of sequence of normal equations $I\!\!N E_{Y,X,n}(\hat\beta^{(\mathrm{LWS},n,w)}) = 0$, $n = 1, 2, \ldots$, is weakly consistent.

P r o o f. To prove the consistency of $\{\hat\beta^{(\mathrm{LWS},n,w)}\}_{n=1}^{\infty}$, we have to show that for any $\varepsilon > 0$ and $\delta > 0$ there is $n_{\varepsilon,\delta} \in \mathcal{N}$ such that for all $n > n_{\varepsilon,\delta}$

$$P\left(\left\{\omega \in \Omega : \left\|\hat\beta^{(\mathrm{LWS},n,w)} - \beta^0\right\| < \delta\right\}\right) > 1-\varepsilon. \tag{60}$$

So fix $\varepsilon_1 > 0$ and $\delta_1 > 0$ and recall that $I\!\!N E_{Y,X,n}(\beta) = \sum_{i=1}^{n} w(F_\beta^{(n)}(|r_i(\beta)|)) \cdot$ $\cdot \beta' X_i(e_i - X_i'\beta)$.

According to Lemma 2.2 there are $\Delta_1 > 0$ and θ_1 so that for ε_1 there is $n_{\Delta_1,\varepsilon_1} \in \mathcal{N}$ so that for any $n > n_{\Delta_1,\varepsilon_1}$

$$P\left(\left\{\omega \in \Omega : \inf_{\|\beta\|\ge\theta_1} -\frac{1}{n}\beta' I\!\!N E_{Y,X,n}(\beta) > \Delta_1\right\}\right) > 1 - \frac{\varepsilon_1}{2} \tag{61}$$

(denote the corresponding set by B_n). It means that for all $n > n_{\Delta_1,\varepsilon_1}$ all solutions of the normal equations $I\!\!N E_{Y,X,n}(\beta) = 0$ with probability at least $1 - \frac{\varepsilon_1}{2}$ are inside the ball $\mathcal{B}(0, \theta_1)$.

Further, consider the compact set $C(\delta_1, \theta_1) = \{\beta \in R^p : \delta_1 \le \|\beta\| \le \theta_1\}$ and find

$$\tau_{C(\delta_1,\theta_1)} = \inf_{\beta \in C(\delta_1,\theta_1)} \left\{ -\beta^{'} I\!\!E \left[w\left(F_\beta(|r(\beta)|)\right) X_1 \left(e - X_1^{'}\beta \right) \right] \right\}.$$

Assume that $\tau_{C(\delta_1,\theta_1)} = 0$. Due to compactness of $C(\delta_1, \theta_1)$, there is a $\{\beta_k\}_{k=1}^{\infty} \subset C(\delta_1, \theta_1)$ such that

$$\lim_{k\to\infty} \beta_k^{'} I\!\!E \left[w\left(F_\beta(|r(\beta)|)\right) X_1 \left(e - X_1^{'}\beta \right) \right] = -\tau_{C(\delta_1,\theta_1)}.$$

Also due to compactness of $C(\delta_1, \theta_1)$, there is a $\bar\beta \in C(\delta_1, \theta_1)$ and a subsequence $\left\{\beta_{k_j}\right\}_{j=1}^{\infty}$ such that

$$\lim_{j\to\infty} \beta_{k_j} = \bar\beta$$

(where the convergence is assumed coordinatewise) and due to the continuity of

$$\beta^{'} I\!\!E \left[w\left(F_\beta(|r(\beta)|)\right) X_1 \left(e - X_1^{'}\beta \right) \right]$$

(see Lemma A.10) we have

$$-\bar\beta' I\!\!E \left[w\left(F_{\bar\beta}(|r(\bar\beta)|)\right) X_1 \left(e - X_1^{'}\bar\beta \right) \right] = \tau_{C(\delta_1,\theta_1)} = 0. \tag{62}$$

Then, Condition $\mathcal{C}3$ implies that $\left|\tau_{C(\delta_1,\theta_1)}\right| \neq 0$.

Now, put $\Delta = \min\left\{ \frac{\left|\tau_{C(\delta_1,\theta_1)}\right|}{2}, \frac{\Delta_1}{2} \right\}$ and utilizing Corollary 2.6 we may find for ε_1, δ_1, θ_1 and Δ such $n_{\varepsilon_1,\delta_1,\theta_1,\Delta} \in \mathcal{N}$ that $n_{\varepsilon_1,\delta_1,\theta_1,\Delta} \ge n_{\varepsilon_1,\delta,\theta_1}$ and for any $n > n_{\varepsilon_1,\delta_1,\theta_1,\Delta}$ there is a set D_n (with $P(D_n) > 1 - \frac{\varepsilon}{2}$) such that for any $\omega \in D_n$

$$\sup_{\|\beta\|\le\theta_1} \left| \frac{1}{n}\sum_{i=1}^{n} w\left(F_\beta^{(n)}(|r_i(\beta)|)\right) \beta^{'} X_i \left(e_i - X_i^{'}\beta \right) \right.$$
$$\left. -\beta^{'} I\!\!E \left[w\left(F_\beta(|r(\beta)|)\right) X_1 \left(e_i - X_1^{'}\beta \right) \right] \right| < \Delta. \tag{63}$$

But (61) and (63) imply that for any $\beta \in R^p, \|\beta\| = \theta_1$ $I\!\!E[w(F_\beta(|r(\beta)|))X_1(e_i - X_1^{'}\beta)] > \Delta$. If then $\tau_{C(\delta_1,\theta_1)} < 0$ there would be a solution of equation (59) inside the compact $C(\delta_1, \theta_1) = \{\beta \in R^p : \delta_1 \le \|\beta\| \le \theta_1\}$. Hence $\tau_{C(\delta_1,\theta_1)} > 0$ (and hence also $\Delta > 0$) and for any $n > n_{\varepsilon_1,\delta_1,\theta_1,\Delta}$ and any $\omega \in B_n \cap D_n$ we have

$$\inf_{\|\beta\|>\delta_1} -\frac{1}{n}\beta^{'} I\!\!N E_{Y,X,n}(\beta) > \Delta. \tag{64}$$

Clearly, $P\left(B_n \cap D_n\right) > 1 - \varepsilon_1$. But it means that all solutions of normal equations (13) are inside the ball of radius δ_1 with probability at least $1 - \varepsilon_1$, i.e. in other words, $\hat\beta^{(\mathrm{LWS},n,w)}$ is weakly consistent. $\square$

3. CONCLUDING REMARKS

As we have already said, the results allow to establish the robustified version of covariance matrix of the estimates by LWS resistant to heteroscadasticity (as a generalization of White estimator of this matrix for OLS) which in turn enable us to make right conclusion about significance of explanatory variables. Empolying them, we can also proceed in study of robustified versions of diagnostic tools and sensitivity characteristics for LWS[2] analogous to the tools and characteristics used by classical econometrics for the OLS.

The results were derived – due to the fact that we assumed the linear regression framework – by simple methods under weak assumptions, usually imposed on corresponding entities in the regression framework. Moreover a brief discussion included them into up to now obtained results on robust regression. Of course, strengthening a bit assumptions would allow to employ results by Vaart, Welner [15] or Koul [9] on empirical processes. Our approach may appear more suitable as the forthcoming research will assume further modifications of the basic method of LWS – in the sense in which econometrics developed a lot of modifications of OLS for regression model for variety of (economic) types of data (e. g. ARCH model) and (economic) frameworks (e. g. errors-in-variables model, limited response variable, etc.).

4. APPENDIX

We need to recall some (general) results.

Lemma A.1. (Štěpán [13], page 420, VII.2.8) Let a and b be positive numbers. Further let ξ be a random variable such that $P(\xi = -a) = \pi$ and $P(\xi = b) = 1 - \pi$ (for a $\pi \in (0,1)$) and $\mathbb{E}\xi = 0$. Moreover let τ be the time for the Wiener process $W(s)$ to exit the interval $(-a, b)$. Then

$$\xi =_{\mathcal{D}} W(\tau)$$

where "$=_{\mathcal{D}}$" denotes the equality of distributions of the corresponding random variables. Moreover, $\mathbb{E}\tau = a \cdot b = var\, \xi$.

Remark A.2. Since the book by Štěpán [13] is in Czech language we refer also to Breiman [2] where however this assertion is not isolated. Nevertheless, the assertion can be found directly in the first lines of the proof of Proposition 13.7 (page 277) of Breiman's book. (See also Theorem 13.6 on the page 276.) The next assertion can be found, in a bit modified form also in Breiman's book, Proposition 12.20 (page 258).

Lemma A.3. (Štěpán [13], page 409, VII.1.6) Let a and b be positive numbers. Then

$$P\left(\max_{0 \le t \le b} |W(t)| > a\right) \le 2 \cdot P\left(|W(b)| > a\right).$$

[2]Some of these studies will require, of course, to derive asymptotic representation (and possibly asymptotic normality) of LWS.

Definition A.4. Let S be a subset of a separable metric space. The stochastic process $V = (V(s),\ s \in S)$ is called *separable* if there is a countable dense subset $T \subset S$ (i. e. T is countable and dense in S) such that for any $(\omega, s) \in \Omega \times S$ there is a sequence such that

$$s_n \in T, \quad \lim_{n\to\infty} s_n = s \quad \text{and} \quad \lim_{n\to\infty} V(\omega, s_n) = V(\omega, s).$$

Lemma A.5. (Štěpán [13], page 85, I.10.4) Let $V = (V(s),\ s \in S)$ be a *separable* stochastic process defined on the probability space $(\Omega, \mathcal{A}, P)$. Moreover, let $G \subset S$ be open and denote by $k(G)$ the set of all finite subsets of G. Then for any close set $K \subset R^p$ we have

$$\{\omega \in \Omega : V(s) \in K, s \in G\} \in \mathcal{A}$$

and

$$P\left(\{\omega \in \Omega : V(s) \in K, s \in G\}\right) = \inf_{J \in k(G)} P\left(\{\omega \in \Omega : V(s) \in K, s \in J\}\right).$$

P r o o f. Since the book by Štěpán is in Czech language and the proof is short, we will give it. Let T be countable dense subset of S. Then we have

$$\{\omega \in \Omega : V(s) \in K, s \in G\} = \{\omega \in \Omega : V(s) \in K, s \in G \cap T\}$$

and

$$\begin{aligned} & P\left(\{\omega \in \Omega : V(s) \in K, s \in G\}\right) \le \inf_{J \in k(G)} P\left(\{\omega \in \Omega : V(s) \in K, s \in J\}\right) \\ \le \quad & \inf_{J \in k(G \cap S)} P\left(\{\omega \in \Omega : V(s) \in K, s \in J\}\right) = P\left(\{\omega \in \Omega : V(s) \in K, s \in G \cap S\}\right) \\ = \quad & P\left(\{\omega \in \Omega : V(s) \in K, s \in G\}\right). \end{aligned}$$

□

Let's recall that we have denoted in (14) the d.f. of $(X_1'\beta)^2$ by $F_{(X'\beta)^2}(u)$ and in (15) the corresponding empirical d.f. by $F^{(n)}_{(X'\beta)^2}(u)$, i. e.

$$F^{(n)}_{(X'\beta)^2}(u) = \frac{1}{n} \sum_{i=1}^{n} I\left\{(X_i'\beta)^2 < u\right\}. \tag{A.65}$$

Lemma A.6. Let the Conditions $\mathcal{C}2$ hold. For any $\varepsilon > 0$ there is a constant K_ε and $n_\varepsilon \in \mathcal{N}$ so that for all $n > n_\varepsilon$

$$P\left(\left\{\omega \in \Omega : \sup_{v \in R^+} \sup_{\beta \in R^p} \sqrt{n} \left| F^{(n)}_{(X'\beta)^2}(u) - F_{(X'\beta)^2}(u) \right| < K_\varepsilon \right\}\right) > 1 - \varepsilon. \tag{A.66}$$

P r o o f. Fix $\varepsilon > 0$ and put $K_\varepsilon = \sqrt{\frac{8}{\varepsilon}} + 1$ together with

$$b_i(u, \beta) = I\left\{\omega \in \Omega : (X_i'\beta)^2 < u\right\}. \tag{A.67}$$

Further put

$$\xi_i(u,\beta) = b_i(u,\beta) - \mathbb{E} b_i(u,\beta) \tag{A.68}$$

and denote

$$\pi_i(u,\beta) = \mathbb{E} b_i(u,\beta) = P\left(b_i(u,\beta) = 1\right) = F_{\left(X_i'\beta\right)^2}(u). \tag{A.69}$$

Then $\{\xi_i(u,\beta)\}_{i=1}^{\infty}$, for any $u \in R^+$ and any $\beta \in R^p$, is a sequence of independently distributed r.v.'s. Finally, (A.65), (A.67) and (A.69) yield

$$\frac{1}{n}\sum_{i=1}^{n} \xi_i(u,\beta) = F^{(n)}_{(X'\beta)^2}(u) - F_{(X'\beta)^2}(u),$$

i. e.

$$\frac{1}{\sqrt{n}}\left|\sum_{i=1}^{n} \xi_i(u,\beta)\right| = \sqrt{n}\left|F^{(n)}_{(X'\beta)^2}(u) - F_{(X'\beta)^2}(u)\right|.$$

Moreover

$$P\left(\xi_i(u,\beta) = 1 - \pi_i(u,\beta)\right) = \pi_i(u,\beta)$$

and

$$P\left(\xi_i(u,\beta) = -\pi_i(u,\beta)\right) = 1 - \pi_i(u,\beta).$$

Now, we are going to employ Lemma A.1. We have already mentioned that $\{\xi_i(u,\beta)\}_{i=1}^{\infty}$ is a sequence of independently distributed r.v.'s. Let us denote by $\{W_i(s)\}_{i=1}^{\infty}$ the sequence of independent Wiener processes (we may assume e. g. that each of them is defined on "an own probability space", say $\{(\Omega_i, \mathcal{A}_i, P_i)\}_{i=1}^{\infty}$ and then consider the product space $(\Omega, \mathcal{A}, P)$ in the same way as it is done in the proof of Daniell-Kolmogorov theorem, see e. g. Tucker [14] and let us define $\tau_i(u,\beta)$ to be the time for the Wiener process $W_i(s)$ to exit the interval $(-\pi_i(u,\beta), 1-\pi_i(u,\beta))$ (please keep in mind that $\tau_i(u,\beta)$ is r.v., i. e. $\tau_i(u,\beta) = \tau_i(u,\beta,\omega)$). Then $\xi_i(u,\beta) =_{\mathcal{D}} W_i(\tau_i(u,\beta))$ and hence for any $\beta \in R^p$

$$n^{-\frac{1}{2}}\sum_{i=1}^{n} \xi_i(u,\beta) =_{\mathcal{D}} n^{-\frac{1}{2}}\sum_{i=1}^{n} W_i(\tau_i(u,\beta)) =_{\mathcal{D}} W_1\left(n^{-1}\sum_{i=1}^{n} \tau_i(u,\beta)\right) \tag{A.70}$$

where the last equality follows from the properties of the Wiener process. Further, let us define U_i to be the time for the Wiener process $W_i(s)$ to exit interval $(-1,1)$. Due to the fact that for all $i = 1, 2, \ldots, n$ for any $u \in R^+$ and any $\beta \in R^p$

$$\pi_i(u,\beta) \le 1 \ \text{ and } \ 1-\pi_i(u,\beta) \le 1, \qquad \text{i. e.} \quad (-\pi_i(u,\beta)), 1-\pi_i(u,\beta)) \subset (-1,1),$$

we conclude that for any $u \in R^+$, any $\beta \in R^p$ and any $\omega \in \Omega$

$$\tau_i(u,\beta) \le U_i$$

and hence (again for any $\omega \in \Omega$)

$$n^{-1}\sum_{i=1}^{n} \tau_i(u,\beta) \le n^{-1}\sum_{i=1}^{n} U_i. \tag{A.71}$$

Of course, $\{U_i\}_{i=1}^{\infty}$ is the sequence of i.i.d r.v.'s and due to Lemma A.1 we have

$$I\!E U_i = 1,$$

so, employing the law of large numbers, we can find n_1 so that for all $n > n_1$ and for

$$B_n = \left\{ \omega \in \Omega \ : \ n^{-1} \sum_{i=1}^{n} U_i \leq 2 \right\}$$

we have

$$P(B_n) \geq 1 - \frac{\varepsilon}{2}. \tag{A.72}$$

Let us consider $n > n_1$ and a fix $\omega_0 \in B_n$ and let us realize that for any $u \in R^+$ and any $\beta \in R^p$ the left hand side of (A.71), i. e. $n^{-1} \sum_{i=1}^{n} \tau_i(u,\beta) = n^{-1} \sum_{i=1}^{n} \tau_i(u,\beta,\omega_0)$, is not larger than $n^{-1} \sum_{i=1}^{n} U_i = n^{-1} \sum_{i=1}^{n} U_i(\omega_0) \in [0,2]$. So for our fix ω_0, we have

$$\left\{ t \in R : t = n^{-1} \sum_{i=1}^{n} \tau_i(v,\beta,\omega_0), v \in R^+, \beta \in R^p \right\} \subset \left\{ t \in R : 0 \leq t \leq n^{-1} \sum_{i=1}^{n} U_i(\omega_0) \right\}.$$

It means that

$$\sup_{v \in R^+} \sup_{\beta \in R^p} W\left(n^{-1} \sum_{i=1}^{n} \tau_i(v,\beta,\omega_0) \right) \leq \sup_{0 \leq t \leq n^{-1} \sum_{i=1}^{n} U_i(\omega_0)} \left| W_1(t,\omega_0) \right|. \tag{A.73}$$

So, we arrived at: We have two processes which are equivalent in distribution, i. e.

$$\sum_{i=1}^{n} \xi_i(u,\beta,\omega) =_{\mathcal{D}} W_1\left(n^{-1} \sum_{i=1}^{n} \tau_i(u,\beta,\omega) \right)$$

with the same index sets, $u \in R, \beta \in R^p$ (see (A.70)), both of them are separable. Then employing Lemma A.5, we obtain

$$n^{-\frac{1}{2}} \sup_{u \in R^+} \sup_{\beta \in R^p} \left| \sum_{i=1}^{n} \xi_i(u,\beta,\omega_0) \right| =_{\mathcal{D}} \sup_{u \in R^+} \sup_{\beta \in R^p} \left| W_1\left(n^{-1} \sum_{i=1}^{n} \tau_i(u,\beta,\omega_0) \right) \right|$$

and due to (A.73)

$$n^{-\frac{1}{2}} \sup_{u \in R^+} \sup_{\beta \in R^p} \left| \sum_{i=1}^{n} \xi_i(u,\beta,\omega_0) \right| \leq \sup_{0 \leq t \leq n^{-1} \sum_{i=1}^{n} U_i(\omega_0)} \left| W_1(t,\omega_0) \right|.$$

In other words, for any $n > n_1$ and any $\omega \in B_n$

$$n^{-\frac{1}{2}} \sup_{u \in R^+} \sup_{\beta \in R^p} \left| \sum_{i=1}^{n} \xi_i(u,\beta) \right| \leq \sup_{0 \leq t \leq n^{-1} \sum_{i=1}^{n} U_i} \left| W_1(t) \right|. \tag{A.74}$$

Further, employing (A.74), we arrive at

$$P\left(\left\{\omega \in \Omega \ : \ n^{-\frac{1}{2}} \sup_{u\in R^+} \sup_{\beta\in R^p} \left|\sum_{i=1}^n \xi_i(u,\beta)\right| > K\right\}\right)$$

$$\leq P\left(\left\{\omega \in \Omega : n^{-\frac{1}{2}} \sup_{u\in R^+} \sup_{\beta\in R^p} \left|\sum_{i=1}^n \xi_i(u,\beta)\right| > K\right\} \cap \left\{\omega \in \Omega \ : \ n^{-1}\sum_{i=1}^n U_i > 2\right\}\right)$$

$$+P\left(\left\{\omega \in \Omega : \sup_{0\leq t\leq n^{-1}\sum_{i=1}^n U_i} |W_1(t)| > K\right\} \cap \left\{\omega \in \Omega \ : \ n^{-1}\sum_{i=1}^n U_i \leq 2\right\}\right)$$

$$\leq P\left(\left\{\omega \in \Omega : n^{-1}\sum_{i=1}^n U_i > 2\right\}\right) + P\left(\left\{\omega \in \Omega : \sup_{0\leq t\leq 2} |W_1(t)| > K\right\}\right). \quad \text{(A.75)}$$

Now, utilizing Lemma A.3, we obtain

$$P\left(\sup_{0\leq t\leq 2} |W_1(t)| > K\right) \leq 2\cdot P\left(|W_1(2)| > K\right). \quad \text{(A.76)}$$

Further, recalling the fact that $\operatorname{var}\{W(2)\} = 2$ and using Chebyshev's inequality, we arrive at

$$2\cdot P\left(|W_1(2)| > K\right) \leq 4\cdot \frac{1}{K^2} = \frac{\varepsilon}{2}. \quad \text{(A.77)}$$

Finally, (A.72), (A.75), (A.76) and (A.77) imply

$$P\left(n^{-\frac{1}{2}} \sup_{u\in R^+,\ \beta\in R^p} \left|\sum_{i=1}^n \xi_i(u,\beta)\right| > K\right) \leq \varepsilon$$

which concludes the proof. □

Let's recall that we have denoted by $F_\beta^{(n)}(v)$ the empirical d. f. of error terms e_i's, i. e.

$$F_\beta^{(n)}(v) = \frac{1}{n}\sum_{i=1}^n I\left\{|e_i - X_i'\beta| < v\right\}$$

and that we have put

$$\overline{F}_{n,\beta}(v) = \frac{1}{n}\sum_{i=1}^n F_{\beta,i}(v)$$

(see (33)) where

$$F_{\beta,i}(v) = P\left(|Y_i - X_i'\beta| < v\right) = P\left(|e_i - X_i'\beta| < v\right).$$

Lemma A.7. Let the Conditions $\mathcal{C}2$ hold. For any $\varepsilon > 0$ there is a constant K_ε and $n_\varepsilon \in \mathcal{N}$ so that for all $n > n_\varepsilon$

$$P\left(\left\{\omega \in \Omega : \sup_{v\in R^+} \sup_{\beta\in R^p} \sqrt{n}\left|F_\beta^{(n)}(v) - \overline{F}_{n,\beta}(v)\right| < K_\varepsilon\right\}\right) > 1 - \varepsilon.$$

For a Proof of the lemma see Víšek [25] (the proof runs along similar lines as the proof of the previous lemma).

Lemma A.8. Under Conditions $\mathcal{C}2$ the distribution functions $F_{\beta,i}(r)$ and $F_{(X'\beta)^2}(r)$ are, uniformly in $i = 1, 2, \ldots$ and in $r \in R$, uniformly continuous in β, i. e. for any $\delta > 0$ there is $\zeta \in (0,1)$ so that for any pair $\beta^{(1)}$ and $\beta^{(2)}$ such that $\left\|\beta^{(1)} - \beta^{(2)}\right\| < \zeta$ we have

$$\sup_{i\in\mathcal{N}} \sup_{r\in R} \left|F_{\beta^{(1)},i}(r) - F_{\beta^{(2)},i}(r)\right| \le \delta$$

and

$$\sup_{r\in R} \left|F_{\left(X'\beta^{(1)}\right)^2}(r) - F_{\left(X'\beta^{(2)}\right)^2}(r)\right| \le \delta.$$

Proof. Let us recall that (see (34))

$$F_{\beta,i}(r) = P\left(\left|e_i - X_i^{'}\beta\right| < r\right) = \int I\{\left|s - x^{'}\beta\right| < r\} \mathrm{d}F_{X,e_i}(x,s)$$

and that (under Conditions $\mathcal{C}2$) there is $f_\sigma < \infty$ so that $\sup_{i\in\mathcal{N}} \sup_{r\in R} f_{e_i}(r) < f_\sigma$. Then

$$\begin{aligned}
& \sup_{i\in\mathcal{N}} \sup_{r\in R} \left|F_{\beta^{(1)},i}(r) - F_{\beta^{(2)},i}(r)\right| \\
\le\ & \sup_{i\in\mathcal{N}} \sup_{r\in R} \int \left|I\{|s - x^{'}\beta^{(1)}| < r\} - I\{|s - x^{'}\beta^{(2)}| < r\}\right| \mathrm{d}F_{X,e_i}(x,s) \\
=\ & \sup_{i\in\mathcal{N}} \sup_{r\in R} \int \left|I\{|s - x^{'}\beta^{(1)}| < r\} - I\{|s - x^{'}\beta^{(2)}| < r\}\right| f_{e_i}(s)\mathrm{d}s\, \mathrm{d}F_X(x).
\end{aligned}$$

Further

$$\begin{aligned}
& \int \left|I\left\{|s - x^{'}\beta^{(1)}| < r\right\} - I\left\{|s - x^{'}\beta^{(2)}| < r\right\}\right| f_{e_i}(s)\mathrm{d}s \\
\le\ & \int_{\min\{-r+x^{'}\beta^{(1)},-r+x^{'}\beta^{(2)}\}}^{\max\{-r+x^{'}\beta^{(1)},-r+x^{'}\beta^{(2)}\}} f_{e_i}(s)\mathrm{d}s + \int_{\min\{r+x^{'}\beta^{(1)},r+x^{'}\beta^{(2)}\}}^{\max\{r+x^{'}\beta^{(1)},r+x^{'}\beta^{(2)}\}} f_{e_i}(s)\mathrm{d}s \\
\le\ & 2 \cdot f_\sigma \cdot \left|x^{'}\beta^{(1)} - x^{'}\beta^{(2)}\right|.
\end{aligned}$$

Hence

$$\begin{aligned}
& \sup_{i\in\mathcal{N}} \sup_{r\in R} \left|F_{\beta^{(1)},i}(r) - F_{\beta^{(2)},i}(r)\right| \le 2 \cdot f_\sigma \int \left|x^{'}\beta^{(1)} - x^{'}\beta^{(2)}\right| f_X(x)\mathrm{d}x \\
\le\ & 2 \cdot f_\sigma \cdot I\!E \left\|X_1\right\| \cdot \left\|\beta^{(1)} - \beta^{(2)}\right\|.
\end{aligned}$$

So, for any $\delta > 0$, putting $\zeta = \frac{1}{2}\delta \cdot f_\sigma^{-1} \cdot I\!E^{-1} \left\|X_1\right\|$, for any $\beta^{(1)}, \beta^{(2)} \in R^p$, $\left\|\beta^{(1)} - \beta^{(2)}\right\| \le \zeta$ we have

$$\sup_{r\in R} \left|F_{\beta^{(1)},i}(r) - F_{\beta^{(2)},i}(r)\right| \le \delta.$$

The proof of the second part of the lemma runs along similar lines. □

Lemma A.9. Let Conditions $\mathcal{C}1$ and $\mathcal{C}2$ hold. Then for any positive ζ

$$\beta^{'} I\!\!E \left[w \left(\overline{F}_{n,\beta}(|r_i(\beta)|) \right) X_i \left(e_i - X_i^{'}\beta \right) \right]$$

is uniformly in $i = 1, 2, \dots$ and uniformly in $n = 1, 2, \dots$ uniformly continuous in β on $\mathcal{B} = \{\beta \in R^p : \|\beta\| \le \zeta\}$.

P r o o f. Fix a positive ζ and ε and for the sake of space write again in a few next lines $w_{n,\beta^{(1)}}(i, \beta^{(2)})$ instead of $w\left(\overline{F}_{n,\beta^{(1)}}(|r_i(\beta^{(2)})|)\right)$. We have to show that then there is $\delta_{\varepsilon,\zeta} > 0$ such that for any pair of $\beta^{(1)}, \beta^{(2)}$ such that $\left\|\beta^{(1)}\right\| \le \zeta$, $\left\|\beta^{(2)}\right\| \le \zeta$ and $\left\|\beta^{(1)} - \beta^{(2)}\right\| < \delta_{\varepsilon,\zeta}$ we have for all $i = 1, 2, \dots$ and for all $n = 1, 2, \dots$

$$\Bigg| \left[\beta^{(1)}\right]^{'} I\!\!E \left[w_{n,\beta^{(1)}}(i, \beta^{(1)}) X_i \left(e_i - X_i^{'}\beta^{(1)} \right) \right] - \left[\beta^{(2)}\right]^{'} I\!\!E \left[w_{n,\beta^{(2)}}(i, \beta^{(2)}) X_i \left(e_i - X_i^{'}\beta^{(2)} \right) \right] \Bigg| \le \varepsilon.$$

Firstly consider

$$\sup_{n\in\mathcal{N}} \sup_{i\in\mathcal{N}} \left| \left[\beta^{(1)}\right]^{'} I\!\!E w_{n,\beta^{(1)}}(i, \beta^{(1)}) X_i \cdot e_i - \left[\beta^{(2)}\right]^{'} I\!\!E w_{n,\beta^{(2)}}(i, \beta^{(2)}) X_i \cdot e_i \right| \quad \text{(A.78)}$$

$$\le \sup_{n\in\mathcal{N}} \sup_{i\in\mathcal{N}} \left\|\beta^{(1)} - \beta^{(2)}\right\| \cdot I\!\!E w_{n,\beta^{(1)}}(i, \beta^{(1)}) \|X_i\| \cdot |e_i| \quad \text{(A.79)}$$

$$+ \sup_{n\in\mathcal{N}} \sup_{i\in\mathcal{N}} \left\|\beta^{(2)}\right\| \cdot I\!\!E \left| w_{n,\beta^{(1)}}(i, \beta^{(1)}) - w_{n,\beta^{(1)}}(i, \beta^{(2)}) \right| \cdot \|X_i\| \cdot |e_i| \quad \text{(A.80)}$$

$$+ \sup_{n\in\mathcal{N}} \sup_{i\in\mathcal{N}} \left\|\beta^{(2)}\right\| \cdot I\!\!E \left| w_{n,\beta^{(1)}}(i, \beta^{(2)}) - w_{n,\beta^{(2)}}(i, \beta^{(2)}) \right| \cdot \|X_i\| \cdot |e_i| . \quad \text{(A.81)}$$

Denoting $\tau_1 = I\!\!E \|X_1\| < \infty$ and finding $A_e = \sup_{i\in\mathcal{N}} I\!\!E |e_i| < \infty$, put $\delta_1 = \frac{1}{6}\varepsilon \cdot \tau_1^{-1} \cdot A_e^{-1}$. Then for any pair $\left\|\beta^{(1)} - \beta^{(2)}\right\| < \delta_1$ (A.79) is less than $\frac{1}{6}\varepsilon$. Putting $\delta_2 = \frac{1}{6}\varepsilon \cdot \zeta^{-2} \cdot \tau_1^{-1} \cdot A_e^{-1} \cdot L^{-1} \cdot f_\sigma^{-1}$ (for f_σ see Remark 1.6), we have also for any pair $\left\|\beta^{(1)} - \beta^{(2)}\right\| < \delta_2$ (A.80) is less than $\frac{1}{6}\varepsilon$. Finally, utilizing Lemma A.8 find δ_3 so that for any pair $\left\|\beta^{(1)} - \beta^{(2)}\right\| < \delta_3$ we have

$$\sup_{i\in\mathcal{N}} \sup_{r\in R} \left| F_{\beta^{(1)},i}(r) - F_{\beta^{(2)},i}(r) \right| \le \frac{1}{6}\varepsilon \cdot \zeta^{-2} \cdot \tau_1^{-1} \cdot A_e^{-1} \cdot L^{-1}.$$

Then for any pair $\left\|\beta^{(1)} - \beta^{(2)}\right\| < \delta_3$ (A.81) is also less than $\frac{1}{6}\varepsilon$. Finally, (A.79), (A.80) and (A.81) imply that for any pair $\left\|\beta^{(1)} - \beta^{(2)}\right\| < \min\{\delta_1, \delta_2, \delta_3\}$ (A.78) is less that $\frac{1}{2}\varepsilon$. The rest of proof employs the same ideas. □

Lemma A.10. Let Conditions $\mathcal{C}1$ and $\mathcal{C}2$ hold. Let e be a r.v. distributed according to $F_e(v)$ and denote for any $\beta \in R^p$ $F_\beta(v) = P(|e - X_1'\beta| < v)$ and $r(\beta) = e - X_1'\beta$.Then for any positive ζ

$$\beta^{'} I\!\!E \left[w\left(F_\beta(|r(\beta)|)\right) X_1 \left(e - X_1^{'}\beta \right) \right]$$

is uniformly continuous in β on $\mathcal{B} = \{\beta \in R^p : \|\beta\| \le \zeta\}$.

P r o o f runs along similar lines as the proof of the previous lemma. $\square$

Lemma A.11. Let Conditions $\mathcal{C}1$ hold. Then for any $\varepsilon > 0$ and $\delta \in (0,1)$ there is $\zeta > 0$ and $n_{\varepsilon,\delta} \in \mathcal{N}$ so that for all $n > n_{\varepsilon,\delta}$

$$P\left(\left\{\omega \in \Omega : \sup_{r \in R} \sup_{\left\|\beta^{(1)}-\beta^{(2)}\right\|<\zeta} \left|F^{(n)}_{\beta^{(1)}}(r) - F^{(n)}_{\beta^{(2)}}(r)\right| < \delta\right\}\right) > 1 - \varepsilon. \tag{A.82}$$

P r o o f. Fix $\varepsilon > 0$ and $\delta \in (0,1)$ and according to Lemma A.8 find $\zeta > 0$ so that for any pair $\left\|\beta^{(1)} - \beta^{(2)}\right\| < \zeta$ we have

$$\sup_{i \in \mathcal{N}} \sup_{r \in R} \left|F_{\beta^{(1)},i}(r) - F_{\beta^{(2)},i}(r)\right| \leq \frac{\delta}{3}.$$

Then also

$$\sup_{r \in R} \left|\overline{F}_{\beta^{(1)}}(r) - \overline{F}_{\beta^{(2)}}(r)\right| \leq \frac{1}{n} \sum_{i=1}^{n} \sup_{i \in \mathcal{N}} \sup_{r \in R} \left|F_{\beta^{(1)},i}(r) - F_{\beta^{(2)},i}(r)\right| \leq \frac{\delta}{3}. \tag{A.83}$$

Employing Lemma A.6 find $K < \infty$ and $n_{\varepsilon,K} \in \mathcal{N}$ so that for any $n > n_{\varepsilon,K}$ and

$$B_n = \left\{\omega \in \Omega : \sup_{r \in R^+} \sup_{\beta \in R^p} \sqrt{n} \left|F^{(n)}_{\beta}(r) - \overline{F}_{\beta}(r)\right| < K\right\} \tag{A.84}$$

we have $P(B_n) > 1 - \varepsilon$.

Further select $n_{\varepsilon,K,\delta} \in \mathcal{N}, n_{\varepsilon,K,\delta} > n_{\varepsilon,K}$ so that

$$\frac{K}{\sqrt{n_{\varepsilon,K,\delta}}} < \frac{\delta}{3}. \tag{A.85}$$

Then, due to (A.83), (A.84) and (A.85), for any $n > n_{\varepsilon,K,\delta}$ and $\omega \in B_n$ we have

$$\sup_{r \in R} \sup_{\left\|\beta^{(1)}-\beta^{(2)}\right\|<\zeta} \left|F^{(n)}_{\beta^{(1)}}(r) - F^{(n)}_{\beta^{(2)}}(r)\right| \leq \sup_{r \in R} \sup_{\beta^{(1)} \in R^p} \left|F^{(n)}_{\beta^{(1)}}(r) - \overline{F}_{\beta^{(1)}}(r)\right|$$

$$+ \sup_{r \in R} \sup_{\left\|\beta^{(1)}-\beta^{(2)}\right\|<\zeta} \left|\overline{F}_{\beta^{(1)}}(r) - \overline{F}_{\beta^{(2)}}(r)\right| + \sup_{r \in R} \sup_{\beta^{(2)} \in R^p} \left|F^{(n)}_{\beta^{(2)}}(r) - \overline{F}_{\beta^{(2)}}(r)\right| < \delta.$$

$\square$

ACKNOWLEDGEMENT

Research was supported by grant of GA ČR number 402/09/0557.

(Received October 12, 2009)

REFERENCES

[1] P. J. Bickel: One-step Huber estimates in the linear model. J. Amer. Statist. Assoc. *70* (1975), 428–433.

[2] L. Breiman: Probability. Addison-Wesley Publishing Company, London 1968.

[3] C. Croux and P. J. Rousseeuw: A class of high-breakdown scale estimators based on subranges. Commun. Statist. Theory Methods *21* (1992), 1935–1951.

[4] J. Hájek and Z. Šidák: Theory of Rank Test. Academic Press, New York 1967.

[5] F. R. Hampel, E. M. Ronchetti, P. J. Rousseeuw, and W. A. Stahel: Robust Statistics – The Approach Based on Influence Functions. Wiley, New York 1986.

[6] T. P. Hettmansperger and S. J. Sheather: A cautionary note on the method of Least Median Squares. Amer. Statist. *46* (1992), 79–83.

[7] G. G. Judge, W. E. Griffiths, R. C. Hill, H. Lütkepohl, and T. C. Lee: The Theory and Practice of Econometrics. Second edition. Wiley, New York 1985.

[8] J. Jurečková and P. K. Sen: Regression rank scores scale statistics and studentization in linear models. In: Proc. Fifth Prague Symposium on Asymptotic Statistics, Physica Verlag 1993, 111–121.

[9] H. L. Koul: Weighted Empirical Processes in Dynamic Nonlinear Models. Lecture Notes in Statist. *166*, Springer Verlag, New York 2002.

[10] J. Mašíček: Consistency of the least weighted squares estimator. In: Statistics for Industry and Technology, Birkhäser Verlag Basel 2004 , 183–194.

[11] P. J. Rousseeuw: Least median of square regression. J. Amer. Statist. Assoc. *79* (1984), 871–880.

[12] P. J. Rousseeuw and A. M. Leroy: Robust Regression and Outlier Detection. Wiley, New York 1987.

[13] J. Štepán: Teorie pravděpodobnosti. Academia, Praha 1987.

[14] H. Tucker: A Graduate Course in Probability. Academic Press, London 1967.

[15] A. W. van der Vaart and J. A. Welner: Weak Convergence and Empirical Processes. Second edition. Springer-Verlag, New York 2000.

[16] J. Á. Víšek: A cautionary note on the method of the Least Median of Squares reconsidered. In: Trans. Twelfth Prague Conference on Information Theory, Statistical Decision Functions and Random Processes 1994, 254–259.

[17] J. Á. Víšek: Sensitivity analysis of M-estimates. Ann. Inst. Statist. Math. *48* (1996), 469–495.

[18] J. Á. Víšek: On high breakdown point estimation. Computational Statistics *11* (1996), 137–146.

[19] J. Á. Víšek: Regression with high breakdown point. In: Robust 2000 (J Antoch and G. Dohnal, eds.), Union of Czech Math. and Phys. 2001, 324–356.

[20] J. Á. Víšek: The least weighted squares I. The asymptotic linearity of normal equations. Bull. Czech Econometr. Soc. *9* (2002), 15, 31–58, 2002

[21] J. Á. Víšek: The least weighted squares II. Consistency and asymptotic normality. Bull. Czech Econom. Soc. *9* (2002), 16, 1–28.

[22] J. Á. Víšek: Sensitivity analysis of M-estimates of nonlinear regression model: Influence of data subsets. Ann. Inst. Statist. Math. *54* (2002), 261–290.

[23] J. Á. Víšek: The least trimmed squares. Sensitivity study. In: Proc. Prague Stochastics 2006, (M. Hušková and M. Janžura, eds.) Matfyzpress 2006, 728–738.

[24] J. Á. Víšek: Robust error-term-scale estimate. IMS Collections. Nonparametrics and Robustness in Modern Statistical Inference and Time Series Analysis: Festschrift for Jana Jurečková, *7* (2010), 254–267.

[25] J. Á. Víšek: Empirical distribution function under heteroscedasticity. To appear in Statistics
http://www.informaworld.com/smpp/content∼db=all∼content=a922317072

Jan Ámos Víšek, Department of Macroeconomics and Econometrics, Institute of Economic Studies, Faculty of Social Siences, Charles University, Smetanovo nábřeží 6, 110 01 Praha 1 and Department of Econometrics, Institute of Information Theory and Automation – Academy of Sciences of the Czech Republic, Pod Vodárenskou věží 4, 182 08 Praha 8. Czech Republic.

e-mail: visek@mbox.fsv.cuni.cz