Nonparametric Statistics, Vol. 3, pp. 355–367
Reprints available directly from the publisher
Photocopying permitted by license only

SOME ASPECTS OF HADAMARD DIFFERENTIABILITY ON REGRESSION *L*-ESTIMATORS

JIAN-JIAN REN

University of Nebraska

We propose a modified version of Welsh's (1987) one-step L-estimator of regression to improve the stability of the estimator and simplify the computation of the estimator. This modified L-estimator of regression is equivalent to a Hadamard differentiable functional defined on the space $D[0, 1] \times D[0, 1]$. Such differentiability property of the estimator leads to its linear approximation in a straightforward way. Thereby, the asymptotic normality of the estimator is derived under less stringent conditions.

KEYWORDS: Asymptotic normality, Hadamard differentiability, L-estimation, linear models, statistical functional, weighted empirical processes.

1. INTRODUCTION

Among the robust alternatives of the classical estimators, which are less sensitive to the deviations from the classical assumptions, three broad classes play the most important role: M-estimators, L-estimators and R-estimators. Each of these estimators has advantages and disadvantages in robustness, efficiency, applicability and usability, depending on the situation (Huber, 1981). Linear combinations of order statistics, called *L-estimators,* have long been of interest as estimates for the classical location model, because of their computational simplicity, scale-equivariance and efficiency. However, the L-estimators do not have any straightforward extension to the linear regression model. Some attempts have been made to develop such extension by various people, such as Bickel (1973), Ruppert and Carroll (1980), Bassett and Koenker (1982), Koenker and Portnoy (1987) and Welsh (1987), among others. Particularly, Welsh (1987) proposed and investigated the asymptotic properties of a general class of one-step L-estimators of regression depending on a preliminary estimator, which carry over the robustness, efficiency and computational simplicity of the L-estimator for the location model to the linear model.

However, for the sake of convenience in proving its asymptotic properties, the one-step L-estimator of regression proposed by Welsh (1987) involves a matrix depending on the preliminary estimator. Simulation shows that such a matrix may cause some instability of the estimator, especially when the sample size is relatively small or iterations are performed (see Ren and Lu, 1994). In this paper, we propose a modified version of Welsh's L-estimator of regression through replacing this matrix by one which depends only on the design matrix. This modified L-estimator retains those robustness, efficiency and reparametrization invariance properties possessed by Welsh's estimator. Additionally, our modified L-estimator improves the stability of Welsh's estimator (Ren and Lu, 1994) and is easier to compute than Welsh's estimator. We should note that Carroll and Welsh (1988) considered an estimator which is similar to our modified L-estimator and that Jurečková and Welsh (1990) studied the asymptotic properties of their estimator. Our estimator in this paper is proposed based on the differentiability of statistical functional induced by the estimator. While more discussions on our modified L-estimator can be found in Ren and Lu's (1994), here we focus on the application of von Mises (1947) method to L-estimators of regression. We use Hadamard differentiability properties of statistical functionals to study the asymptotic properties of our modified L-estimator of regression under weaker conditions.

We consider the following simple linear model:

$$X_i = \mathbf{c}_i^T \boldsymbol{\theta}_0 + e_i, \qquad i \geq 1 \tag{1.1}$$

where $\mathbf{c}_i^T = (1, c_{i2}, \ldots, c_{ip}) = (1, \mathbf{c}_{i*}^T)$ with $\sum_{i=1}^n c_{ik} = 0,\ k = 2, \ldots, p,$ are known p-vectors of regression constants with $p > 1$, $\boldsymbol{\theta}_0$ is the vector of unknown (regression) parameter to be estimated, and e_i are independent and identically distributed random variables (i.i.d.r.v.'s) with distribution function (d.f.) F. Our modified version of the L-estimators of regression proposed by Welsh (1987) is given by

$$\begin{aligned} \boldsymbol{\lambda}_n &= \hat{\boldsymbol{\theta}}_N + \begin{bmatrix} T(G_n) \\ \mathbf{0} \end{bmatrix} - \begin{bmatrix} 0 & \mathbf{0}^T \\ \mathbf{0} & \mathbf{C}_{n*}^{-1} \end{bmatrix} \sum_{i=1}^n \mathbf{c}_i \bigg\{ \int (I\{r_i \leq y\} - G_n(y)) h(G_n(y))\, \mathrm{d}y \\ &\quad + \sum_{j=1}^m \omega_j \phi_n(q_j)(I\{r_i \leq G_n^{-1}(q_j)\} - q_j) \bigg\} \\ &= \hat{\boldsymbol{\theta}}_n + \begin{bmatrix} T(G_n) \\ -\mathbf{C}_{n*}^{-1} \sum_{i=1}^n \mathbf{c}_{i*} \Big\{ \int I\{r_i \leq y\} h(G_n(y))\, \mathrm{d}y + \sum_{j=1}^m \omega_j \phi_n(q_j) I\{r_i \leq G_n^{-1}(q_j)\} \Big\} \end{bmatrix} \end{aligned} \tag{1.2}$$

where $r_i = X_i - \mathbf{c}_i^T \hat{\boldsymbol{\theta}}_n$, $1 \leq i \leq n$, are the residuals from a preliminary estimator $\hat{\boldsymbol{\theta}}_n$, $G_n(x) = n^{-1} \sum_{i=1}^n I\{r_i \leq x\}$; $\phi_n(t)$ is any pointwise consistent estimator of $\phi(t) = [F'(F^{-1}(t))]^{-1}$; h is the weight function; $\omega_1, \ldots, \omega_m$ are constant weights and $0 < q_1 < \cdots < q_m < 1$ with $m < \infty$;

$$T(G_n) = \int_0^1 G_n^{-1}(t) h(t)\, \mathrm{d}t + \int_0^1 G_n^{-1}(t)\, \mathrm{d}m(t), \qquad \text{with } m(t) = \sum_{i=1}^m \omega_i I\{q_i \leq t\};$$

$$\mathbf{C}_n = \sum_{i=1}^n \mathbf{c}_i \mathbf{c}_i^T = \mathbf{D}_n \mathbf{D}_n^T = \begin{bmatrix} n & \mathbf{0}^T \\ \mathbf{0} & \mathbf{C}_{n*} \end{bmatrix}, \qquad \text{with } \mathbf{D}_n = (\mathbf{c}_1, \ldots, \mathbf{c}_n)^T.$$

Note that, compared with the L-estimator of regression by Welsh (1987), who uses an inverse of the matrix: $\mathbf{C}_W = \sum_{i=1}^n \mathbf{c}_i \mathbf{c}_i^T \{h(G_n(r_i)) + \sum_{j=1}^m \omega_j\}$ in (1.2) instead of our matrix

$$\begin{bmatrix} n & \mathbf{0}^T \\ \mathbf{0} & \mathbf{C}_{n*}^{-1} \end{bmatrix},$$

our $\mathbf{C}_{n*}$ does not depend on the preliminary estimator $\hat{\boldsymbol{\theta}}_n$. This simplifies the computation of Welsh's estimator. Also note that our estimator given by (1.2) would be the same as the one given by Carroll and Welsh (1988) if we could put $G_n(G_n^{-1}(q))=q,\ 0<q<1$, and this holds to a good degree of approximation when n is large.

To consider a normalized form of (1.2), we denote a matrix

$$\mathbf{C}_n^o=\mathrm{Diag}\{\ \|\mathbf{d}_1\|,\ldots,\|\mathbf{d}_p\|\}=\begin{bmatrix}\sqrt{n} & \mathbf{0}^T\\ \mathbf{0} & \mathbf{C}_{n*}^o\end{bmatrix},$$

where $\mathbf{d}_i$ is the column vector of $\mathbf{D}_n$. Then, we notice that the normalized form of (1.2), $\mathbf{C}_n^o(\boldsymbol{\lambda}_n-\hat{\boldsymbol{\theta}}_n)$, is a statistical functional of two weighted empirical processes $\mathbf{S}_n^*(t,\mathbf{u})$ and $J_n^*(t,\mathbf{u})$, $t\in[0,1]$ for $\mathbf{u}=\mathbf{C}_n^o(\hat{\boldsymbol{\theta}}_n-\boldsymbol{\theta}_o)\in R^p$, viz.,

$$\begin{aligned}
&\mathbf{C}_n^o(\boldsymbol{\lambda}_n-\hat{\boldsymbol{\theta}}_n)\\
&\quad=\begin{bmatrix}\sqrt{n}\,T(G_n)\\ -\mathbf{Q}_{n*}^{-1}\sum_{i=1}^n\mathbf{c}_{ni*}\left\{\int I\{r_i\le y\}h(G_n(y))\,\mathrm{d}y+\sum_{j=1}^m\omega_j\phi_n(q_j)I\{r_i\le G_n^{-1}(q_j)\}\right\}\end{bmatrix}\\
&\quad=\tau_n(\mathbf{S}_n^*(\cdot,\mathbf{u}),J_n^*(\cdot,\mathbf{u}))\\
&\quad=\begin{bmatrix}\sqrt{n}\,\tau_1(J_n^*(\cdot,\mathbf{u}))\\ -\mathbf{Q}_{n*}^{-1}\{\tau_2(\mathbf{S}_n^*(\cdot,\mathbf{u}),J_n^*(\cdot,\mathbf{u}))+\tau_{3n}(\mathbf{S}_n^*(\cdot,\mathbf{u}),J_n^*(\cdot,\mathbf{u}))\}\end{bmatrix},
\end{aligned}\tag{1.3}$$

where

$$\mathbf{c}_{ni}=(\mathbf{C}_n^o)^{-1}\mathbf{c}_i=(1/\sqrt{n},\mathbf{c}_{ni*}^T)^T,\qquad \mathbf{Q}_n=(\mathbf{C}_n^o)^{-1}\mathbf{C}_n(\mathbf{C}_n^o)^{-1}=\begin{bmatrix}1 & \mathbf{0}^T\\ \mathbf{0} & \mathbf{Q}_{n*}\end{bmatrix};$$

$$\mathbf{S}_n^*(t,\mathbf{u})=\sum_{i=1}^n\mathbf{c}_{ni*}\,I\{Y_i\le\mathbf{c}_{ni}^T\mathbf{u}+F^{-1}(t)\},$$

$$J_n^*(t,\mathbf{u})=n^{-1}\sum_{i=1}^n I\{Y_i\le\mathbf{c}_{ni}^T\mathbf{u}+F^{-1}(t)\},$$

with $Y_i=X_i-\mathbf{c}_i^T\boldsymbol{\theta}_o$ (i.i.d.r.v.'s with d.f. F), and

$$\tau_1(J_n^*(\cdot,\mathbf{u}))=\int_0^1F^{-1}(J_n^{*-1}(t,\mathbf{u}),\mathbf{u})h(t)\,\mathrm{d}t+\int_0^1F^{-1}(J_n^{*-1}(t,\mathbf{u}),\mathbf{u})\,\mathrm{d}m(t),$$

$$\tau_2(\mathbf{S}_n^*(\cdot,\mathbf{u}),J_n^*(\cdot,\mathbf{u}))=\int_0^1\phi(t)\mathbf{S}_n^*(t,\mathbf{u})h(J_n^*(t,\mathbf{u}))\,\mathrm{d}t$$

$$\tau_{3n}(\mathbf{S}_n^*(\cdot,\mathbf{u}),J_n^*(\cdot,\mathbf{u}))=\int_0^1\mathbf{S}_n^*(J_n^{*-1}(t,\mathbf{u}),\mathbf{u})\phi_n(t)\,\mathrm{d}m(t),$$

with $J_n^{*-1}(x,\mathbf{u})=\inf\{y;J_n^*(y,\mathbf{u})\ge x\}$. We will show, in the proof of our main theorem, that the functional given by (1.3) is uniformly equivalent to a Hadamard differentiable functional τ defined on $D[0,1]\times D[0,1]$, i.e.,

$$\begin{aligned}
\mathbf{C}_n^o(\boldsymbol{\lambda}_n-\boldsymbol{\theta}_n)&\doteq\tau(\mathbf{S}_n^*(\cdot,\mathbf{u}),J_n^*(\cdot,\mathbf{u}))\\
&=\begin{bmatrix}\sqrt{n}\,\tau_1(J_n^*(\cdot,\mathbf{u}))\\ -\mathbf{Q}_{n*}^{-1}\{\tau_2(\mathbf{S}_n^*(\cdot,\mathbf{u}),J_n^*(\cdot,\mathbf{u}))+\tau_3(\mathbf{S}_n^*(\cdot,\mathbf{u}),J_n^*(\cdot,\mathbf{u}))\}\end{bmatrix},
\end{aligned}\tag{1.4}$$

where

$$\tau_3(\mathbf{S}_n^*(\cdot,\mathbf{u}),J_n^*(\cdot,\mathbf{u}))=\int_0^1 \mathbf{S}_n^*(J_n^{*-1}(t,\mathbf{u}),\mathbf{u})\phi(t)\,\mathrm{d}m(t).$$

Note that the essential idea of Von Mises method is linear approximation. Hence, from the results on Hadamard differentiability by Ren and Sen (1991) or Ren (1994), we immediately have that such functional τ is equivalent to a linear functional L as below:

$$\sup_{|\mathbf{u}|\le K}|\tau(\mathbf{S}_n^*(\cdot,\mathbf{u}),J_n^*(\cdot,\mathbf{u}))-\mathbf{C}_n^o(T(F),\mathbf{0}^T)T-L(\hat{\mathbf{S}}_n^*(\cdot,\mathbf{u}))|\xrightarrow{P}0,\qquad \text{as } n\to\infty, \tag{1.5}$$

where $K>0$ and $\hat{\mathbf{S}}_n^*(t,\mathbf{u})=\sum_{i=1}^n \mathbf{c}_{ni}I\{Y_i\le \mathbf{c}_{ni}^T\mathbf{u}+F^{-1}(t)\}$. Since the following Jurečková-uniformly asymptotic linearity for a linear functional was already established (Ren, 1994) under certain regularity conditions:

$$\sup_{|\mathbf{u}|\le K}|L(\hat{\mathbf{S}}_n^*(\cdot,\mathbf{u}))-L(\hat{\mathbf{S}}_n^*(\cdot,\mathbf{0}))-\mathbf{u}|\xrightarrow{P}0,\qquad \text{as } n\to\infty, \tag{1.6}$$

where $L(\hat{\mathbf{S}}_n^*(\cdot,\mathbf{0}))=\sum_{i=1}^n \mathbf{c}_{ni}\psi(Y_i)$ for a real function ψ, therefore the asymptotic properties of our modified L-estimator of regression follow easily from (1.5) and (1.6).

Along with the preliminary notion, the basic assumptions for our research are presented in Section 2. In Section 3, the asymptotic normality of our modified L-estimator of regression is derived by Hadamard differentiability approach under weaker conditions on the design matrix $\mathbf{D}_n$ than those required by Welsh (1987) and than those by Jurečková and Welsh (1990).

2. PRELIMINARIES AND ASSUMPTIONS

We write the components of $\mathbf{S}_n^*(t,\mathbf{u})$ for $t\in[0,1]$, $\mathbf{u}\in R^p$ and $\mathbf{c}_{ni}=(c_{ni1},\ldots,c_{nip})^T$ as below,

$$S_{nk}^*(t,\mathbf{u})=\sum_{i=1}^n c_{nik}I\{Y_i\le \mathbf{c}_{ni}^T\mathbf{u}+F^{-1}(t)\},\qquad k=2,\ldots,p.$$

Then,

$$\tau_2(S_{nk}^*(\cdot,\mathbf{u}),J_n^*(\cdot,\mathbf{u}))=\int_0^1 \phi(t)S_{nk}^*(t,\mathbf{u})h(J_n^*(t,\mathbf{u}))\,\mathrm{d}t,$$

$$\tau_{3n}(S_{nk}^*(\cdot,\mathbf{u}),J_n^*(\cdot,\mathbf{u}))=\int_0^1 S_{nk}^*(J_n^{*-1}(t,\mathbf{u}),\mathbf{u})\phi_n(t)\,\mathrm{d}m(t),$$

$$\tau_3(S_{nk}^*(\cdot,\mathbf{u}),J_n^*(\cdot,\mathbf{u}))=\int_0^1 S_{nk}^*(J_n^{*-1}(t,\mathbf{u}),\mathbf{u})\phi(t)\,\mathrm{d}m(t).$$

It is clear that for any fixed $\mathbf{u}$, τ_2 and τ_3 are functionals defined on $D[0,1]\times D[0,1]$. We will always denote U as the classical uniform d.f. on $[0,1]$ and

$\phi(t) = [F'(F^{-1}(t))]^{-1}$. Also, $\|\cdot\|$ and $|\cdot|$ will always stand for Euclidean norm and uniform norm for space R^p, respectively.

We impose throughout the following basic conditions on the model (1.1).

BASIC CONDITIONS ON THE MODEL

(i) $\mathbf{C}_n^o(\hat{\boldsymbol{\theta}}_n - \boldsymbol{\theta}_o)$ is bounded in probability;
(ii) $\sum_{i=1}^n c_{ik} = 0$, for $k = 2, \ldots, p$;
(iii) $\{\phi_n(q_i) - \phi(q_i)\} \xrightarrow{P} 0$, as $n \to \infty$, for $1 \le i \le m$;
(iv) $\int_0^1 h(t)\,\mathrm{d}t + \sum_{i=1}^m \omega_i = 1$;
(v) h is the difference of two nonnegative functions.

More assumptions required by our main theorem is given as below.

ASSUMPTIONS

(A1) $\overline{\lim}_{n\to\infty} \max_{1\le i\le n} \|\mathbf{c}_{ni}\|^2 = 0$;
(A2) There exists a positive definite $p \times p$ matrix $\mathbf{Q}$ such that $\lim_{n\to\infty} \mathbf{Q}_n = \mathbf{Q}$;
(B) F has a positive and uniformly continuous derivative F';
(C) h is bounded and continuous, a.e., with $h(t) = 0$ for $t \notin [\alpha, \beta]$, where $0 < \alpha < \beta < 1$.

Remark 1. It can be easily shown that our requirements on the design matrix $\mathbf{D}_n$ is weaker than Welsh's (1987) basic condition (ii) and than those by Jurečková and Welsh (1990). We should note however that Jurečková and Welsh (1990) studied stronger asymptotic properties of one-step L-estimator of regression. It is clear that meanwhile we do not require any stronger conditions on weight and error distribution F.

Remark 2. The choice of the pointwise consistent estimator ϕ_n of ϕ is discussed in Welsh's (1987).

We give the definition of Hadamard differentiability as below. More about Hadamard differentiability theory can be found in Fernholz's (1983).

Definition. Let V and W be the topological vector spaces and $L(V, W)$ be the set of continuous linear transformation from V to W. Let $\mathscr{A}$ be an open set of V, a functional τ: $\mathscr{A} \to W$ is *Hadamard Differentiable* (or *Compact Differentiable*) at $S \in \mathscr{A}$ if there exists $\tau'_S \in L(V, W)$ such that for any compact set Γ of V,

$$\lim_{t\to 0} \frac{\tau(S + tH) - \tau(S) - \tau'_S(tH)}{t} = 0$$

uniformly for any $H \in \Gamma$. The linear function τ'_S is called the *Hadamard Derivative* of τ at S.

For our current study, we consider the functional τ defined on the space $D[0, 1] \times D[0, 1]$, and denote the remainder term of the first order expansion as

$$\mathrm{Rem}(tH; \tau) = \tau(S + tH) - \tau(S) - \tau'_S(tH),$$

where $S = (U, U)$, $t \in R$ and $H \in D[0, 1] \times D[0, 1]$.

3. MAIN RESULTS

MAIN THEOREM. Assume the basic conditions (i)–(v) on the model (1.1), and assume (A1), (A2), (B) and (C), we have for $\boldsymbol{\lambda}_0 = \boldsymbol{\theta}_0 + (T(F), 0, \ldots, 0)^T$

$$\left\{\mathbf{C}_n^0(\boldsymbol{\lambda}_n - \boldsymbol{\lambda}_0) - \mathbf{Q}_n^{-1}\sum_{i=1}^{n} \mathbf{c}_{ni}\psi(Y_i)\right\} \xrightarrow{P} 0, \qquad \text{as } n \to \infty, \tag{3.1}$$

where

$$\psi(x) = -\int [I\{x \le y\} - F(y)]h(F(y))\,\mathrm{d}y - \sum_{i=1}^{m} \omega_i \phi(q_i)(I\{x \le F^{-1}(q_i)\} - q_i).$$

Therefore

$$\mathbf{C}_n^0(\boldsymbol{\lambda}_n - \boldsymbol{\lambda}_0) \xrightarrow{D} N_p(\mathbf{0}, \sigma^2\mathbf{Q}^{-1}), \qquad \text{as } n \to \infty \tag{3.2}$$

where $0 < \sigma^2 = \int \psi^2(x)\,\mathrm{d}F(x) < \infty$.

REMARK. Note that our Main Theorem for the modified Welsh's L-estimator of regression is essentially the same as Theorem 1 of Welsh (1987). The use of $\mathbf{C}_n^0$ (instead of $\sqrt{n}$) allows us to weaken the conditions on $\mathbf{D}_n$ for the asymptotic normality of L-estimator of regression.

The proof of Main Theorem will be given after the following lemmas and theorems.

LEMMA 1. Let $\tau_1 : D[0, 1] \to R$ be a functional defined by

$$\tau_1(G) = \int_0^1 F^{-1}(G^{-1}(t))h(t)\,\mathrm{d}t + \int_0^1 F^{-1}(G^{-1}(t))\,\mathrm{d}m(t),$$

where $h : [0, 1] \to R$ is bounded with $h(t) = 0$ for $0 < \alpha < t < \beta < 1$. Then τ_1 is Hadamard differentiable at U with derivative

$$\tau'_{1_U}(G) = -\int_0^1 \phi(t)G(t)h(t)\,\mathrm{d}t - \int_0^1 \phi(t)G(t)\,\mathrm{d}m(t), \qquad G \in D[0, 1].$$

Proof. The proof follows from Proposition 7.2.1 of Fernholz (1983). □

LEMMA 2. Let τ_2: $D[0, 1] \times D[0, 1] \to R$ be a functional defined by

$$\tau_2(G, H) = \int_0^1 \phi(t)G(t)h(H(t))\,\mathrm{d}t,$$

where h: $R \to R$ is continuous piecewise differentiable with bounded derivative, then τ_2 is Hadamard differentiable at (U, U) with derivative

$$\tau'_{2_{(U,U)}}(G, H) = \int_0^1 \phi(t)\{G(t)h(t) + h'(t)H(t)\}\,\mathrm{d}t, \qquad G, H \in D[0, 1].$$

Proof. The functional τ_2 can be expressed as a composition of the following Hadamard differentiable transformations:

γ_1: $D[0, 1] \times D[0, 1] \to D[0, 1] \times L^1[0, 1]$ defined by $\gamma_1(G, H) = (G, h(H))$, is, by Proposition 6.1.2 of Fernholz (1983), Hadamard differentiable at (U, U) with derivative

$$\gamma'_{1_{(U,U)}}(G, H) = (G, h'H).$$

γ_2: $D[0, 1] \times L^1[0, 1] \to L^1[0, 1]$ defined by $\gamma_2(G, H) = GH$, can be easily shown to be Hadamard differentiable at (U, h) with derivative

$$\gamma'_{2_{(U,h)}}(G, H) = Gh + H.$$

γ_3: $L^1[0, 1] \to R$ defined by $\gamma_3(G) = \int_0^1 G(t)\phi(t)\,\mathrm{d}t$, is Fréchet differentiable at Uh with derivative

$$\gamma'_{3_{Uh}}(G) = \int_0^1 G(t)\phi(t)\,\mathrm{d}t,$$

because it is a linear and continuous functional.

We have

$$\tau_2(G, H) = \gamma_3(\gamma_2(\gamma_1(G, H))).$$

Hence, by chain rule (Fernholz, 1983), τ_2 is Hadamard differentiable at (U, U) with derivative

$$\begin{aligned}\tau'_{2_{(U,U)}}(G, H) &= \gamma'_{3_{Uh}} \circ \gamma'_{2_{(U,h)}} \circ \gamma'_{1_{(U,U)}}(G, H)\\ &= \gamma'_{3_{Uh}} \circ \gamma'_{2_{(U,h)}}(G, h'H)\\ &= \gamma'_{3_{Uh}}(Gh + h'H) = \int_0^1 [G(t)h(t) + h'(t)H(t)]\phi(t)\,\mathrm{d}t. \quad \square\end{aligned}$$

LEMMA 3. Let τ_3: $D[0, 1] \times D[0, 1] \to R$ be a functional defined by

$$\tau_3(G, H) = \int_0^1 G(H^{-1}(t))\phi(t)\,\mathrm{d}m(t),$$

then τ_3 is Hadamard differentiable at (U, U) with derivative

$$\tau'_{3_{(U,U)}}(G, H) = \int_0^1 (G(t) - H(t))\phi(t)\,\mathrm{d}m(t), \qquad G, H \in D[0, 1].$$

Proof. The functional τ_3 can be expressed as a composition of the following Hadamard differentiable transformations:

γ_1: $D[0, 1] \times D[0, 1] \to D[0, 1] \times L^1[0, 1]$ defined by $\gamma_1(G, H) = (G, H^{-1})$, is, by Proposition 6.1.1 of Fernholz (1983), Hadamard differentiable at (U, U) with derivative

$$\gamma'_{1_{(U,U)}}(G, H) = (G, -H).$$

γ_2: $D[0, 1] \times L^1[0, 1] \to L^1[0, 1]$ defined by $\gamma_2(G, H) = G \circ H$ is, by Proposition 6.1.6 of Fernholz (1983), Hadamard differentiable at (U, U) with derivative

$$\gamma'_{2_{(U,U)}}(G, H) = G + H.$$

γ_3: $L^1[0, 1] \to R$ defined by $\gamma_3(G) = \int_0^1 G(t)\phi(t)\,\mathrm{d}m(t)$, is Fréchet differentiable with derivative

$$\gamma'_{3_U} = \int_0^1 G(t)\phi(t)\,\mathrm{d}m(t),$$

because it is a linear and continuous functional.

We have

$$\tau_3(G, H) = \gamma_3(\gamma_2(\gamma_1(G, H))).$$

Hence, by chain rule (Fernholz, 1983), τ_3 is Hadamard differentiable at (U, U) with derivative

$$\begin{aligned}\tau'_{3_{(U,U)}}(G, H) &= \gamma'_{3_U} \circ \gamma'_{2_{(U,U)}} \circ \gamma'_{1_{(U,U)}}(G, H)\\ &= \gamma'_{3_U} \circ \gamma'_{2_{(U,U)}}(G, -H) = \gamma'_{3_U}(G - H) = \int_0^1 [G(t) - H(t)]\phi(t)\,\mathrm{d}m(t). \quad \square\end{aligned}$$

LEMMA 4. Assume (B), we have for any $K > 0$,

$$\sup_{t \in [0,1],\, |\mathbf{u}| \le K} |J_n^*(t, \mathbf{u}) - t| \xrightarrow{P} 0, \qquad \text{as } n \to \infty.$$

Proof. By the proof of Theorem 3.1 of Ren (1994), we have $\sup\{\sqrt{n}(J_n^*(t, \mathbf{u}) - E\{J_n^*(t, \mathbf{u})\});\ t \in [0, 1], |\mathbf{u}| \le K\}$ is bounded in probability. Therefore, the proof follows from the facts:

$$\sup_{t \in [0,\,1],\, |\mathbf{u}| \le K} |E\{J_n^*(t, \mathbf{u})\} - E\{J_n^*(t, \mathbf{0})\}| \to 0, \qquad \text{as } n \to \infty,$$

and $E\{J_n^*(t, \mathbf{0})\} = t$, for $t \in [0, 1]$. $\square$

For the sake of convenience, we state, without proof, the following theorems of Ren (1994) on Hadamard differentiability and uniform linearity of a linear functional.

THEOREM 5. Suppose τ: $D[0, 1] \times D[0, 1] \to R$ is a functional and is Hadamard differentiable at (U, U). Assume (A1) and (B). Then, for any $K > 0$, $2 \le k \le p$, as $n \to \infty$

$$\sup_{|\mathbf{u}| \le K} \left| \sum_{i=1}^n c_{nik}^+ \operatorname{Rem}\left(\frac{S_{nk}^{*+}(\cdot, \mathbf{u})}{\sum_{i=1}^n c_{nik}^+} - U(\cdot), J_n^*(\cdot, \mathbf{u}) - U(\cdot); \tau\right)\right| \xrightarrow{P} 0,$$

and

$$\sup_{|\mathbf{u}| \le K} \left| \sum_{i=1}^n c_{nik}^- \operatorname{Rem}\left(\frac{S_{nk}^{*-}(\cdot, \mathbf{u})}{\sum_{i=1}^n c_{nik}^-} - U(\cdot), J_n^*(\cdot, \mathbf{u}) - U(\cdot); \tau\right)\right] \xrightarrow{P} 0,$$

for

$$c_{nik}^+ = \max\{0, c_{nik}\},\ c_{nik}^- = -\min\{0, c_{nik}\}, S_{nk}^{*+}(t, \mathbf{u}) = \sum_{i=1}^n c_{nik}^+ I\{Y_i \le \mathbf{c}_{ni}^T \mathbf{u} + F^{-1}(t)\}$$

and

$$S_{nk}^{*-}(t, \mathbf{u}) = \sum_{i=1}^n c_{nik}^- I\{Y_i \le \mathbf{c}_{ni}^T \mathbf{u} + F^{-1}(t)\}.$$

Therefore, as $n \to \infty$

$$\sup_{|\mathbf{u}|\le K}\left|\sum_{i=1}^{n} c_{nik}^{+}\left\{\tau\left(\frac{S_{nk}^{*+}(\cdot,\mathbf{u})}{\sum_{i=1}^{n} c_{nik}^{+}}, J_n^*(\cdot,\mathbf{u})\right) - \tau(U(\cdot),U(\cdot))\right\}\right.$$
$$\left. - \sum_{i=1}^{n} c_{nik}^{+}\tau'_{(U,U)}\left(\frac{S_{nk}^{*+}(\cdot,\mathbf{u})}{\sum_{i=1}^{n} c_{nik}^{+}} - U(\cdot), J_n^*(\cdot,\mathbf{u}) - U(\cdot)\right)\right| \xrightarrow{P} 0,$$

$$\sup_{|\mathbf{u}|\le K}\left|\sum_{i=1}^{n} c_{nik}^{-}\left\{\tau\left(\frac{S_{nk}^{*-}(\cdot,\mathbf{u})}{\sum_{i=1}^{n} c_{nik}^{-}}, J_n^*(\cdot,\mathbf{u})\right) - \tau(U(\cdot),U(\cdot))\right\}\right.$$
$$\left. - \sum_{i=1}^{n} c_{nik}^{-}\tau'_{(U,U)}\left(\frac{S_{nk}^{*-}(\cdot,\mathbf{u})}{\sum_{i=1}^{n} c_{nik}^{-}} - U(\cdot), J_n^*(\cdot,\mathbf{u}) - U(\cdot)\right)\right| \xrightarrow{P} 0,$$

and

$$\sup_{|\mathbf{u}|\le K}\left|\sum_{i=1}^{n} c_{nik}^{+}\tau\left(\frac{S_{nk}^{*+}(\cdot,\mathbf{u})}{\sum_{i=1}^{n} c_{nik}^{+}}, J_n^*(\cdot,\mathbf{u})\right)\right.$$
$$\left. - \sum_{i=1}^{n} c_{nik}^{-}\left\{\tau\left(\frac{S_{nk}^{*-}(\cdot,\mathbf{u})}{\sum_{i=1}^{n} c_{nik}^{-}}, J_n^*(\cdot,\mathbf{u})\right) - \tau'_{(U,U)}(S_{nk}^*(\cdot,\mathbf{u}),0)\right|\right. \xrightarrow{P} 0.$$

THEOREM 6. Assume (B) and assume a real function ϕ: $R \to R$, is bounded, nondecreasing and right or left continuous with $\int \phi\, dF = 0$. Let

$$\mathbf{L}_n(\mathbf{u}) = \sum_{i=1}^{n} \boldsymbol{\alpha}_{ni}\phi(Y_i - \boldsymbol{\beta}_{ni}^T\mathbf{u}),$$

where $\mathbf{u}, \boldsymbol{\alpha}_{ni}, \boldsymbol{\beta}_{ni} \in R^p$ with $\lim_{n\to\infty}\max_{1\le i\le n}\|\boldsymbol{\alpha}_{ni}\|^2 = 0$, $\lim_{n\to\infty}\max_{1\le i\le n}\|\boldsymbol{\beta}_{ni}\|^2 = 0$ and $\sum_{i=1}^{n}\|\boldsymbol{\alpha}_{ni}\|^2 \le M$, $\sum_{i=1}^{n}\|\boldsymbol{\beta}_{ni}\|^2 \le M$ for a constant $M > 0$. Then, for any $K > 0$,

$$\sup_{|\mathbf{u}|\le K}\left|[\mathbf{L}_n(\mathbf{u}) - \mathbf{L}_n(\mathbf{0})] + \sum_{i=1}^{n}\boldsymbol{\alpha}_{ni}\boldsymbol{\beta}_{ni}^T\mathbf{u}\gamma_\phi\right| \xrightarrow{P} 0, \qquad \text{as } n \to \infty,$$

where $0 < \gamma_\phi = \int F'\, d\phi < \infty$.

Proof of Main Theorem. First, we notice that, by the proof of Theorem 3.1 of Ren and Sen (1991) and Lemma 3.5, Lemma 3.6 of Ren (1994), we have for any $K > 0$,

$$\sup_{|\mathbf{u}|\le K,\, t\in[0,1]} |S_{nk}^*(t,\mathbf{u})| = O_p(1), \qquad k = 2, \ldots, p. \tag{3.3}$$

Therefore, the basic condition (iii), we have

$$\sup_{|\mathbf{u}|\le K} |\tau_{3n}(S_{nk}^*(\cdot,\mathbf{u}), J_n^*(\cdot,\mathbf{u})) - \tau_3(S_{nk}^*(\cdot,\mathbf{u}), J_n^*(\cdot,\mathbf{u}))| \xrightarrow{P} 0, \qquad \text{as } n \to \infty, \tag{3.4}$$

which gives us (1.4). We also notice that, by Lemma 1 and Theorem 5 for $c_{nik}^{+} = 1/\sqrt{n}$, we have,

$$\sup_{|\mathbf{u}|\le K} |\sqrt{n}\{\tau_1(J_n^*(\cdot,\mathbf{u})) - \tau_1(U(\cdot)) - \tau'_{1_U}(J_n^*(\cdot,\mathbf{u}) - U(\cdot))\}| \xrightarrow{P} 0, \qquad \text{as } n \to \infty.$$

Since $\tau_1(U) = T(F)$, and

$$\tau'_{1_U}(J_n^*(\cdot, \mathbf{u}) - U(\cdot)) = -\int_0^1 \phi(t)[J_n^*(t, \mathbf{u}) - t]h(t)\,\mathrm{d}t - \int_0^1 \phi(t)[J_n^*(t, \mathbf{u}) - t]\,\mathrm{d}m(t),$$
$$= n^{-1}\sum_{i=1}^n \psi(Y_i - \mathbf{c}_{ni}^T\mathbf{u}),$$

we have, as $n \to \infty$

$$\sup_{|\mathbf{u}|\le K}\left|\sqrt{n}\{\tau_1(J_n^*(\cdot, \mathbf{u}))) - T(F)\} - \frac{1}{\sqrt{n}}\sum_{i=1}^n \psi(Y_i - \mathbf{c}_{ni}^T\mathbf{u})\right| \xrightarrow{P} 0. \tag{3.5}$$

We will first prove the theorem under the assumption that h is continuous and piecewise differentiable with bounded derivative. From Lemma 1 and Lemma 2, we know that functional τ_j for $j = 2, 3$, are Hadamard differentiable at (U, U). Hence, by Theorem 5, we have that, for any $K > 0$ and $j = 2, 3$; $k = 2, \dots, p$,

$$\sup_{|\mathbf{u}|\le K}\left|\sum_{i=1}^n c_{nik}^+\,\tau_j\left(\frac{S_{nk}^{*+}(\cdot, \mathbf{u})}{\sum_{i=1}^n c_{nik}^+}, J_n^*(\cdot, \mathbf{u})\right)\right.$$
$$\left.-\sum_{i=1}^n c_{nik}^-\,\tau_j\left(\frac{S_{nk}^{*-}(\cdot, \mathbf{u})}{\sum_{i=1}^n c_{nik}^-}, J_n^*(\cdot, \mathbf{u})\right) - \tau'_{j_{(U,U)}}(S_{nk}^*(\cdot, \mathbf{u}), 0)\right| \xrightarrow{P} 0, \qquad \text{as } n \to \infty.$$

Since, for $j = 2, 3$ and $k = 2, \dots, p$,

$$\tau_j(S_{nk}^*(\cdot, \mathbf{u}), J_n^*(\cdot, \mathbf{u})) = \sum_{i=1}^n c_{nik}^+\tau_j\left(\frac{S_{nk}^{*+}(\cdot, \mathbf{u})}{\sum_{i=1}^n c_{nik}^+}, J_n^*(\cdot, \mathbf{u})\right)$$
$$-\sum_{i=1}^n c_{nik}^-\tau_j\left(\frac{S_{nk}^{*-}(\cdot, \mathbf{u})}{\sum_{i=1}^n c_{nik}^-}, J_n^*(\cdot, \mathbf{u})\right),$$

and

$$\tau'_{2_{(U,U)}}(S_{nk}^*(\cdot, \mathbf{u}), 0) = \int_0^1 \phi(t)S_{nk}^*(t, \mathbf{u})h(t)\,\mathrm{d}t = \sum_{i=1}^n c_{nik}\int_0^1 I\{F(Y_i - \mathbf{c}_{ni}^T\mathbf{u}\} \le t\}h(t)\phi(t)\,\mathrm{d}t,$$

$$\tau'_{3_{(U,U)}}(S_{nk}^*(\cdot, \mathbf{u}), 0) = \int_0^1 S_{nk}^*(t, \mathbf{u})\phi(t)\,\mathrm{d}m(t) = \sum_{i=1}^n\sum_{j=1}^m c_{nik}I\{F(Y_i - \mathbf{c}_{ni}^T\mathbf{u}) \le q_j\}\omega_j\phi(q_j),$$

by (1.3), (A2), (3.4) and (3.5), we have, as $n \to \infty$

$$\sup_{|\mathbf{u}|\le K}\left|\mathbf{C}_n^0(\boldsymbol{\lambda}_n - \boldsymbol{\theta}_n) - \mathbf{Q}_n^{-1}\begin{bmatrix}\sqrt{n}\,T(F) + \dfrac{1}{\sqrt{n}}\displaystyle\sum_{i=1}^n \psi(Y_i - \mathbf{c}_{ni}^T\mathbf{u}) \\ -\{\tau'_{1_{(U,U)}}(\mathbf{S}_n^*(\cdot, \mathbf{u}), 0) + \tau'_{2_{(U,U)}}(\mathbf{S}_n^*(\cdot, \mathbf{u}), 0)\}\end{bmatrix}\right| \xrightarrow{P} 0. \tag{3.6}$$

Note that if we denote, for $k = 2, \dots, p$

$$L(S_{nk}^*(\cdot, \mathbf{u})) = \tau'_{1_{(U,U)}}(S_{nk}^*(\cdot, \mathbf{u}), 0) + \tau'_{2_{(U,U)}}(S_{nk}^*(\cdot, \mathbf{u}), 0)$$
$$= \int_0^1 \phi(t)S_{nk}^*(t, \mathbf{u})h(t)\,\mathrm{d}t + \int_0^1 S_{nk}^*(t, \mathbf{u})\phi(t)\,\mathrm{d}m(t) = -\sum_{i=1}^n c_{nik}\psi(Y_i - \mathbf{c}_{ni}^T\mathbf{u}), \tag{3.7}$$

and for $k=1$,

$$L(S_{n1}^*(\cdot,\mathbf{u})) = -\sum_{i=1}^{n} c_{ni1}\psi(Y_i-\mathbf{c}_{ni}^T\mathbf{u}) = -n^{-1/2}\sum_{i=1}^{n}\psi(Y_i-\mathbf{c}_{ni}^T\mathbf{u}),$$

then L is a linear functional of $S_{nk}^*(\cdot,\mathbf{u})$ for any fixed $\mathbf{u}$. Since ψ is the sum of two monotone functions and left continuous with

$$\gamma_\psi = \int F'(t)\,\mathrm{d}\psi(t) = \int_0^1 h(t)\,\mathrm{d}t + \sum_{i=1}^{n}\omega_i = 1$$

and $\int \psi(x)\,\mathrm{d}F(x)=0$, from Theorem 6, we have for $k=1,2,\ldots,p$,

$$\sup_{|\mathbf{u}|\le K}\left|L(S_{nk}^*(\cdot,\mathbf{u})) - L(S_{nk}^*(\cdot,\mathbf{0})) - \sum_{i=1}^{n} c_{nik}\mathbf{c}_{ni}^T\mathbf{u}\right| \xrightarrow{P} 0, \qquad \text{as } n\to\infty. \tag{3.8}$$

Therefore, by (3.6) and (3.7), we have

$$\sup_{|\mathbf{u}|\le K}\left|\mathbf{C}_n^0(\boldsymbol{\lambda}_n-\hat{\boldsymbol{\theta}}_n) - \mathbf{Q}_n^{-1}\begin{bmatrix}\sqrt{n}\,T(F) + \dfrac{1}{\sqrt{n}}\displaystyle\sum_{i=1}^{n}\psi(Y_i-\mathbf{c}_{ni}^T\mathbf{u}) \\ -L(\mathbf{S}_n^*(\cdot,\mathbf{u}))\end{bmatrix}\right| \xrightarrow{P} 0,$$

furthermore, by (3.8), as $n\to\infty$

$$\sup_{|\mathbf{u}|\le K}\left|\mathbf{C}_n^0(\boldsymbol{\lambda}_n-\hat{\boldsymbol{\theta}}_n) - \mathbf{C}_n^0\begin{bmatrix}T(F)\\ \mathbf{0}\end{bmatrix} + \mathbf{Q}_n^{-1}\sum_{i=1}^{n}\mathbf{c}_{ni}\mathbf{c}_{ni}^T\mathbf{u} - \mathbf{Q}_n^{-1}\sum_{i=1}^{n}\mathbf{c}_{ni}\psi(Y_i)\right| \xrightarrow{P} 0.$$

Since $\mathbf{Q}_n=\sum_{i=1}^n \mathbf{c}_{ni}\mathbf{c}_{ni}^T$, and since $\mathbf{u}=\mathbf{C}_n^0(\hat{\boldsymbol{\theta}}_n-\boldsymbol{\theta}_0)$, by the basic condition (i), is bounded in probability, then

$$\mathbf{C}_n^0(\boldsymbol{\lambda}_n-\hat{\boldsymbol{\theta}}_n) + \mathbf{C}_n^0(\hat{\boldsymbol{\theta}}_n-\boldsymbol{\lambda}_0) - \mathbf{Q}_n^{-1}\sum_{i=1}^{n}\mathbf{c}_{ni}\psi(Y_i) = o_p(1).$$

Therefore (3.1) follows. By Hájek and Sidák (1967, page 153), (3.2) follows immediately from $\sum_{i=1}^n \mathbf{c}_{ni}\psi(Y_i) \xrightarrow{D} N_p(\mathbf{0},\sigma^2\mathbf{Q})$, as $n\to\infty$ and (A2).

To prove the theorem assuming that only condition (C) holds for h, it suffices to establish for $k=2,\ldots,p$

$$\sup_{|\mathbf{u}|\le K}\left|\tau_2(S_{nk}^*(\cdot,\mathbf{u}),J_n^*(\cdot,\mathbf{u})) - \int_0^1 \phi(t)S_{nk}^*(t,\mathbf{u})h(t)\,\mathrm{d}t\right| \xrightarrow{P} 0, \qquad \text{as } n\to\infty. \tag{3.9}$$

For any $\epsilon>0$, from Royden's (1968, page 68) and the continuity of h at 0 and 1, there exists a continuous function h_ϵ and a set $E\subset(\delta,1-\delta)$ for some $\delta>0$ such that

$$|h(t)-h_\epsilon(t)|\le\epsilon, \qquad \text{for all } t\in E^c,$$

and $\mu(E)<\epsilon/M_\delta$, where M_δ is an upper bound of $[F'(F^{-1}(t))]^{-1}$ for $t\in(\delta,1-\delta)$ and μ is Lebesgue measure on real line. Since, for a constant $M>0$,

$$Z_n(\mathbf{u}) = \int_0^1 |h(J_n^*(t,\mathbf{u})) - h_\epsilon(J_n^*(t,\mathbf{u}))|\,\phi(t)\,\mathrm{d}t \le \int_0^1 M\phi(t)\,\mathrm{d}t = M, \quad \text{for any } |\mathbf{u}|\le K,$$

and $Z_n(\mathbf{u}_n) = \sup_{|\mathbf{u}|\le K} Z_n(\mathbf{u})$, by Dominated Convergence Theorem and Lemma 4,

$$\sup_{|\mathbf{u}|\le K} \int_0^1 |h(J_n^*(t,\mathbf{u})) - h_\epsilon(J_n^*(t,\mathbf{u}))|\,\phi(t)\,\mathrm{d}t \xrightarrow{P} \int_0^1 |h(t) - h_\epsilon(t)|\phi(t)\,\mathrm{d}t, \quad as\ n\to\infty. \tag{3.10}$$

Since

$$\int_0^1 |h(t) - h_\epsilon(t)|\,\phi(t)\,\mathrm{d}t \le \epsilon + M\int_E \phi(t)\,\mathrm{d}t = \epsilon + M\int_{F(x)\in E} \mathrm{d}x = \epsilon + M\mu(F^{-1}(E)),$$

and since there exists $I_i = (a_i, b_i] \subset [\delta, 1-\delta]$, $i = 1, 2, \ldots,$ such that $E \subset \bigcup_{i=1}^\infty I_i$ and $\mu(E) \le \sum_{n=1}^\infty \mu(I_i) < 2\epsilon/M_\delta$, we have

$$\mu(F^{-1}(E)) \le \sum_{i=1}^\infty \mu(F^{-1}(I_i)) = \sum_{i=1}^\infty \{F^{-1}(b_i) - F^{-1}(a_i)\} \le M_\delta \sum_{i=1}^\infty \mu(I_i) < 2\epsilon,$$

therefore,

$$\int_0^1 |h(t) - h_\epsilon(t)|\,\phi(t)\,\mathrm{d}t \le (2M+1)\epsilon.$$

By (3.3) and (3.10), we have for large n and a constant $M_1 > 0$,

$$\sup_{|\mathbf{u}|\le K} \left| \tau_2(S_{nk}^*(\cdot,\mathbf{u}), J_n^*(\cdot,\mathbf{u})) - \int_0^1 \phi(t) S_{nk}^*(t,\mathbf{u}) h_\epsilon(J_n^*(t,\mathbf{u}))\,\mathrm{d}t \right| \le M_1\epsilon$$

in probability. For h_ϵ, there exists a continuous g_ϵ such that g_ϵ is piecewise differentiable with bounded derivative, say $|g'_\epsilon| \le M_\epsilon$, and $|h_\epsilon(t) - g_\epsilon(t)| \le \epsilon$, for all $t \in [0, 1]$. Therefore, by (3.3), we have for a constant $M_2 > 0$,

$$\sup_{|\mathbf{u}|\le K} \left| \tau_2(S_{nk}^*(\cdot,\mathbf{u}), J_n^*(\cdot,\mathbf{u})) - \int_0^1 \phi(t) S_{nk}^*(t,\mathbf{u}) g_\epsilon(J_n^*(t,\mathbf{u}))\,\mathrm{d}t \right| \le M_2\epsilon \tag{3.11}$$

in probability, and

$$\left| \int_0^1 \phi(t) S_{nk}^*(t,\mathbf{u}) h(t)\,\mathrm{d}t - \int_0^1 \phi(t) S_{nk}^*(t,\mathbf{u}) g_\epsilon(t)\,\mathrm{d}t \right| \le M_2\epsilon \tag{3.12}$$

in probability. From the first part of the proof of this theorem, we have (3.9) for g_ϵ. Therefore, (3.9) follows from (3.11) and (3.12). □

References

Bassett, G. Jr. and Koenker, R. (1982). An empirical quantile function for linear models with iid errors. *JASA*. **77,** 407–415.

Bickel, P. J. (1973). On some analogues to linear combinations of order statistics in the linear model. *Ann. Stat.* **1,** No. 4, 597–616.

Carroll, R. J. and Welsh, A. H. (1988). A note on asymmetry and robustness in linear regression. *Amer. Statist.* Vol. 42, No. 4, 285–287.

Fernholz, L. T. (1983). *Von Mises Calculus for Statistical Functionals.* Lecture notes in statistics 19. Springer, New York. [1.3f, 2.1b, 2.3a].

Hájek, J. and Šidák, Z. (1967). *Theory of Rank Tests.* Academia, Prague.

Huber, P. J. (1981). *Robust Statistics.* John Wiley & Sons, New York.

Jurečková, J. and Welsh, A. H. (1990). Asymptotic relations between *L*- and *M*-estimators in the linear model. *Ann. Inst. Statist. Math.* Vol. 42, No. 4, 671–698.

Koenker, R. and Portnoy, S. (1987). *L*-estimation for linear models. *JASA* Vol. 82, No. 399, 851–857.

Ren, J. and Sen, P. K. (1991). On Hadamard differentiability of extended statistical functional. *Multi. Anal.* **39:1,** 30–43.

Ren, J. (1994). On Hadamard differentiability and its application to *R*-estimation in linear models. *Stat. & Dec.*, **12,** No. 1.

Ren, J. and Lu, K. L. (1992). On *L*-estimation in linear models. *Comm. Stat., Theory and Methods*, **23,** No. 1, 137–151.

Royden, H. L. (1968). *Real Analysis.* MacMillan Publishing Co., Inc., New York.

Ruppert, D. and Carroll, R. J. (1980). Trimmed least squares estimation in the linear model. *JASA* **75,** 828–837.

Von Mises, R. (1947). On the asymptotic distribution of differentiable statistical functions. *AMS* **18,** 309–348.

Welsh, A. H. (1987). One-step *L*-estimators for the linear model. *Ann. Stat.* **15,** No. 2, 626–641.