arXiv:1411.3609v2 [math.ST] 5 Feb 2016

Bernoulli **22**(2), 2016, 1093–1112
DOI: 10.3150/14-BEJ687

Behavior of R-estimators under measurement errors

JANA JUREČKOVÁ[1,*,**], HIRA L. KOUL[2],
RADIM NAVRÁTIL[1,3,†] and JAN PICEK[4]

[1] *Faculty of Mathematics and Physics, Charles University, Sokolovská 83, CZ-186 75 Prague 8, Czech Republic. E-mail:* [*] *jurecko@karlin.mff.cuni.cz;* [†] *navratil@karlin.mff.cuni.cz; url:* [**] *http://www.karlin.mff.cuni.cz/~jurecko*
[2] *Statistics and Probability, Michigan State University, A435 Wells Hall, East Lansing, Michigan, USA.*
E-mail: koul@stt.msu.edu
[3] *Department of Mathematics and Statistics, Masaryk University Brno, Kotlářská 2, 611 37 Brno,*
Czech Republic
[4] *Applied Mathematics, Technical University, Voroněžská 13, Liberec, Czech Republic.*
E-mail: jan.picek@tul.cz

As was shown recently, the measurement errors in regressors affect only the power of the rank test, but not its critical region. Noting that, we study the effect of measurement errors on R-estimators in linear model. It is demonstrated that while an R-estimator admits a local asymptotic bias, its bias surprisingly depends only on the precision of measurements and does neither depend on the chosen rank test score-generating function nor on the regression model error distribution. The R-estimators are numerically illustrated and compared with the LSE and L_1 estimators in this situation.

Keywords: contiguity; linear rank statistic; linear regression model; local asymptotic bias; measurement error; R-estimate

1. Introduction

Measurement technologies are often affected by random errors; if the goal of the experiment is to estimate a parameter, then the estimate is biased, and thus inconsistent. This problem appears in the analytic chemistry, in environmental monitoring, in modeling astronomical data, in biometrics, and practically in all parts of the reality. Moreover, some observations can be undetected, for example, when the measured flux (light, magnetic) in the experiment falls below some flux limit. In econometrics, the errors can be a result of misreporting by subjects, miscoding by the collectors of the data, or by incorrect

This is an electronic reprint of the original article published by the ISI/BS in *Bernoulli*, 2016, Vol. 22, No. 2, 1093–1112. This reprint differs from the original in pagination and typographic detail.

1350-7265 © 2016 ISI/BS

transformation from initial reports. An essential part of measuring techniques, used, for example, in the analytic chemistry, is the construction of a calibration curve – the result for an unknown sample is then determined by interpolation. Robust calibration methods were developed in [24]. However, even the calibration can be affected by measurement errors. The mismeasurements make the statistical inference biased, and they distort the trends in the data.

A variety of functional models have been proposed for handling measurement errors in regression models. Either the regressor or the response or both can be affected by random errors. Technicians, geologists and other specialists are aware of this problem, and try to reduce the bias with various ad hoc procedures. The bias cannot be completely eliminated or substantially reduced unless we have some additional knowledge on the behavior of measurement errors. The papers dealing with practical aspects of measurement error models include [2, 15, 21, 23, 30], among others.

Adcock [1] was probably the first to realize the importance of the situation. There exists a rich literature on the statistical inference in the error-in-variables (EV) models as is evidenced by the monographs of Fuller [9], Carroll et al. [6], and Cheng and Van Ness [7], and the references therein. The monographs [9] and [7] deal mostly with classical Gaussian set up while [6] discusses numerous inference procedure under semi-parametric set up. Nonparametric methods in EV models are considered in [4, 5] and in references therein, and in [8], among others. The regression quantile theory in the area of EV models was started by He and Liang [12]. Arias, Hallock and Sosa-Escudero [3] used an instrumental variable estimator for quantile regression, considering biases arising from unmeasured ability and measurement errors. The problem of mismeasurement is also of interest in the econometric literature: [11] and [16] described the recent developments in treating the effect of mismeasurement on econometric models.

The advantage of *rank and signed rank procedures* in the measurement errors models was discovered recently in [20, 25, 26, 31] and in [32]; the latter made a detailed analysis of rank procedures in the linear model with a nonlinear nuisance regressor and under various kinds of measurement errors. Namely the rank tests can be recommended in this situation: it is shown in [20] that the critical region of the rank test for regression is insensitive to measurement errors in regressors under very general conditions; the errors affect only the power of the test. However, against expectations following from the invariance of the ranks, due to which an estimate of a nuisance parameter in [20] was consistent for every fixed value of the same, we show that the R-estimator of slope parameter $\boldsymbol{\beta}$ in linear model is biased. More precisely, we show that, unless $\boldsymbol{\beta} = \mathbf{0}$, the R-estimator is biased even in a local neighborhood of $\mathbf{0}$. Hence, we cannot have an unbiased estimator of any kind in this situation, unless we have some additional information on the measurement errors.

As we further show in the present paper, surprisingly the local asymptotic bias of R-estimators neither depends on the chosen rank test score-generating functions nor on the unknown distribution of the model errors. It depends only on value of slope parameter vector and on the covariance matrix of the measurement error distribution of regressors.

2. Model and preliminary considerations

Consider the linear regression model

$$Y_{ni} = \beta_0 + \mathbf{x}_{ni}^\top \boldsymbol{\beta} + e_{ni}, \qquad i = 1, \dots, n \tag{2.1}$$

with unknown parameters $\beta_0 \in \mathbb{R}^1$, $\boldsymbol{\beta} \in \mathbb{R}^p$. The regressors $\mathbf{x}_{ni}$ are either deterministic or random and affected by additive random measurement errors, so that instead of $\mathbf{x}_{ni}$ we observe $\mathbf{w}_{ni} = \mathbf{x}_{ni} + \mathbf{v}_{ni}, i = 1, \dots, n,$ where $\mathbf{v}_{n1}, \dots, \mathbf{v}_{nn}$ are p-dimensional random errors, identically distributed with an unknown distribution, and independent of the errors $e_{ni}, 1 \le i \le n$. Moreover, there are additive measurement errors in the responses, thus instead of Y_{ni} we observe $Y_{ni}^* = Y_{ni} + u_{ni}$, where $u_{n1}, \dots, u_{nn}$ are i.i.d. random variables. Thus in terms of the observable responses and predicting variables, our regression model becomes

$$Y_{ni}^* = \beta_0 + \mathbf{w}_{ni}^\top \boldsymbol{\beta} + e_{ni}^*, \qquad i = 1, \dots, n, \tag{2.2}$$

where $e_{ni}^* = e_{ni}^*(\boldsymbol{\beta}) = e_{ni} + u_{ni} - \mathbf{v}_{ni}^\top \boldsymbol{\beta}, i = 1, \dots, n$ are i.i.d random variables.

We are interested in R-estimator of the slope vector $\boldsymbol{\beta}$, considering β_0 as nuisance parameter. To define these estimators, let $R_{ni}(\mathbf{b})$ be the rank of the residual

$$\begin{aligned} Y_{ni}^* - \mathbf{w}_{ni}^\top \mathbf{b} &= e_{ni} + u_{ni} + \mathbf{x}_{ni}^\top \boldsymbol{\beta} - \mathbf{w}_{ni}^\top \mathbf{b} \\ &= e_{ni} + u_{ni} - \mathbf{w}_{ni}^\top \mathbf{b}^* - \mathbf{v}_{ni}^\top \boldsymbol{\beta}, \qquad i = 1, \dots, n, \end{aligned}$$

where $\mathbf{b}^* = \mathbf{b} - \boldsymbol{\beta}$. We shall work with the vector of linear rank statistics

$$\mathbf{S}_n(\mathbf{b}) = (S_{nj}(\mathbf{b}); j = 1, \dots, p)^\top = n^{-1/2} \sum_{i=1}^{n} (\mathbf{w}_{ni} - \bar{\mathbf{w}}_n) a_n(R_{ni}(\mathbf{b})), \tag{2.3}$$

where the scores $a_n(i), 1 \le i \le n$ are nondecreasing in i and $\sum_{i=1}^n a_n(i) = 0$.

Hodges and Lehmann [14] introduced a class of estimators of the location parameter θ in one- and two-sample location models, by inverting a class of rank tests for θ. This methodology was extended to linear regression models without measurement error by Jurečková [19], where an estimator of $\boldsymbol{\beta}$ is defined as

$$\widehat{\boldsymbol{\beta}}_n = \underset{\mathbf{b} \in \mathbb{R}_p}{\arg\min} \sum_{j=1}^{p} |S_{nj}(\mathbf{b})|.$$

This estimator can be seen to be asymptotically equivalent to an estimator obtained by inverting the equations $S_{nj}(\mathbf{b}) = 0, j = 1, \dots, p$. Note that this latter estimator is precisely an extension of the Hodges–Lehmann estimator from one- and two-sample location models to linear regression models without measurement error. Under more general conditions, the R-estimators are studied by Koul [22].

On the other hand, Jaeckel [17] called an analog of the function

$$\mathcal{D}_n(\mathbf{b}) = \sum_{i=1}^{n} [Y_{ni}^* - \mathbf{w}_{ni}^\top \mathbf{b}](a_n(R_{ni}(\mathbf{b})) - \bar{a}_n), \tag{2.4}$$

as a measure of rank dispersion of residuals, in the case of no measurement error where $\mathbf{w}_{ni}$'s are replaced by $\mathbf{x}_{ni}$'s. He showed that $\mathcal{D}_n(\mathbf{b})$ is convex and piecewise linear in $\mathbf{b} \in \mathbb{R}^p$. He also showed that $-n^{1/2}\mathbf{S}_n(\mathbf{b})$ is the subgradient of $\mathcal{D}_n(\mathbf{b})$; hence the estimator defined as a minimizer of $\mathcal{D}_n$ exists and is equivalent to the above estimators based on $\mathbf{S}_n$. Both of these estimators are asymptotically equivalent, and Jaeckel's definition of R-estimator is now generally used in the literature. We are using this definition of R-estimator throughout this paper.

In the absence of measurement errors, that is, if $\mathbf{w}_{ni} = \mathbf{x}_{ni}, u_{ni} = 0, i = 1, \ldots, n$, the estimator $\widehat{\boldsymbol{\beta}}_n$ is consistent and asymptotically normal. However, $\widehat{\boldsymbol{\beta}}_n$ is biased in the presence of measurement errors, even asymptotically, unless the true $\boldsymbol{\beta} = \mathbf{0}$. Furthermore, we show that it is even asymptotically locally biased in the sense that the asymptotic distribution of $n^{1/2}(\widehat{\boldsymbol{\beta}}_n - n^{-1/2}\boldsymbol{\beta}^0)$, with a fixed $\boldsymbol{\beta}^0 \in \mathbb{R}^p$, converges to a normal distribution with nonzero mean vector and some positive definite covariance matrix.

In the sequel, all limits are taken as $n \to \infty$, unless mentioned otherwise, $\xrightarrow{p}$ denotes the convergence in probability. We shall now describe the needed assumptions on the underlying entities.

(A.1) The score generating function $\varphi : (0,1) \mapsto \mathbb{R}$ is nondecreasing, square-integrable and skew-symmetric on $(0,1)$, that is, satisfies $\varphi(1-t) = -\varphi(t), 0 < t < 1$. The scores $a_n(i), i = 1, \ldots, n$ are generated by φ in either of the following two ways:

$$a_n(i) = \varphi\left(\frac{i}{n+1}\right) \quad \text{or} \quad a_n(i) = \mathbb{E}\varphi(U_{n:i}), \qquad i = 1, \ldots, n,$$

where $U_{n:1} \leq \cdots \leq U_{n;n}$ are order statistics pertaining to the sample of size n from the uniform $(0,1)$ distribution.

(F.1) Distribution function F of the model errors e_{ni} has an absolutely continuous density f with a.e. derivative f'.

(F.2) For every $u \in \mathbb{R}$, $\int (|f'(x-tu)|^j / f^{j-1}(x))\,\mathrm{d}x \to \int (|f'(x)|^j / f^{j-1}(x))\,\mathrm{d}x < \infty$, as $t \to 0$, $j = 2, 3$.

(V.1) The measurement errors $\{u_{ni}, 1 \leq i \leq n\}$ are independent of $\{e_{ni}, \mathbf{v}_{ni}, 1 \leq i \leq n\}$ and i.i.d. with generally an unknown absolutely continuous density h, having finite Fisher's information for location.

(V.2) The measurement error $\mathbf{v}_{ni}$ is independent of e_{ni} and its p-dimensional distribution function G has a continuous density g, generally unknown, $i = 1, \ldots, n$.

(V.3) $\mathbb{E}\mathbf{V}_n \to \mathbf{V}$ where $\mathbf{V}_n = n^{-1} \sum_{i=1}^n (\mathbf{v}_{ni} - \bar{\mathbf{v}}_n)(\mathbf{v}_{ni} - \bar{\mathbf{v}}_n)^\top$ and $\mathbf{V}$ is a positive definite $p \times p$ matrix. Moreover, $\sup_{n \geq 1} \mathbb{E}(\|\mathbf{v}_{n1}\|^3 + \|\mathbf{x}_{n1}\|^3) < \infty$.

(V.4) $\mathbb{E}[n^{-1} \sum_{i=1}^n (\mathbf{v}_{ni} - \bar{\mathbf{v}}_n)(\mathbf{x}_{ni} - \bar{\mathbf{x}}_n)^\top] \to \mathbf{0}$.

(X.1) If the regressors $\mathbf{x}_{ni}$ are nonrandom, then assume that $\mathbf{Q}_n \to \mathbf{Q}$, where

$$\mathbf{Q}_n = n^{-1} \sum_{i=1}^{n} (\mathbf{x}_{ni} - \bar{\mathbf{x}}_n)(\mathbf{x}_{ni} - \bar{\mathbf{x}}_n)^\top,$$

and $\mathbf{Q}$ is positive definite $p \times p$ matrix. Moreover,

$$\frac{1}{n} \max_{1 \le i \le n} (\mathbf{x}_{ni} - \bar{\mathbf{x}}_n)^\top (\mathbf{Q}_n)^{-1} (\mathbf{x}_{ni} - \bar{\mathbf{x}}_n) \to 0.$$

(X.2) If the regressors $\mathbf{x}_{ni}$ are random, then assume that they are independent of $e_{ni}, u_{ni}, \mathbf{v}_{ni}$, $i = 1, \dots, n$, and

$$\mathbb{E}\left[n^{-1} \sum_{i=1}^{n} (\mathbf{x}_{ni} - \bar{\mathbf{x}}_n)(\mathbf{x}_{ni} - \bar{\mathbf{x}}_n)^\top \right] \to \mathbf{Q},$$

where $\mathbf{Q}$ is positive definite $p \times p$ matrix.

Let $m(\cdot), M(\cdot)$ denote the density and distribution function of $e_{ni} + u_{ni}, i = 1, \dots, n$, that is, $m(z) = \int f(z-t)h(t)\,\mathrm{d}t$. The density is absolutely continuous and has finite Fisher's information $\mathcal{I}(m)$. We need to define

$$\begin{aligned} \gamma_m &= -\int_{\mathbb{R}^1} \varphi(M(z))\,\mathrm{d}m(z), \qquad A_m^2(\varphi) = \gamma_m^{-2} \int_0^1 \varphi^2(u)\,\mathrm{d}u, \\ \mathbf{B} &= -(\mathbf{Q} + \mathbf{V})^{-1} \mathbf{V} \boldsymbol{\beta}^0. \end{aligned} \tag{2.5}$$

The following theorem gives the asymptotic distribution of the estimator $\widehat{\boldsymbol{\beta}}_n$ when the true parameter value is

$$\boldsymbol{\beta}_n = n^{-1/2} \boldsymbol{\beta}^0, \qquad \boldsymbol{\beta}^0 \in \mathbb{R}^p \text{ fixed.} \tag{2.6}$$

Theorem 2.1. *Assume the conditions* (A.1), (F.1)–(F.2), (V.1)–(V.4), (X.1)–(X.2) *hold. When the true parameter value is $\boldsymbol{\beta}_n$, the R-estimator $\widehat{\boldsymbol{\beta}}_n$ is asymptotically normally distributed with the bias $\mathbf{B} = -(\mathbf{Q} + \mathbf{V})^{-1} \mathbf{V} \boldsymbol{\beta}^0$, that is,*

$$n^{1/2} (\widehat{\boldsymbol{\beta}}_n - \boldsymbol{\beta}_n) \xrightarrow{\mathcal{D}} \mathcal{N}_p(\mathbf{B}, (\mathbf{Q} + \mathbf{V})^{-1} A_m^2(\varphi)). \tag{2.7}$$

Theorem 2.1 will be proved in several steps; the proof is given in Section 3. The numerical illustrations of the results are given in subsequent Section 4.

Corollary 2.1. *Under the conditions of Theorem 2.1 and under $\boldsymbol{\beta} = \boldsymbol{\beta}_n = n^{-1/2} \boldsymbol{\beta}^0$, the R-estimator $\widehat{\boldsymbol{\beta}}_n$ has asymptotic normal distribution*

$$n^{1/2} (\widehat{\boldsymbol{\beta}}_n - (\mathbf{Q} + \mathbf{V})^{-1} \mathbf{Q} \boldsymbol{\beta}_n) \xrightarrow{\mathcal{D}} \mathcal{N}_p(\mathbf{0}, (\mathbf{Q} + \mathbf{V})^{-1} A_m^2(\varphi)). \tag{2.8}$$

Notice that the local asymptotic bias cannot be controlled by the choice of the score-generating function φ; this choice can only influence the asymptotic variance factor of the estimator. The magnitude of the bias fully depends on the precision of the measurements, namely on the matrix $\mathbf{V}$. The measurement errors in the responses Y_{ni} affect only the asymptotic variance, not the bias. The result is entirely nonparametric, valid for classes of distributions of model and measurement errors, demanding only finite first moment and finite (and positive) Fisher's information for location of the model error distributions, and finite third moment for measurement error distributions.

Consider the two measurement methods with the same regressors (random or nonrandom), with the respective limiting covariance matrices $\mathbf{V}_1, \mathbf{V}_2$. Comparing the biases in (2.7) for $\mathbf{V}_1$ and $\mathbf{V}_2$, the first method is considered being more precise than the second one if the matrix $(\mathbf{V}_2 + \mathbf{Q})^{-1} \prec (\mathbf{V}_1 + \mathbf{Q})^{-1}$; otherwise speaking, if $\mathbf{Q}^{-1}\mathbf{V}_1 \prec \mathbf{Q}^{-1}\mathbf{V}_2$, where the ordering $\mathbf{A} \prec \mathbf{B}$ means that $\mathbf{B} - \mathbf{A}$ is a positive definite matrix.

3. Proof of Theorem 2.1

We shall prove Theorem 2.1 in several steps. Notice that if we observe $Y_{ni}^* = Y_{ni} + u_{ni}$ instead of Y_{ni}, then $e_{ni}^* = e_{ni} + u_{ni}, i = 1, \ldots, n$ are still i.i.d. random variables with density $m(z) = \int f(z-t)h(t)\,\mathrm{d}t$. The steps of the proof are parallel for both densities f and m of model errors; measurement errors in the Y_{ni} affect only the asymptotic variance of the estimate, not the bias. Noting this, we shall prove the theorem assuming $u_{ni} \equiv 0 . i = 1, \ldots, n$. In the sequel, we shall suppress the subscript n whenever it does not cause a confusion.

The steps of the proof are as follows:

(1) Asymptotic representation of the linear rank statistic

$$\mathbf{S}_n(\mathbf{0}, \mathbf{0}) = n^{-1/2} \sum_{i=1}^{n} (\mathbf{w}_{ni} - \bar{\mathbf{w}}_n) a_n(R_{ni}(\mathbf{0})) \tag{3.1}$$

with the sum of independent summands. Here $\mathbf{w}_{ni} = \mathbf{x}_{ni} + \mathbf{v}_{ni}, i = 1, \ldots, n,$ while $\mathbf{x}_{n1}, \ldots, \mathbf{x}_{nn}$ are either i.i.d. random vectors or nonrandom vectors, and $\mathbf{v}_{n1}, \ldots, \mathbf{v}_{nn}$ are i.i.d. random vectors.

(2) Contiguity of the sequence $\{Q_n\}$ of distributions of $(e_{ni} - (\mathbf{w}_{ni} - \bar{\mathbf{w}}_n)^\top \mathbf{b}_n^* - (\mathbf{v}_{ni} - \bar{\mathbf{v}}_n)^\top \boldsymbol{\beta}_n)$, with $\mathbf{b}_n^* = n^{-1/2}\mathbf{b}^0, \boldsymbol{\beta}_n = n^{-1/2}\boldsymbol{\beta}^0$ for $\mathbf{b}^0, \boldsymbol{\beta}^0 \in \mathbb{R}^p$ fixed, with respect to the sequence $\{P_n\}$ of distributions of $e_{ni},\ i = 1, \ldots, n$.

(3) Asymptotic representation of the linear rank statistic (2.3) under contiguous sequence of distribution $\{Q_n\}$, and the resulting asymptotic linearity of (2.3) in parameters $\mathbf{b}^0, \boldsymbol{\beta}^0$.

(4) Uniform asymptotic quadraticity of $\mathcal{D}_n$ in parameters $\mathbf{b}^0, \boldsymbol{\beta}^0$ under $\{Q_n\}$, as a result of (3) and of the convexity of $\mathcal{D}_n$.

(5) Resulting asymptotic distribution and bias of $\widehat{\boldsymbol{\beta}}_n$ in the case $u_{ni} \equiv 0, i = 1, \ldots, n$.

(6) Asymptotic distribution and bias of $\widehat{\boldsymbol{\beta}}_n$ in the case of nonzero $u_{ni}, i = 1, \ldots, n$.

3.1. Asymptotic representation of $\mathrm{S}_n(0,0)$

Assume that $u_{ni} = 0, i = 1, \ldots, n$. That is, for now we assume that there is no measurement error in the response variables Y_{ni}. Let

$$\mathbf{Z}_n = n^{-1/2} \sum_{i=1}^{n} (\mathbf{w}_{ni} - \bar{\mathbf{w}}_n) \varphi(F(e_{ni})).$$

We are ready to state and prove the following lemma.

Lemma 3.1. *Under the conditions of Theorem 2.1, the statistic* $\mathbf{S}_n(\mathbf{0},\mathbf{0})$ *admits the asymptotic representation*

$$\mathbf{S}_n(\mathbf{0},\mathbf{0}) = \mathbf{Z}_n + \mathrm{o}_p(1). \tag{3.2}$$

Proof. The proof is adapted from [28]. If $\mathbf{b} = \boldsymbol{\beta} = \mathbf{0}$, then $(Y_{n1}, \ldots, Y_{nn}) = (e_{n1}, \ldots, e_{nn})$. Let $R_{n1}, \ldots, R_{nn}$ denote their ranks. Further denote $r_{ni} = a_n(R_{ni}) - \varphi(F(e_{ni})), i = 1, \ldots, n$.

Let σ_j^2 be the variance of $w_{ij}, i = 1, \ldots, n$, for $j = 1, \ldots, p$, and let $s^2 = \sum_{j=1}^{p} \sigma_j^2$.

Notice that $(r_{n1}, \ldots, r_{nn})$ and $(\mathbf{w}_1, \ldots, \mathbf{w}_n)$ are independent. Consider the conditional squared distance

$$\begin{aligned}
&\mathbb{E}_G\{(\mathbf{S}_n - \mathbf{Z}_n)^\top (\mathbf{S}_n - \mathbf{Z}_n) | e_1, \ldots, e_n\} \\
&= n^{-1} \mathbb{E}_G \left\{ \sum_{i=1}^{n} \sum_{k=1}^{n} (\mathbf{w}_i - \bar{\mathbf{w}}_n)^\top (\mathbf{w}_k - \bar{\mathbf{w}}_n) r_i r_k \middle| e_1, \ldots, e_n \right\} \\
&= n^{-1} \sum_{i=1}^{n} \sum_{k=1}^{n} r_i r_k \mathbb{E}_G \left\{ \sum_{j=1}^{p} (w_{ij} - \bar{w}_j)(w_{kj} - \bar{w}_j) \middle| e_1, \ldots, e_n \right\} \\
&= n^{-1} \left\{ \sum_{i=1}^{n} \sum_{k=1}^{n} r_i r_k \sum_{j=1}^{p} (x_{ij} - \bar{x}_j)(x_{kj} - \bar{x}_j) + s^2 \sum_{i=1}^{n} (r_i - \bar{r})^2 \right\} \\
&= \sum_{j=1}^{p} \left[n^{-1/2} \sum_{i=1}^{n} (x_{ij} - \bar{x}_j) r_i \right]^2 + s^2 \sum_{i=1}^{n} (r_i - \bar{r})^2.
\end{aligned}$$

Then (3.2) follows from [10] [Theorems V.1.4.a,b, V.1.6.a]. □

3.2. Contiguity

For any two probability measures P and Q, absolutely continuous with respect to a σ-finite measure ν with $p = \mathrm{d}P/\mathrm{d}\nu$, $q = \mathrm{d}Q/\mathrm{d}\nu$, let

$$H(P,Q) = \left[\int (\sqrt{p} - \sqrt{q})^2 \, \mathrm{d}\mu \right]^{1/2} = \left[2 \int (1 - \sqrt{pq}) \, \mathrm{d}\mu \right]^{1/2}$$

denote the Hellinger distance between P and Q.

Let $\{P_{n1},\ldots,P_{nn}\}$ and $\{Q_{n1},\ldots,Q_{nn}\}$ be two triangular arrays of probability measures defined on measurable space $(\mathcal{X},\mathcal{A})$ with densities p_{ni},q_{ni} with respect to σ-finite measures μ_i [which can be also $\mu_i=P_{ni}+Q_{ni}, i=1,\ldots,n$]. Denote $P_n^{(n)}=\prod_{i=1}^n P_{ni}$ and $Q_n^{(n)}=\prod_{i=1}^n Q_{ni}$ the product measures, $n=1,2,\ldots$.

Oosterhoff and van Zwet [27] proved that $\{Q_n^{(n)}\}$ is contiguous with respect to $\{P_n^{(n)}\}$ if and only if

$$\limsup_{n\to\infty}\sum_{i=1}^n H^2(P_{ni},Q_{ni})<\infty, \tag{3.3}$$

$$\lim_{n\to\infty}\sum_{i=1}^n Q_{ni}\Big\{\frac{q_{ni}(X_{ni})}{p_{ni}(X_{ni})}\geq c_n\Big\}=0 \qquad \forall c_n\to\infty. \tag{3.4}$$

Note that in the case $P_{ni}\equiv P_n, p_{ni}\equiv p_n$, and $Q_{ni}\equiv Q_n, q_{ni}\equiv q_n$, not depending on i,

$$\begin{aligned}\sum_{i=1}^n H^2(P_{ni},Q_{ni}) &= n\int[\sqrt{q_n(z)}-\sqrt{p_n(z)}]^2\,\mathrm{d}z\\ &= n\int\frac{(q_n(z)-p_n(z))^2}{[\sqrt{q_n(z)}+\sqrt{p_n(z)}]^2}\,\mathrm{d}z\\ &\leq n\int\frac{(q_n(z)-p_n(z))^2}{p_n(z)}\,\mathrm{d}z.\end{aligned} \tag{3.5}$$

Moreover, for $c_n>1$ and with $d_n=c_n-1$,

$$\begin{aligned}\sum_{i=1}^n Q_{ni}\Big\{\frac{q_{ni}(X_{ni})}{p_{ni}(X_{ni})}\geq c_n\Big\} &= nQ_n\Big\{\frac{q_n(X_{n1})-p_n(X_{n1})}{p_n(X_{n1})}\geq d_n\Big\}\\ &\leq d_n^{-2}n\int\frac{|q_n(x)-p_n(x)|^2}{p_n^2(x)}q_n(x)\,\mathrm{d}x\\ &\leq d_n^{-2}n\int\frac{|q_n(x)-p_n(x)|^3}{p_n^2(x)}\,\mathrm{d}x\\ &\quad+d_n^{-2}n\int\frac{|q_n(x)-p_n(x)|^2}{p_n(x)}\,\mathrm{d}x.\end{aligned} \tag{3.6}$$

Now, let $Y_{ni}=\mathbf{x}_{ni}^\top\boldsymbol{\beta}+e_{ni}, i=1,\ldots,n.$ where the e_{ni} are i.i.d. with distribution function F and density f, satisfying (F.1) and (F.2). Consider the residuals

$$\begin{aligned}Y_{ni}-(\mathbf{w}_{ni}-\bar{\mathbf{w}}_n)^\top\mathbf{b}_n &= e_{ni}+(\mathbf{x}_{ni}-\bar{x}_n)^\top\boldsymbol{\beta}_n-(\mathbf{w}_{ni}-\bar{\mathbf{w}}_n)^\top\mathbf{b}_n\\ &= e_{ni}-(\mathbf{w}_{ni}-\bar{\mathbf{w}}_n)^\top\mathbf{b}_n^*-(\mathbf{v}_{ni}-\bar{\mathbf{v}}_n)^\top\boldsymbol{\beta}_n,\end{aligned}$$

$i=1,\dots,n$, where $\mathbf{b}_n = n^{-1/2}\mathbf{b}^0, \boldsymbol{\beta}_n = n^{-1/2}\boldsymbol{\beta}^0, \mathbf{b}_n^* = n^{-1/2}\mathbf{b}^{0*}, \mathbf{b}_0^* = \mathbf{b}^0 - \boldsymbol{\beta}^0$, with fixed $\mathbf{b}^0, \boldsymbol{\beta}^0 \in \mathbb{R}^p$. Using (3.3) and (3.4), we shall prove the following lemma.

Lemma 3.2. *Under the conditions of Theorem 2.1, the sequence* $\{Q_n^{(n)}\}$ *is contiguous with respect to* $\{P_n^{(n)}\}$, *where* $Q_n^{(n)} = \prod_{i=1}^n Q_{ni}$, $P_n^{(n)} = \prod_{i=1}^n P_{ni}$, *where* P_{ni} *is the distribution of* e_{ni} *and* Q_{ni} *is the distribution of* $(e_{ni} - (\mathbf{w}_{ni} - \bar{\mathbf{w}}_n)^\top \mathbf{b}_n^* - (\mathbf{v}_{ni} - \bar{\mathbf{v}}_n)^\top \boldsymbol{\beta}_n), i = 1,\dots,n$.

Proof. We shall distinguish the two cases: the $\mathbf{x}_{ni}$ are either i.i.d. random vectors or nonrandom vector components.

We start with the first case, where $\mathbf{w}_{n1},\dots,\mathbf{w}_{nn}$ are i.i.d. random vectors. Note that $U_i := (\mathbf{w}_{ni} - \bar{\mathbf{w}}_n)^\top \mathbf{b}^{0*} + (\mathbf{v}_{ni} - \bar{\mathbf{v}}_n)^\top \boldsymbol{\beta}^0, i=1,\dots,n$, are i.i.d. r.v.'s. Let k_1 denote the common density function of U_i. Then, Q_{ni}, P_{ni} do not depend on i and $q_n(x) \equiv \int f(x - n^{-1/2}u) k_1(u)\,\mathrm{d}u, p_n(x) \equiv f(x)$. Hence, by the Cauchy–Schwarz inequality, and the Fubini theorem,

$$\begin{aligned}
n\int \frac{(q_n(x)-p_n(x))^2}{p_n(x)}\,\mathrm{d}x &= n\int\left\{\int [f(x-n^{-1/2}u)-f(x)]k_1(u)\,\mathrm{d}u\right\}^2 \frac{\mathrm{d}x}{f(x)} \\
&\le n\int\!\!\int [f(x-n^{-1/2}u)-f(x)]^2 \frac{k_1(u)}{f(x)}\,\mathrm{d}u\,\mathrm{d}x \\
&\le n\int\!\!\int \left[\int_{-n^{-1/2}}^{n^{-1/2}} |uf'(x-tu)|\,\mathrm{d}t\right]^2 \frac{k_1(u)}{f(x)}\,\mathrm{d}u\,\mathrm{d}x \\
&\le 2n^{1/2}\int\!\!\int\!\!\int_{-n^{-1/2}}^{n^{-1/2}} |f'(x-tu)|^2\,\mathrm{d}t\,u^2 \frac{k_1(u)}{f(x)}\,\mathrm{d}u\,\mathrm{d}x \\
&\le 2n^{1/2}\int\!\!\int_{-n^{-1/2}}^{n^{-1/2}} \int \frac{|f'(x-tu)|^2}{f(x)}\,\mathrm{d}x\,u^2 k_1(u)\,\mathrm{d}u\,\mathrm{d}t \qquad \forall n\ge 1.
\end{aligned}$$

Hence, by (3.5), (F.2) applied with $j=2$, and by (V.3), which guaranteed $\int u^2 k_1(u)\,\mathrm{d}u < \infty$,

$$\limsup_n \sum_{i=1}^n H^2(P_{ni}, Q_{ni}) \le 2I(f)\int u^2 k_1(u)\,\mathrm{d}u < \infty. \tag{3.7}$$

Similarly, the bound

$$n\int \frac{(q_n(x)-p_n(x))^3}{p_n^2(x)}\,\mathrm{d}x \le 2n^{1/2}\int\!\!\int_{-n^{-1/2}}^{n^{-1/2}} \int \frac{|f'(x-tu)|^3}{f^2(x)}\,\mathrm{d}x |u|^3 k_1(u)\,\mathrm{d}u\,\mathrm{d}t \qquad \forall n\ge 1$$

together with (3.6), (F.2) applied with $j = 3$, and (V.3), which guaranteed $\int |u|^3 k_1(u)\,\mathrm{d}u < \infty$, yield

$$\begin{aligned}
&\lim_n \sum_{i=1}^n Q_{ni}\left\{\frac{q_{ni}(Y_{ni})}{p_{ni}(Y_{ni})} \ge c_n\right\} \\
&\quad \le 2\lim_n d_n^{-2}\left\{\int\left(\frac{|f'(x)|}{f(x)}\right)^3 f(x)\,\mathrm{d}x \int |u|^3 k_1(u)\,\mathrm{d}u + I(f)\int u^2 k_1(u)\,\mathrm{d}u\right\} = 0.
\end{aligned}$$

This ensures the validity of (3.4), and completes the proof of the contiguity in present case.

Next, consider the case where $\mathbf{x}_{n1}, \ldots, \mathbf{x}_{nn}$ are nonrandom, and we observe $\mathbf{w}_{ni} = \mathbf{x}_{ni} + \mathbf{v}_{ni}, i = 1, \ldots, n$. Let k_2 denote the density of $(\mathbf{v}_{ni} - \bar{\mathbf{v}}_n)^\top \mathbf{b}^0, i = 1, \ldots, n$. Again, by (3.5),

$$\begin{aligned}
&\sum_{i=1}^n H^2(P_{ni}, Q_{ni}) \\
&\quad \le \sum_{i=1}^n \int\left\{\int [f(e - n^{-1/2}u) - f(e)] k_2(u + (\mathbf{x}_{ni} - \bar{\mathbf{x}}_n)^\top \mathbf{b}^{0*})\,\mathrm{d}u\right\}^2 \frac{\mathrm{d}e}{f(e)} \\
&\quad \le \sum_{i=1}^n \int\left\{\int [f(e - n^{-1/2}u) - f(e)]^2 k_2(u - (\mathbf{x}_{ni} - \bar{\mathbf{x}}_n)^\top \mathbf{b}_0^*)\,\mathrm{d}u\right\} \frac{\mathrm{d}e}{f(e)} \\
&\quad \le 2n^{1/2} \int\int_{n^{-1/2}}^{n^{-1/2}} \int \frac{|f'(e - tu)|^2}{f(e)}\,\mathrm{d}e\,\mathrm{d}t\, n^{-1}\sum_{i=1}^n u^2 k_2(u - (\mathbf{x}_{ni} - \bar{\mathbf{x}}_n)^\top \mathbf{b}^{0*})\,\mathrm{d}u.
\end{aligned}$$

Hence, by (F.2) and by the change of variable formula,

$$\limsup_n \sum_{i=1}^n H^2(P_{ni}, Q_{ni}) \le C\left[\int u^2 k_2(u)\,\mathrm{d}u + \mathbf{b}^{0*\top}\mathbf{Q}_n \mathbf{b}^{0*}\right] < \infty. \tag{3.8}$$

Similarly one verifies (3.4) here. □

Lemmas 3.1 and 3.2 enable us to extend the approximation of the rank statistic $\mathbf{S}_n(\mathbf{b}_n^*, \boldsymbol{\beta}_n)$ by a sum of independent r.v.'s under the contiguous sequence of distributions. Let

$$\mathbf{T}_n(\mathbf{b}_n^*, \boldsymbol{\beta}_n) = n^{-1/2}\sum_{i=1}^n (\mathbf{w}_{ni} - \bar{\mathbf{w}}_n)\varphi(F(e_{ni} - (\mathbf{w}_{ni} - \bar{\mathbf{w}}_n)^\top \mathbf{b}_n^* - (\mathbf{v}_{ni} - \bar{\mathbf{v}}_n)^\top \boldsymbol{\beta}_n)).$$

We have the following corollary.

Corollary 3.1. *Under the conditions of Theorem 2.1, and under* $\{Q_n^{(n)}\}$,

$$\begin{aligned}\mathbf{S}_n(\mathbf{b}_n^*,\boldsymbol{\beta}_n) &= n^{-1/2}\sum_{i=1}^{n}(\mathbf{w}_{ni}-\bar{\mathbf{w}}_n)a_n(R(e_{ni}-(\mathbf{w}_{ni}-\bar{\mathbf{w}}_n)^\top\mathbf{b}_n^*-(\mathbf{v}_{ni}-\bar{\mathbf{v}}_n)^\top\boldsymbol{\beta}_n)) \\ &= \mathbf{T}_n(\mathbf{b}_n^*,\boldsymbol{\beta}_n)+\mathrm{o}_p(1).\end{aligned} \tag{3.9}$$

Hence,

$$\mathbf{S}_n(\mathbf{b}_n^*,\boldsymbol{\beta}_n)-\mathbf{S}_n(\mathbf{0},\mathbf{0})=\mathbf{T}_n(\mathbf{b}_n^*,\boldsymbol{\beta}_n)-\mathbf{T}_n(\mathbf{0},\mathbf{0})+\mathrm{o}_p(1).$$

3.3. Asymptotic linearity of $\mathbf{S}_n(\mathbf{b}_n^*,\boldsymbol{\beta}_n)$

Lemma 3.3. *Under the conditions of Theorem 2.1,*

$$\|\mathbf{S}_n(\mathbf{b}_n^*,\boldsymbol{\beta}_n)-\mathbf{S}_n(\mathbf{0},\mathbf{0})+\gamma[(\mathbf{Q}+\mathbf{V})\mathbf{b}^{0*}+\mathbf{V}\boldsymbol{\beta}^0]\|\xrightarrow{p}0, \tag{3.10}$$

where

$$\gamma=\int_0^1-\frac{f'(F^{-1}(u))}{f(F^{-1}(u))}\varphi(u)\,\mathrm{d}u=-\int_{\mathbb{R}^1}\varphi(F(z))\,\mathrm{d}f(z). \tag{3.11}$$

Proof. Consider the sequence of functions $\{\varphi^{(k)}(\cdot)\}_{k=1}^{\infty}$

$$\varphi^{(k)}(u)=\varphi\left(\frac{1}{k+1}\right)\mathbb{I}\left[u<\frac{1}{k}\right]+\varphi(u)\mathbb{I}\left[\frac{i-1}{k+1}<u\le\frac{i}{k+1}\right],\qquad i=2,\dots,k. \tag{3.12}$$

Then, by Lemma V.1.6.a [10], $\varphi^{(k)}$ is nondecreasing and bounded on $(0,1)$ and

$$\lim_{n\to\infty}\int_0^1[\varphi^{(k)}(u)-\varphi(u)]^2\,\mathrm{d}u=0. \tag{3.13}$$

The function $\varphi^{(k)}$ has at most countable set B_k of discontinuity points. Observe that assumption (V.3) implies that $n^{-1/2}\max_{1\le i\le n}\{\|\mathbf{w}_{ni}-\bar{\mathbf{w}}_n\|+\|\mathbf{v}_{ni}-\bar{\mathbf{v}}_n\|\}\xrightarrow{p}0$. This fact together with the uniform continuity of F implies that

$$\sup_{e\in\mathbb{R},1\le i\le n}|F(e-n^{-1/2}(\mathbf{w}_{ni}-\bar{\mathbf{w}}_n)^\top\mathbf{b}^{0*}-n^{-1/2}(\mathbf{v}_{ni}-\bar{\mathbf{v}}_n)^\top\boldsymbol{\beta}^0)-F(e)|\xrightarrow{p}0.$$

Hence,

$$\varphi^{(k)}(F(e-n^{-1/2}(\mathbf{w}_{ni}-\bar{\mathbf{w}}_n)^\top\mathbf{b}^{0*}-n^{-1/2}(\mathbf{v}_{ni}-\bar{\mathbf{v}}_n)^\top\boldsymbol{\beta}^0))$$

converges to $\varphi^{(k)}(F(e))$, in probability, uniformly in $i=1,\dots,n$. It, in turn, implies that the conditional expectation

$$\mathbb{E}[(\varphi^{(k)}(F(e_{ni}-n^{-1/2}(\mathbf{w}_{ni}-\bar{\mathbf{w}}_n)^\top\mathbf{b}^{0*}-n^{-1/2}(\mathbf{v}_{ni}-\bar{\mathbf{v}}_n)^\top\boldsymbol{\beta}^0))-\varphi^{(k)}(F(e_{ni})))^2|\mathbf{v}_{ni},\mathbf{x}_{ni}]$$

converges to 0, in probability, uniformly in $i=1,\dots,n$ and k.

Let $\mathbf{S}_n^{(k)}(\mathbf{b}^*,\boldsymbol{\beta})$ and $\mathbf{T}_n^{(k)}(\mathbf{b}^*,\boldsymbol{\beta})$ be analogous to $\mathbf{S}_n(\mathbf{b}^*,\boldsymbol{\beta}),\mathbf{T}_n(\mathbf{b}^*,\boldsymbol{\beta})$, respectively, with φ replaced with $\varphi^{(k)}$. Then we can bound the norm of the covariance matrix of $\mathbf{T}_n^{(k)}(\mathbf{b}^*,\boldsymbol{\beta})-\mathbf{T}_n^{(k)}(\mathbf{0},\mathbf{0})$ for any fixed k in the following way: denote

$$\mathbf{A}_n^{(k)}=\mathbb{E}\{[\mathbf{T}_n^{(k)}(\mathbf{b}_n^*,\boldsymbol{\beta}_n)-\mathbf{T}_n^{(k)}(\mathbf{0},\mathbf{0})][\mathbf{T}_n^{(k)}(\mathbf{b}_n^*,\boldsymbol{\beta}_n)-\mathbf{T}_n^{(k)}(\mathbf{0},\mathbf{0})]^\top\}.$$

Then

$$\begin{aligned}\mathbf{A}_n^{(k)}&=\mathbb{E}\Bigg\{n^{-1}\sum_{i=1}^n(\mathbf{w}_{ni}-\bar{\mathbf{w}}_n)(\mathbf{w}_{ni}-\bar{\mathbf{w}}_n)^\top\\&\qquad\times[\varphi^{(k)}(F(e_{ni}-(\mathbf{w}_{ni}-\bar{\mathbf{w}}_n)^\top\mathbf{b}_n^*-(\mathbf{v}_{ni}-\bar{\mathbf{v}}_n)^\top\boldsymbol{\beta}_n))-\varphi^{(k)}(F(e_{ni}))]^2\Bigg\}\\&=n^{-1}\sum_{i=1}^n\mathbb{E}\{(\mathbf{w}_{ni}-\bar{\mathbf{w}}_n)(\mathbf{w}_{ni}-\bar{\mathbf{w}}_n)^\top\\&\qquad\times\mathbb{E}[(\varphi^{(k)}(F(e_{ni}-(\mathbf{w}_{ni}-\bar{\mathbf{w}}_n)^\top\mathbf{b}_n^*-(\mathbf{v}_{ni}-\bar{\mathbf{v}}_n)^\top\boldsymbol{\beta}_n))\\&\qquad\qquad-\varphi^{(k)}(F(e_{ni})))^2|\mathbf{v}_{ni},\mathbf{x}_{ni}]\}.\end{aligned}\tag{3.14}$$

Hence,

$$\begin{aligned}\|\mathbf{A}_n^{(k)}\|&\le\left\{\left\|n^{-1}\sum_{i=1}^n(\mathbf{w}_{ni}-\bar{\mathbf{w}}_n)(\mathbf{w}_{ni}-\bar{\mathbf{w}}_n)^\top-(\mathbf{Q}+\mathbf{V})\right\|+\|\mathbf{Q}+\mathbf{V}\|\right\}\cdot\mathrm{o}(1)\\&=\{\|\mathbf{Q}+\mathbf{V}\|+\mathrm{o}(1)\}\cdot\mathrm{o}(1).\end{aligned}$$

This, together with the fact $\mathbb{E}\mathbf{T}_n^{(k)}(\mathbf{0},\mathbf{0})=\mathbf{0}$, implies

$$\|\mathbf{T}_n^{(k)}(\mathbf{b}_n^*,\boldsymbol{\beta}_n)-\mathbf{T}_n^{(k)}(\mathbf{0},\mathbf{0})-\mathbb{E}\mathbf{T}_n^{(k)}(\mathbf{b}_n^*,\boldsymbol{\beta}_n)\|\xrightarrow{p}0.\tag{3.15}$$

Furthermore, for any fixed k and for fixed $\mathbf{b}^{0*},\boldsymbol{\beta}^0$,

$$\mathbf{T}_n^{(k)}(\mathbf{b}_n^*,\boldsymbol{\beta}_n)-\mathbf{T}_n^{(k)}(\mathbf{0},\mathbf{0})+\gamma_k[(\mathbf{Q}+\mathbf{V})\mathbf{b}^{0*}+\mathbf{V}\boldsymbol{\beta}^0]\xrightarrow{p}\mathbf{0},\tag{3.16}$$

where

$$\gamma_k=-\int_{\mathbb{R}^1}\varphi^{(k)}(F(e))f'(e)\,\mathrm{d}e=-\int_0^1\varphi^{(k)}(u)\frac{f'(F^{-1}(u))}{f(F^{-1}(u))}\,\mathrm{d}u.$$

Indeed, (we put $\bar{\mathbf{x}}_n=\bar{\mathbf{v}}_n=\mathbf{0}$, for the sake of brevity)

$$n^{-1/2}\sum_{i=1}^n\mathbb{E}\{\mathbf{w}_{ni}(\mathbb{E}[\varphi^{(k)}(F(e_{ni}-n^{-1/2}(\mathbf{w}_{ni}^\top\mathbf{b}^{0*}-n^{-1/2}\mathbf{v}_{ni}^\top\boldsymbol{\beta}^0)-\varphi^k(F(e_{ni}))$$

$$
\begin{aligned}
&\qquad - \gamma_k(n^{-1/2}[\mathbf{w}_{ni}^\top \mathbf{b}^{0*} + \mathbf{v}_{ni}^\top \boldsymbol{\beta}^0])|\mathbf{v}_{ni}, \mathbf{x}_{ni}])\} \\
&= n^{-1/2} \sum_{i=1}^n \mathbb{E}\bigg\{\mathbf{w}_{ni}\bigg(\int_{\mathbb{R}^1} \varphi^{(k)}(F(z))\,\mathrm{d}[F(z + n^{-1/2}\mathbf{w}_{ni}^\top \mathbf{b}^{0*} + n^{-1/2}\mathbf{v}_{ni}^\top \boldsymbol{\beta}^0) - F(z)] \\
&\qquad - n^{-1/2}[\mathbf{w}_{ni}^\top \mathbf{b}^{0*} + \mathbf{v}_{ni}^\top \boldsymbol{\beta}^0] \int_{\mathbb{R}^1} \varphi^{(k)}(F(z)) f'(z)\,\mathrm{d}z\bigg)\bigg\} \\
&= n^{-1/2} \sum_{i=1}^n \mathbb{E}\bigg\{\mathbf{w}_{ni}\bigg(\int_{\mathbb{R}^1} \varphi^{(k)}(F(z))\,\mathrm{d}[F(z + n^{-1/2}\mathbf{w}_{ni}^\top \mathbf{b}^{0*} + n^{-1/2}\mathbf{v}_{ni}^\top \boldsymbol{\beta}^0) - F(z) \\
&\qquad - n^{-1/2}(\mathbf{w}_{ni}^\top \mathbf{b}^{0*} + \mathbf{v}_{ni}^\top \boldsymbol{\beta}^0) f(z)]\bigg)\bigg\} \to \mathbf{0}.
\end{aligned}
$$

Moreover, we have

$$
\begin{aligned}
(\gamma_k - \gamma)^2 &= \left\langle (\varphi_k - \varphi), -\frac{f'(F^{-1}(\cdot))}{f(F^{-1}(\cdot))} \right\rangle^2 \\
&\le \|\varphi_k - \varphi\|^2 \left\| -\frac{f'(F^{-1}(\cdot))}{f(F^{-1}(\cdot))} \right\|^2 \\
&= \mathcal{I}(f)\|\varphi_k - \varphi\|^2 \to 0 \qquad \text{as } k \to \infty.
\end{aligned} \tag{3.17}
$$

Using (3.16), (3.17), Lemmas 3.1 and 3.2, Corollary 3.1 and Lemma 3.5 in [18], we obtain that

$$
P(\|\mathbf{S}_n(\mathbf{b}_n^*, \boldsymbol{\beta}_n) - \mathbf{S}_n^{(k)}(\mathbf{b}_n^*, \boldsymbol{\beta}_n)\| > \varepsilon) < \varepsilon
$$

for $\forall k > k(\varepsilon)$, $\forall n > n(k)$, and finally we arrive at (3.10). □

3.4. Uniform asymptotic quadraticity of the Jaeckel dispersion

Recall that $\bar{a}_n = 0$ under (A.1). Rewrite the Jaeckel dispersion in the presence of measurement errors in the form

$$
\mathcal{D}_n(\mathbf{b}) = \sum_{i=1}^n [Y_{ni} - \mathbf{w}_{ni}^\top \mathbf{b}] a_n(R_{ni}(\mathbf{b}))
$$

or eventually in the form

$$
\mathcal{D}_n(\mathbf{b}^*, \boldsymbol{\beta}) = \sum_{i=1}^n [e_{ni} - (\mathbf{w}_{ni} - \bar{\mathbf{w}}_n)^\top \mathbf{b}^* - (\mathbf{v}_{ni} - \bar{\mathbf{v}}_n)^\top \boldsymbol{\beta}] a_n(R(e_{ni} - \mathbf{w}_{ni}^\top \mathbf{b}^* - \mathbf{v}_{ni}^\top \boldsymbol{\beta})),
$$

where $\mathbf{b}^* = \mathbf{b} - \boldsymbol{\beta}$. It is a piecewise linear, convex function of $\mathbf{b}$ and $\mathbf{b}^*$, respectively. Hence, the minimum $\widehat{\boldsymbol{\beta}}_n = \arg\min_{\mathbf{b} \in \mathbb{R}^p} \mathcal{D}_n(\mathbf{b})$ exists, and is considered as an estimate of

$\boldsymbol{\beta}$ in model (2.1). By [17], the partial derivatives of $\mathcal{D}_n(\mathbf{b})$ exist for almost all $\mathbf{b}$, and where they exist, are equal to

$$\frac{\partial}{\partial b_j}\mathcal{D}_n(\mathbf{b}) = -n^{1/2}S_{nj}(\mathbf{b}) = -\sum_{i=1}^n (w_{nij}-\bar{w}_j)a_n(Y_i - \mathbf{w}_{ni}\mathbf{b}), \qquad j=1,\dots,p.$$

Otherwise speaking,

$$\nabla\mathcal{D}_n(\mathbf{b}^*,\boldsymbol{\beta}) = -n^{1/2}\mathbf{S}_n(\mathbf{b}^*,\boldsymbol{\beta}) = -\sum_{i=1}^n (\mathbf{w}_{ni}-\bar{\mathbf{w}}_n)a_n(R(e_{ni}-\mathbf{w}_{ni}^\top\mathbf{b}^* - \mathbf{v}_{ni}^\top\boldsymbol{\beta})),$$

where ∇ denotes the subgradient.

Consider the quadratic function

$$\mathcal{C}_n(\mathbf{b}^*,\boldsymbol{\beta}) = \tfrac{1}{2}\gamma\mathbf{b}^{*\top}(\mathbf{Q}+\mathbf{V})\mathbf{b}^* - \mathbf{b}^{*\top}\mathbf{S}_n(\mathbf{0}) + \gamma\mathbf{b}^*\mathbf{V}\boldsymbol{\beta} + \mathcal{D}_n(\mathbf{0}).$$

Then $\mathcal{D}_n(\mathbf{b})$ and $\mathcal{C}_n(\mathbf{b})$ are both convex functions and $\mathcal{D}_n(\mathbf{0}) = \mathcal{C}_n(\mathbf{0})$. Moreover,

$$\nabla[\mathbf{D}_n(\mathbf{b}^*,\boldsymbol{\beta}) - \mathcal{C}_n(\mathbf{b}^*,\boldsymbol{\beta})] = -[n^{1/2}(\mathbf{S}_n((\mathbf{b}^*,\boldsymbol{\beta})) - \mathbf{S}_n(\mathbf{0},\mathbf{0}) + \gamma(\mathbf{Q}+\mathbf{V})\mathbf{b}^* + \gamma\mathbf{V}\boldsymbol{\beta})].$$

Hence, it follows from (3.10) that for $\mathbf{b}_n^* = n^{-1/2}\mathbf{b}^{0*}, \boldsymbol{\beta}_n = n^{-1/2}\boldsymbol{\beta}^0$ with $\mathbf{b}^{0*},\boldsymbol{\beta}^0 \in \mathbb{R}^p$ fixed that

$$\|\nabla[\mathbf{D}_n(n^{-1/2}\mathbf{b}^{0*}, n^{-1/2}\boldsymbol{\beta}^0) - \mathcal{C}_n(n^{-1/2}\mathbf{b}^{0*}, n^{-1/2}\boldsymbol{\beta}^0)]\| \xrightarrow{p} 0.$$

Using the convexity arguments in [13] (Appendix) and [29] (Convexity lemma), we conclude that

$$\sup|\mathcal{D}_n(n^{-1/2}\mathbf{b}^{0*}, n^{-1/2}\boldsymbol{\beta}^0) - \tfrac{1}{2}\gamma\mathbf{b}^{0*\top}(\mathbf{Q}+\mathbf{V})\mathbf{b}^{0*} + \mathbf{b}^{0*\top}\mathbf{S}_n(\mathbf{0}) - \gamma\mathbf{b}^{0*}\mathbf{V}\boldsymbol{\beta}^0 + \mathcal{D}_n(\mathbf{0})| = o_p(1),$$

where the supremum is taken over the set $\{\|\mathbf{b}^{0*}\| \le C, \|\boldsymbol{\beta}^0\| \le C\}$. Hence, following the arguments in the proof of Theorem 1 in [29], we conclude that, under the local alternative $\boldsymbol{\beta}_n = n^{-1/2}\boldsymbol{\beta}^0$,

$$\arg\min_{\mathbf{b}^{0*}} \mathcal{D}_n(n^{-1/2}\mathbf{b}^{0*}, n^{-1/2}\boldsymbol{\beta}^0)$$

is asymptotically equivalent to

$$\arg\min_{\mathbf{b}^{0*}}\left[\frac{1}{2}\gamma\mathbf{b}^{0*\top}(\mathbf{Q}+\mathbf{V})\mathbf{b}^{0*} - \mathbf{b}^{0*\top}\mathbf{S}_n(\mathbf{0}) + \gamma\mathbf{b}^{0*}\mathbf{V}\boldsymbol{\beta}^0\right]. \tag{3.18}$$

The solution of (3.18) equals to

$$\mathbf{b}^{0*} = \mathbf{b}^0 - \boldsymbol{\beta}^0 = n^{1/2}(\widehat{\boldsymbol{\beta}}_n - \boldsymbol{\beta}_n) = \gamma^{-1}(\mathbf{Q}+\mathbf{V})^{-1}\mathbf{S}_n(\mathbf{0},\mathbf{0}) - (\mathbf{Q}+\mathbf{V})^{-1}\mathbf{V}\boldsymbol{\beta}^0.$$

Hence, in the linear model with local value of regression parameter β, when

$$Y_{ni} = \mathbf{x}_{ni}^\top\boldsymbol{\beta}_n + e_{ni}, \qquad \boldsymbol{\beta}_n = n^{-1/2}\boldsymbol{\beta}^0,$$

when we observe only $\mathbf{w}_{ni} = \mathbf{x}_{ni} + \mathbf{v}_{ni}$ instead of $\mathbf{x}_{ni}, i = 1, \ldots, n$, the R-estimator is asymptotically normally distributed with a bias $\mathbf{B} = -(\mathbf{Q} + \mathbf{V})^{-1}\mathbf{V}\boldsymbol{\beta}^0$, that is,

$$n^{1/2}(\widehat{\boldsymbol{\beta}}_n - n^{-1/2}\boldsymbol{\beta}^0) \xrightarrow{\mathcal{D}} \mathcal{N}_p(B, (\mathbf{Q} + \mathbf{V})^{-1}A^2(\varphi)), \qquad \mathbf{B} = -(\mathbf{Q} + \mathbf{V})^{-1}\mathbf{V}\boldsymbol{\beta}^0. \tag{3.19}$$

Finally, as we have already mentioned, all of the above arguments and motivations are valid when we replace e_{ni} with $e_{ni} + u_{ni}, i = 1, \ldots, n$. This completes the proof of Theorem 2.1.

4. Numerical illustration

The following simulation study illustrates the effect of measurement errors in regressors on the finite-sample performance of R-estimates. Empirical bias (and variance) of R-estimates are computed and compared for various measurement error models. For the sake of comparison, the biases and variances are also computed for the least squares estimate (LSE) and the least absolute deviation (L_1) estimate, under the same setup. Moreover, we compare the deterministic and random regressors.

All the simulations were performed in the statistical software `R` using standard tools and libraries. For minimization of (2.4) functions `optimize` and `optim` with initial estimate 0.5 – regression quantile were used. The random numbers generator was setup with the initial value `set.seed(15).`

The results illustrate that the bias of R-estimate is surprisingly stable with respect to the sample size; the bias corresponding to small n is comparable to the asymptotic one derived in Theorem 2.1.

Notice that the bias of R-estimator only slightly differs from the biases of LSE and L_1-estimators.

4.1. Regression line

Consider first the model of regression line

$$Y_i = \beta_0 + x_i\beta_1 + e_i, \qquad i = 1, \ldots, n,$$

where the Y_i are measured accurately, while instead of x_i we observe only $w_i = x_i + v_i, i = 1, \ldots, n$. The R-estimator of parameter β_1 is based on Wilcoxon scores generated by score function $\varphi(u) = u - 1/2$.

All the simulation results are based on 10 000 replications, parameters were chosen as $\beta_0 = 1, \beta_1 = 2$, and model errors e_i follow the logistic distribution. In Tables 1 and 2, the empirical bias of R-estimator based on Wilcoxon scores is compared for various sample sizes ($n = 10, \ldots, 1000$) and with the theoretical asymptotic result ($n = \infty$). The regressors x_i are deterministic in Table 1; they were generated from uniform $\mathcal{U}(-3, 9)$ distribution once for all experiment and then considered as fixed. The regressors in Table 2 are random; each time they were generated also from uniform distribution $\mathcal{U}(-3, 9)$. This

Table 1. Empirical bias of R-estimator for various n and measurement errors v_i; nonrandom regressors x_i

	n							
v_i	10	20	50	100	200	500	1000	∞
0	0.002	0.001	0.000	0.000	0.000	0.000	0.000	0.000
$\mathcal{U}(-5,0)$	−0.264	−0.295	−0.305	−0.297	−0.302	−0.306	−0.307	−0.296
$\mathcal{U}(0,9)$	−0.684	−0.727	−0.732	−0.714	−0.719	−0.727	−0.728	−0.720
$\mathcal{U}(-3,9)$	−0.982	−1.013	−1.006	−0.983	−0.986	−0.995	−0.995	−1.000
$\mathcal{N}(0,1)$	−0.128	−0.148	−0.150	−0.146	−0.148	−0.151	−0.152	−0.154
$\mathcal{N}(0,2)$	−0.440	−0.483	−0.488	−0.476	−0.480	−0.487	−0.488	−0.500
$\mathcal{N}(0,3)$	−0.790	−0.836	−0.837	−0.819	−0.822	−0.832	−0.833	−0.857

enables to see the difference between deterministic and random regressors: the bias differs more from its asymptotic value in case of deterministic regressors than in case of random regressors; it can be caused by the slower rate of convergence. The measurement errors v_i are either uniformly or normally distributed ($i = 1, \dots, n$).

Table 3 compares empirical bias and variance (in parenthesis) of R-estimator based on Wilcoxon scores, of LSE and L_1-estimate for the sample size $n = 50$ and when regressors x_i are random, generated from uniform $\mathcal{U}(-3,9)$ distribution; model errors e_i are generated from normal, logistic, Laplace, Pareto with parameter $\alpha = 0.9$ and Cauchy distributions. The measurement errors v_i follow various distributions, similarly as in Tables 1 and 2.

Table 2. Empirical bias of R-estimator for various n and measurement errors v_i; random regressors x_i

	n							
v_i	10	20	50	100	200	500	1000	∞
0	0.004	−0.001	0.000	0.000	0.000	0.000	0.000	0.000
$\mathcal{U}(-5,0)$	−0.283	−0.297	−0.305	−0.306	−0.307	−0.309	−0.309	−0.296
$\mathcal{U}(0,9)$	−0.711	−0.722	−0.728	−0.730	−0.730	−0.732	−0.732	−0.720
$\mathcal{U}(-3,9)$	−0.998	−1.000	−0.999	−1.000	−0.999	−1.000	−1.000	−1.000
$\mathcal{N}(0,1)$	−0.138	−0.149	−0.150	−0.153	−0.153	−0.153	−0.153	−0.154
$\mathcal{N}(0,2)$	−0.462	−0.481	−0.487	−0.489	−0.491	−0.492	−0.492	−0.500
$\mathcal{N}(0,3)$	−0.813	−0.830	−0.833	−0.835	−0.837	−0.837	−0.838	−0.857

Table 3. Empirical bias (variance) of R-estimator, LSE and L_1-estimator for various measurement errors v_i and model errors e_i; $n = 50$

	e_i				
v_i	Normal	Logistic	Laplace	Pareto	Cauchy
	0.002 (0.182)	0.008 (0.527)	−0.002 (0.254)	0.000 (0.416)	0.018 (0.672)
0	0.004 (0.172)	0.009 (0.567)	−0.003 (0.355)	5.047 (89 200)	−3.289 (85 700)
	0.002 (0.274)	0.010 (0.708)	−0.002 (0.249)	0.000 (1.191)	0.018 (0.526)
	−0.399 (0.155)	−0.398 (0.438)	−0.396 (0.214)	−0.401 (0.422)	−0.404 (0.568)
$\mathcal{U}(-3,3)$	−0.395 (0.147)	−0.396 (0.466)	−0.394 (0.283)	−7.137 (604 000)	22.62 (4 000 000)
	−0.401 (0.232)	−0.400 (0.591)	−0.400 (0.235)	−0.422 (0.932)	−0.405 (0.456)
	−0.995 (0.101)	−1.006 (0.278)	−0.997 (0.142)	−1.001 (0.309)	−1.010 (0.397)
$\mathcal{U}(-6,6)$	−0.995 (0.096)	−1.009 (0.294)	−0.998 (0.182)	−7.259 (401 000)	0.933 (36 400)
	−0.995 (0.151)	−1.006 (0.376)	−0.995 (0.157)	−1.001 (0.587)	−1.014 (0.320)
	−0.153 (0.174)	−0.161 (0.493)	−0.145 (0.243)	−0.149 (0.439)	−0.147 (0.638)
$\mathcal{N}(0,1)$	−0.152 (0.163)	−0.158 (0.523)	−0.145 (0.328)	−2.136 (281 000)	−7.380 (739 000)
	−0.153 (0.261)	−0.159 (0.675)	−0.147 (0.259)	−0.165 (1.092)	−0.138 (0.510)

4.2. Model of two regressors

Consider the model

$$Y_i = \beta_0 + x_{i,1}\beta_1 + x_{i,2}\beta_2 + e_i, \qquad i = 1, \ldots, n,$$

where again the Y_i are measured accurately, but instead of $\mathbf{x}_i$ we observe only $\mathbf{w}_i = \mathbf{x}_i + \mathbf{v}_i, i = 1, \ldots, n$. The R-estimator of parameter $\boldsymbol{\beta} = (\beta_1, \beta_2)^\top$ is based on Wilcoxon scores generated by score function $\varphi(u) = u - 1/2$.

Here we chose $n = 50$, parameters $\beta_0 = 1, \beta_1 = 2, \beta_2 = 1$, random regressors $\mathbf{x}_i = (x_{i,1}, x_{i,2})^\top$ are generated from 2-dimensional normal distributions $\mathcal{N}_2(\boldsymbol{\mu}, \mathbf{S}_\nu), \nu = 1, 2, 3$, where $\boldsymbol{\mu} = (0, 1)^\top$ and

$$\mathbf{S}_1 = \begin{pmatrix} 4 & 0.5 \\ 0.5 & 2 \end{pmatrix}, \qquad \mathbf{S}_2 = \begin{pmatrix} 2 & 0.2 \\ 0.2 & 2 \end{pmatrix}, \qquad \mathbf{S}_3 = \begin{pmatrix} 1 & 0.9 \\ 0.9 & 1 \end{pmatrix}.$$

Table 4 compares empirical bias and variance (in parentheses) of R-estimator based on Wilcoxon scores, with those of the LSE and L_1-estimator for various distributions of the measurement errors $\mathbf{v}_i$ and model errors e_i.

We have also computed R-estimates generated by other score functions, for example, van der Waerden, median; also another simulation design was considered – various sample sizes n, values of the parameters, distributions of regressors, measurement errors v_i and u_i and model errors. It is of interest that the results for corresponding R-estimates are quite similar to those presented in the previous tables.

The simulation study confirms that R-estimates in measurement error models are biased, as well as other usual estimates. The bias is relatively stable with respect to the

Table 4. Empirical bias (variance) of R-estimator, LSE and L_1-estimator for various measurement errors $\mathbf{v}_i$ and model errors e_i; $n=50$

		e_i				
v_i		Normal	Logistic	Laplace	Pareto	Cauchy
$\mathbf{0}$	$\widehat{\beta}_1$	0.000 (0.600)	−0.015 (1.688)	0.002 (0.821)	−0.002 (0.017)	0.026 (2.326)
		−0.001 (0.569)	−0.019 (1.789)	0.006 (1.120)	21.73 (5 330 000)	−3.698 (48 700)
		−0.007 (0.864)	−0.008 (2.295)	0.002 (0.857)	−0.003 (0.034)	0.035 (1.843)
	$\widehat{\beta}_2$	0.020 (1.176)	0.026 (3.497)	0.001 (1.678)	−0.003 (0.033)	0.037 (4.634)
		0.014 (1.117)	0.027 (3.725)	−0.001 (2.250)	−29.28 (9 770 000)	0.695 (66 200)
		0.031 (1.744)	0.017 (4.588)	−0.001 (1.758)	−0.005 (0.067)	0.007 (3.618)
$\mathcal{N}_2(\boldsymbol{\mu},\mathbf{S}_3)$	$\widehat{\beta}_1$	−0.362 (0.554)	−0.379 (1.603)	−0.379 (0.798)	−0.372 (0.087)	−0.347 (2.161)
		−0.359 (0.528)	−0.385 (1.693)	−0.376 (1.030)	22.81 (6 270 000)	−4.118 (48 500)
		−0.365 (0.823)	−0.369 (2.134)	−0.375 (0.881)	−0.370 (0.102)	−0.338 (1.758)
	$\widehat{\beta}_2$	−0.774 (0.936)	−0.738 (2.662)	−0.754 (1.295)	−0.770 (0.136)	−0.738 (3.618)
		−0.776 (0.881)	−0.734 (2.832)	−0.757 (1.696)	−26.76 (7 920 000)	−0.268 (78 500)
		−0.769 (1.402)	−0.754 (3.366)	−0.754 (1.409)	−0.769 (0.164)	−0.752 (2.935)
$\mathcal{N}_2(\boldsymbol{\mu},\mathbf{S}_2)$	$\widehat{\beta}_1$	−0.643 (0.419)	−0.652 (1.155)	−0.648 (0.579)	−0.649 (0.076)	−0.626 (1.622)
		−0.640 (0.399)	−0.655 (1.216)	−0.653 (0.750)	14.66 (3 070 000)	−3.987 (47 200)
		−0.647 (0.615)	−0.645 (1.573)	−0.642 (0.657)	−0.650 (0.089)	−0.623 (1.317)
	$\widehat{\beta}_2$	−0.495 (0.643)	−0.474 (1.730)	−0.477 (0.866)	−0.494 (0.107)	−0.466 (2.426)
		−0.497 (0.605)	−0.469 (1.843)	−0.471 (1.139)	−17.59 (3 660 000)	−0.646 (55 900)
		−0.489 (0.948)	−0.477 (2.282)	−0.486 (0.944)	−0.492 (0.129)	−0.478 (1.992)
$\mathcal{N}_2(\boldsymbol{\mu},\mathbf{S}_1)$	$\widehat{\beta}_1$	−0.997 (0.329)	−1.013 (0.879)	−1.010 (0.448)	−1.005 (0.071)	−0.987 (1.264)
		−0.999 (0.311)	−1.015 (0.931)	−1.011 (0.577)	9.435 (1 260 000)	−2.474 (40 500)
		−0.994 (0.484)	−1.009 (1.192)	−1.011 (0.509)	−1.005 (0.086)	−0.998 (1.037)
	$\widehat{\beta}_2$	−0.505 (0.670)	−0.493 (1.797)	−0.496 (0.917)	−0.499 (0.141)	−0.482 (2.489)
		−0.505 (0.630)	−0.486 (1.883)	−0.487 (1.168)	−21.75 (5 680 000)	0.106 (49 600)
		−0.501 (0.980)	−0.520 (2.374)	−0.504 (1.024)	−0.500 (0.170)	−0.501 (2.064)

sample size and to distribution of model errors. The R-estimates provide meaningful results as long as the e_i have a finite Fisher information; even under the normal errors are their empirical variances only slightly greater than that of LSE. The bias and other properties of R-estimates are comparable with those of the least squares and of L_1 estimates unless the distribution of model errors e_i is heavy, where the LSE fails. Generally, the reduction of the bias is rather a matter of measurement precision, of calibration and repeated measurements.

Acknowledgements

The authors thank the Referee for his/her comments, which helped to better understanding the text.

The work of R. Navrátil and J. Jurečková was partially done during their visits to Michigan State University, hosted by the Department of Statistics and Probability. Research of J. Jurečková was partially supported by the Grant GAČR 201/12/0083. Research of H.L. Koul was supported in part by the NSF DMS Grant 1250271. Research of R. Navrátil was supported by the Grant SVV-2013-267 315, and J. Picek and R. Navrátil were supported by the Project Klimatext CZ.1.07/2.3.00/20.0086.

References

[1] Adcock, R.J. (1877). Note on the method of least squares. *The Analyst* **4** 183–184.

[2] Akritas, M.G. and Bershady, M.A. (1996). Linear regression for astronomical data with measurement errors and intrinsic scatter. *Astrophysical Journal* **470** 706–728.

[3] Arias, O., Hallock, K.F. and Sosa-Escudero, W. (2001). Individual heterogeneity in the returns to schooling: Instrumental variables quantile regression using twins data. *Empirical Economics* **26** 7–40.

[4] Carroll, R.J., Delaigle, A. and Hall, P. (2007). Non-parametric regression estimation from data contaminated by a mixture of Berkson and classical errors. *J. R. Stat. Soc. Ser. B Stat. Methodol.* **69** 859–878. MR2368574

[5] Carroll, R.J., Maca, J.D. and Ruppert, D. (1999). Nonparametric regression in the presence of measurement error. *Biometrika* **86** 541–554. MR1723777

[6] Carroll, R.J., Ruppert, D., Stefanski, L.A. and Crainiceanu, C.M. (2006). *Measurement Error in Nonlinear Models. A Modern Perspective*, 2nd ed. *Monographs on Statistics and Applied Probability* **105**. Boca Raton, FL: Chapman & Hall/CRC. MR2243417

[7] Cheng, C.-L. and Van Ness, J.W. (1999). *Statistical Regression with Measurement Error. Kendall's Library of Statistics* **6**. London: Arnold. MR1719513

[8] Fan, J. and Truong, Y.K. (1993). Nonparametric regression estimation involving errors-in-variables. *Ann. Statist.* **21** 23–37.

[9] Fuller, W.A. (1987). *Measurement Error Models. Wiley Series in Probability and Mathematical Statistics: Probability and Mathematical Statistics*. New York: Wiley. MR0898653

[10] Hájek, J. and Šidák, Z. (1967). *Theory of Rank Tests*. New York: Academic Press. MR0229351

[11] Hausman, J. (2001). Mismeasured variables in econometric analysis: Problems from the right and problems from the left. *J. Econ. Perspect.* **15** 57–67.

[12] He, X. and Liang, H. (2000). Quantile regression estimates for a class of linear and partially linear errors-in-variables models. *Statist. Sinica* **10** 129–140. MR1742104

[13] Heiler, S. and Willers, R. (1988). Asymptotic normality of R-estimates in the linear model. *Statistics* **19** 173–184. MR0945375

[14] Hodges, J.L. Jr. and Lehmann, E.L. (1963). Estimates of location based on rank tests. *Ann. Math. Statist.* **34** 598–611. MR0152070

[15] Hyk, W. and Stojek, Z. (2013). Quantifying uncertainty of determination by standard additions and serial dilutions methods taking into account standard uncertainties in both axes. *Anal. Chem.* **85** 5933–5939.

[16] Hyslop, D.R. and Imbens, G.W. (2001). Bias from classical and other forms of measurement error. *J. Bus. Econom. Statist.* **19** 475–481. MR1963378

[17] Jaeckel, L.A. (1972). Estimating regression coefficients by minimizing the dispersion of the residuals. *Ann. Math. Statist.* **43** 1449–1458. MR0348930

[18] Jurečková, J. (1969). Asymptotic linearity of a rank statistic in regression parameter. *Ann. Math. Statist.* **40** 1889–1900. MR0248931

[19] Jurečková, J. (1971). Nonparametric estimate of regression coefficients. *Ann. Math. Statist.* **42** 1328–1338. MR0295487

[20] Jurečková, J., Picek, J. and Saleh, A.K.Md.E. (2010). Rank tests and regression and rank score tests in measurement error models. *Comput. Statist. Data Anal.* **54** 3108–3120. MR2727738

[21] Kelly, B.C. (2007). Some aspects of measurement error in linear regression of astronomical data. *The Astrophysical Journal* **665** 1489–1506.

[22] Koul, H.L. (2002). *Weighted Empirical Processes in Dynamic Nonlinear Models. Lecture Notes in Statistics* **166**. New York: Springer. MR1911855

[23] Marques, T.A. (2004). Predicting and correcting bias caused by measurement error in line transect sampling using multiplicative error models. *Biometrics* **60** 757–763. MR2089452

[24] Müller, I. (1996). Robust methods in the linear calibration model. Ph.D. thesis, Charles Univ. in Prague.

[25] Navrátil, R. (2012). Rank Tests and R-estimates in Location Model with Measurement errors. In *Proceedings of Workshop of the Jaroslav Hájek Center and Financial Mathematics in Practice I. Book of Short Papers* (J. Zelinka and J. Horová, eds.). Brno: Masaryk Univ.

[26] Navrátil, R. and Saleh, A.K.Md.E. (2011). Rank tests of symmetry and R-estimation of location parameter under measurement errors. *Acta Univ. Palack. Olomuc. Fac. Rerum Natur. Math.* **50** 95–102. MR2920711

[27] Oosterhoff, J. and van Zwet, W.R. (1979). A note on contiguity and Hellinger distance. In *Contributions to Statistics. Jaroslav Hájek Memorial Volume* (J. Jurečková, ed.) 157–166. Dordrecht: Reidel. MR0561267

[28] Picek, J. (1996). Statistical procedures based on regression rank scores. Ph.D. thesis, Charles Univ. in Prague.

[29] Pollard, D. (1991). Asymptotics for least absolute deviation regression estimators. *Econometric Theory* **7** 186–199. MR1128411

[30] Rocke, D.M. and Lorenzato, S. (1995). A two-component model for measurement error in analytical chemistry. *Technometrics* **37** 176–184.

[31] Saleh, A.K.Md.E., Picek, J. and Kalina, J. (2012). R-estimation of the parameters of a multiple regression model with measurement errors. *Metrika* **75** 311–328. MR2909549

[32] Sen, P.K., Jurečková, J. and Picek, J. (2013). Rank tests for corrupted linear models. *J. Indian Statist. Assoc.* **51** 201–229. MR3234614

Received February 2014 and revised October 2014