Ann. Inst. Statist. Math.
Vol. 42, No. 4, 671–698 (1990)

ASYMPTOTIC RELATIONS BETWEEN *L*- AND *M*-ESTIMATORS IN THE LINEAR MODEL

JANA JUREČKOVÁ[1] AND A. H. WELSH[2]

[1]*Department of Probability and Statistics, Charles University, Sokolovska 83, 18600 Prague, Czechoslovakia*
[2]*Department of Statistics, The Australian National University, GPO Box 4, Canberra, ACT 2601, Australia*

(Received February 20, 1989; revised September 11, 1989)

Abstract. We obtain Bahadur-type representations for one-step *L*-estimators, *M*- and one-step *M*-estimators in the linear model. The order of the remainder terms in these representations depends on the smoothness of the weight function for *L*-estimators and on the smoothness of the ψ-function for *M*- and one-step *M*-estimators. We use the representations to investigate the asymptotic relations between these estimators. In particular, we show that asymptotically equivalent *L*- and *M*-estimators of the slope parameter exist even when the underlying distribution is asymmetric. It is important to consider the asymmetric case for both practical and robustness reasons: first, there is no compelling argument which precludes asymmetric distributions from arising in practice, and, secondly, even if a symmetric model can be posited, it is important to allow for the possibility of mild (and therefore difficult to detect) departures from the symmetric model.

Key words and phrases: Bahadur representations, multiple regression, robust estimators, uniform asymptotic linearity.

1. Introduction

Consider the usual linear model in which we observe $Y_1,\dots, Y_n$, where

$$Y_j = x_j'\theta_0 + Z_j, \qquad 1 \le j \le n\,, \tag{1.1}$$

with $\{x_j' = (1, x_{j2},\dots, x_{jp})\}$ a sequence of known p-vectors ($p \ge 1$), $\theta_0' = (0, \theta_{02},\dots, \theta_{0p})$ an unknown slope parameter to be estimated and $\{Z_j\}$ a sequence of independent and identically distributed random variables with common distribution function F. Since we will not assume that F is symmetric about some point, the model will not have a unique identifiable intercept. Instead, we define an "intercept" for each estimator of $\theta_0 \in \mathbb{R}^p$

which is a function of the estimator and F. Since we can center each component of x_j about its sample mean, there is no loss of generality in taking $\sum_{j=1}^{n} x_{jk} = 0$, $2 \le k \le p$. We suppress the consequent dependence of $\{x_j\}$ on n for notational simplicity.

While there is an extensive literature on the relationships between classes of location estimators, there is little beyond Jurečková (1977) on the relationships between classes of regression estimators. The regression problem has content beyond that contained in the location problem in that the slope parameters are identifiable when F is asymmetric. Thus the regression problem can and should be treated separately from the location problem. Recently, Welsh (1987*a*, 1987*b*) introduced a general class of one-step L-estimators of the regression parameter θ_0 in the model (1.1). In the discussion to Welsh (1987*a*), Koenker raised important questions about the relationship between these estimators and M-estimators (Relles (1968), Huber (1973)) or one-step M-estimators (Bickel (1975)). In this paper, we investigate the asymptotic relationship between these classes of estimators and clarify some of the issues raised by Koenker's discussion.

Let

$$H(u) = \int_0^u h(t)dt, \qquad 0 < u < 1 ,$$

be a fixed, bounded, signed measure on (0,1) with a weight function h which is the sum of an absolutely continuous function and a step function. Also let $w_1, \ldots, w_m$ be constant weights and $0 < q_1 < \cdots < q_m < 1$ for some $m < \infty$. It is convenient to normalise so that

$$H(1) + \sum_{i=1}^{m} w_i = 1 .$$

Then for any distribution function G define an L-functional

$$T_L(G) = \int_0^1 G^{-1}(u)dH(u) + \sum_{i=1}^{m} w_i G^{-1}(q_i) , \tag{1.2}$$

where $G^{-1}(t) = \inf\{s\colon G(s) \ge t\}$. Let $r_j = Y_j - x_j'\theta_n$, $1 \le j \le n$, denote the residuals from $\theta_n' = (0, \tau_{n2}') \in \mathbb{R}^p$, where $\tau_n' = (\tau_{n1}, \tau_{n2}') \in \mathbb{R}^p$ is an initial estimator, and set

$$G_n(y) = n^{-1} \sum_{j=1}^{n} I(r_j \le y), \qquad y \in \mathbb{R} .$$

Following Welsh (1987*b*), a one-step L-estimator of θ_0 is defined by

(1.3)
$$\lambda_n = \theta_n + T_L(G_n)a_1 - D_n^{-1}\sum_{j=1}^n x_j\left[\int_{-\infty}^{\infty}\{I(r_j \le y) - G_n(y)\}h(G_n(y))dy + \sum_{i=1}^m \{w_i/\phi_n(q_i)\}\{I(r_j \le G_n^{-1}(q_i)) - q_i\}\right],$$

where $a_1' = (1, 0, \ldots, 0) \in \mathbb{R}^p$,

$$D_n = \begin{cases} \sum_{j=1}^n x_j x_j'\left\{h(G_n(r_j)) + \sum_{i=1}^m w_i\right\} & \text{for a type I estimator} \\ X'X = \sum_{j=1}^n x_j x_j' & \text{for a type II estimator} \end{cases}$$

and $\phi_n(q_i) \xrightarrow{P} \phi(q_i)$, $1 \le i \le m$, with $\phi(q)^{-1} = \partial F^{-1}(q)/\partial q$. It makes sense to replace τ_{n1} by 0 after fitting (1.1) with an intercept and to adopt the convention that $G_n(G_n^{-1}(q)) = q$ because then $\lambda_{n1} = T(G_n)$, the usual L-estimator in the location problem ($p = 1$). We discuss possible choices of θ_n and ϕ_n in Section 2. Alternative formulations of λ_n are given in Welsh (1987*b*).

Let $\psi: \mathbb{R} \to \mathbb{R}$ be a fixed real function of the form

$$\psi = \psi_a + \psi_c + \psi_s,$$

where ψ_a is absolutely continuous with absolutely continuous derivative $\psi_a'(z) = d\psi_a(z)/dz$, ψ_c is a continuous, piecewise linear function which is constant in a neighbourhood of $\pm\infty$ and ψ_s is a monotone step function. Then we define an M-estimator μ_n of θ_0 to be a p-vector satisfying

(1.4)
$$n^{-1/2}\sum_{j=1}^n x_j\psi((Y_j - x_j'\mu_n)/\sigma_n) = \begin{cases} O_p(n^{-1/2}) & \text{if } \psi_s = 0 \\ O_p(n^{-1/4}) & \text{otherwise} \end{cases}$$

and

$$n^{1/2}(\mu_n - \theta_0 - T_M(F)a_1) = O_p(1), \quad \text{as} \quad n \to \infty,$$

where σ_n is a translation invariant and scale equivariant scale statistic satisfying $n^{1/2}(\sigma_n - \sigma) = O_p(1)$ as $n \to \infty$, for some positive functional $\sigma = \sigma(F) > 0$, $T_M(F)$ is a real M-functional defined by

(1.5)
$$\int_{-\infty}^{\infty}\psi((z - T_M(F))/\sigma)dF(z) = 0$$

and $a_1' = (1, 0, \ldots, 0) \in \mathbb{R}^p$. If $\psi_s = 0$, we obtain an estimator μ_n satisfying

(1.4) by solving the system of equations

$$\sum_{j=1}^{n} x_j \psi((Y_j - x_j't)/\sigma_n) = 0 . \tag{1.6}$$

If $\psi_s \neq 0$, then (1.6) may have no solution. In this case, if ρ is a real function such that $\rho' = \psi$ is monotone increasing and skew symmetric, we can use the argument from the appendix to Ruppert and Carroll (1980) to show that the minimum of $\sum_{j=1}^{n} \rho((Y_j - x_j't)/\sigma_n)$ satisfies (1.4). Of course, other criteria can be used to find estimators satisfying (1.4) in this case. The requirement that ψ be decomposable into smooth functions and a step function is not unduly restrictive and defines the class of M-estimators which can be related to L-estimators.

A one-step M-estimator is the result of a single iteration from an initial estimator τ_n of a Newton-Raphson procedure for solving (1.4). We define a one-step M-estimator to be

$$\tilde{\mu}_n = \tau_n + \Delta_n^{-1} \sum_{j=1}^{n} x_j \psi((r_j - \tau_{n1})/\sigma_n) ,$$

where r_j, $1 \leq j \leq n$, are the residuals from $\theta_n' = (0, \tau_{n2}') \in \mathbb{R}^p$, σ_n is a translation invariant and scale equivariant scale statistic satisfying $n^{1/2}(\sigma_n - \sigma) = O_p(1)$ as $n \to \infty$, for some positive functional $\sigma = \sigma(F) > 0$ and Δ_n^{-1} is a generalised inverse of

$$\Delta_n = \begin{cases} \sum_{j=1}^{n} x_j x_j' \{\psi_s'((r_j - \tau_{n1})/\delta_n)/\sigma_n + \xi_{1n}\} & \text{for a type I estimator} \\ (\gamma_{1n} + \xi_{1n}) X'X & \text{for a type II estimator ,} \end{cases}$$

where ξ_{1n} estimates ξ_1 defined in (2.4) and

$$\gamma_{1n} = \frac{1}{2} n^{-1/2} \sum_{j=1}^{n} \{\psi((r_j - \tau_{n1} + n^{-1/2})/\sigma_n) - \psi((r_j - \tau_{n1} - n^{-1/2})/\sigma_n)\} .$$

The type I estimator is less appealing to work with because the analysis of Δ_n requires conditions on ψ_s' which are not required for M-estimators or type II one-step M-estimators. Consequently, in the sequel we will restrict attention to type II one-step M-estimation.

In this paper, we investigate the behaviour of the vector differences $\lambda_n - \mu_n$ and $\lambda_n - \tilde{\mu}_n$. We are particularly interested in obtaining conditions under which these vectors converge in probability to zero, in which case we need to relate h and $\{(q_i, w_i)\colon 1 \leq i \leq m\}$ to ψ. In addition, we are interested in the rate of this convergence.

In the location problem ($p = 1$), the relationship between *L*- and *M*-estimators has been studied by Jaeckel (1971), Rivest (1982), van Eeden (1983) and Jurečková (1986) (see Jurečková (1986) for a review of the earlier work). Essentially, these papers differ in the nature of the conditions and the strength of the results. For example, Jaeckel (1971) and van Eeden (1983) treat the scale as known whereas Rivest (1982) and Jurečková (1986) treat the scale as unknown and Rivest (1982) and van Eeden (1983) establish conditions for the asymptotic equivalence of λ_n and μ_n but do not obtain a rate of convergence whereas Jaeckel (1971) and Jurečková (1986) do obtain rates of convergence. The present results for the more general regression problem are most closely related to those of Jurečková (1986). The present conditions are of course slightly different but the nature and type of result is similar to those in Jurečková (1986).

We approach the problem of relating *L*- and *M*-estimators by deriving asymptotic Bahadur-type representations for these estimators and examining the order of the remainder terms in the representations. Of course, these representations are of independent interest. The representation for λ_n is derived by extending the arguments of Welsh (1987*b*) using arguments from Jurečková (1986). Representations for μ_n have been obtained by a number of authors including Huber (1973), Jurečková (1977), Yohai and Maronna (1979) and Jurečková and Sen (1984). These representations are obtained under various conditions, usually without examining the order of the remainder term (see Jurečková and Sen (1984)). Representations for $\tilde{\mu}_n$ have been obtained by Bickel (1975) and Jurečková and Portnoy (1987). We will derive new representations (including the order of the remainder term) for *M*-estimators and one-step *M*-estimators which are particularly useful for relating these estimators to *L*-estimators. The main technical tool is the recent result on multiparameter stochastic processes established in Jurečková and Sen (1989).

Finally, the regression quantiles of Koenker and Bassett (1978) can be used to construct alternative *L*-estimators of θ_0. General alternatives to λ_n have been considered by Koenker and Bassett (1978) and Koenker and Portnoy (1987) while trimmed and Winsorised means have been considered by Ruppert and Carroll (1980) and Jurečková (1983*a*, 1983*b*, 1983*c*, 1984). We will not consider these estimators in this paper beyond noting that the representations for these estimators can be applied to obtain results which are analogous to those we present.

We introduce the notation and conditions we require in Section 2 before deriving asymptotic representations for *L*-, *M*- and one-step *M*-estimators in Sections 3–5 respectively. Finally, in Section 6 we examine the relations between these estimators.

2. Notation and conditions

We assume throughout that the basic linear model (1.1) holds and that the design sequence $\{x_j\}$ satisfies the following basic conditions:

CONDITION. (i) $x_{j1} = 1$ and $\sum_{j=1}^{n} x_{jk} = 0,\ 2 \le k \le p$;
(ii) $n^{-1/4} \max_{1 \le j \le n} |x_j| = O(1)$ and
(iii) $n^{-1} \sum_{j=1}^{n} |x_j|^4 = O(1)$ and $\lim_{n\to\infty} n^{-1} \sum_{j=1}^{n} x_j x_j' = \Gamma$
for some non-singular matrix Γ.

These are standard design conditions; Condition (ii) and the first part of Condition (iii) are needed to obtain the order of the remainder term in the asymptotic representations.

To obtain an asymptotic representation for λ_n, we require conditions on the initial estimator θ_n and on F and h. It is convenient to treat L-functionals and L-estimators as the sum of a smooth term depending on $h = h_a + h_s$, where h_a is absolutely continuous and h_s is a step function, and a quantile term depending on $\{(w_i, q_i)\colon 1 \le i \le m\}$. If the support of h is a proper subset of (0,1), we describe h (and λ_n) as trimmed, otherwise we describe h (and λ_n) as untrimmed. The conditions on F depend on whether h is trimmed or not.

We will impose the following conditions:

CONDITION. (L1) There is an estimator $\tau_n \in \mathbb{R}^p$, satisfying $n^{1/2}(\tau_n - \theta_0 - \tau_0 a_1) = O_p(1)$ for some fixed $\tau_0 \in \mathbb{R}$, where $a_1' = (1, 0, \ldots, 0) \in \mathbb{R}^p$;
(L2) $f(y) = dF(y)/dy$ is uniformly continuous, positive and bounded;
(L3) h_a is bounded and satisfies $\int_{-\infty}^{\infty} \left\{ \sup_{|u| \le \delta} |h_a'(z+u)| \right\}^{\kappa} dF(z) < \infty$, for all $0 < \delta \le \delta_0$. If $h_a(u) \ne 0$ for $u < \alpha$ or $u > \beta$, for some $0 < \alpha < \beta < 1$, then

$$\int_{-\infty}^{\infty} [F(y+\varepsilon)\{1 - F(y-\varepsilon)\}]^{1/2} dy < \infty, \qquad \varepsilon > 0\,,$$

holds;
(L4) $h_s(u) = 0$ for $u < \alpha$ or $u > \beta$, $0 < \alpha < \beta < 1$, and h_s is a bounded step function with a finite number of steps at $\alpha < s_1 < \cdots < s_c < \beta$; and
(L5) (i) there is a $\phi_n(q_i)$ such that $n^{1/4}\{\phi_n(q_i) - \phi(q_i)\} = O_p(1)$, $1 \le i \le m$ and (ii) $f'(y) = d^2F(y)/dy^2$ is bounded in a neighbourhood of $F^{-1}(q_i)$, $1 \le i \le m$.

Finally, the asymptotic representations for λ_n involve the function

$$\psi_L(z) = -\int_{-\infty}^{\infty} \{I(\sigma z \le y) - F(y)\}h(F(y))dy - \sum_{i=1}^{m} \{w_i/\phi(q_i)\}\{I(\sigma z \le F^{-1}(q_i)) - q_i\} . \tag{2.1}$$

In Condition (L1) we permit $\tau_n' = (\tau_{n1}, \tau_{n2}') \in \mathbb{R}^p$ to estimate an arbitrary intercept τ_0 which may depend both on τ_n and F. Although for one-step L-estimators we discard the intercept component τ_{n1} and use $\theta_n' = (0, \tau_{n2}') \in \mathbb{R}^p$ as the initial estimator, the form of Condition (L1) emphasises that F can be asymmetric and that using θ_n does not entail that the regression surface in (1.1) is constrained to pass through the origin. Under mild conditions, Condition (L1) is satisfied by a rich class of estimators including M-estimators (see Section 4). Condition (L2) is stronger than the smoothness conditions required on F in the location problem but is required for the weak convergence of the empirical process based on regression residuals. The second part of Condition (L3) is a tail condition which is close to requiring $EZ^2 < \infty$. Under Condition (L5)(ii), it is straightforward to show that the simple histogram estimator discussed by Welsh (1987*b*) satisfies Condition (L5)(i). An alternative kernel estimator which also satisfies Condition (L5)(i) under mild conditions is considered by Welsh (1987*c*).

As noted in the introduction, we will restrict attention to M-estimators and one-step M-estimators based on bounded ψ-functions of the form

$$\psi = \psi_a + \psi_c + \psi_s ,$$

where ψ_a is absolutely continuous with absolutely continuous derivative $\psi_a'(z) = d\psi_a(z)/dz$, ψ_c is a continuous, piecewise linear function which is constant in a neighbourhood of $\pm\infty$ and ψ_s is a monotone step function. Let $\tau \in \mathbb{R}$ be a fixed number; we will take $\tau = T_M(F)$ in Section 4 and $\tau = \tau_0$ in Section 5. Then we impose the following conditions:

CONDITION. (M1) There is a location invariant and scale equivariant statistic $\sigma_n > 0$ satisfying $n^{1/2}(\sigma_n - \sigma) = O_p(1)$;

(M2) $\int_{-\infty}^{\infty} \psi((z-\tau)/\sigma)dF(z) = 0$ and $\int_{-\infty}^{\infty} \psi((z-\tau)/\sigma)^2 dF(z) < \infty$;

(M3) For some $\kappa > 1$ and $\delta_0 > 0$,

$$\int_{-\infty}^{\infty} \left\{ |z| \sup_{|v|\le\delta} \sup_{|u|\le\delta} |\psi_a''(e^{-v}(z-\tau+u)/\sigma)| \right\}^{\kappa} dF(z) < \infty$$

and

$$\int_{-\infty}^{\infty} \left\{ |z|^2 \sup_{|u|\le\delta} |\psi_a''((z-\tau+u)/\sigma)| \right\}^{\kappa} dF(z) < \infty ,$$

for all $0 < \delta \le \delta_0$;

(M4) ψ_c is a continuous, piecewise linear function with knots at $-\infty = S_0 < S_1 < \cdots < S_c < S_{c+1} = \infty$ which is constant in a neighbourhood of $\pm\infty$ (so that ψ_c' is a step function with jump discontinuities at $-\infty = S_0 < S_1 < \cdots < S_c < S_{c+1} = \infty$ which equals zero in a neighbourhood of $\pm\infty$) and $f(y) = dF(y)/dy$ is bounded in neighbourhoods of the discontinuities of ψ_c';

(M5) $\psi_s(z) = \sum_{l=1}^{k} d_l I(Q_l < z \le Q_{l+1})$, where $-\infty = Q_0 < Q_1 < \cdots < Q_{k+1} = \infty$ and $d_0 < d_1 < \cdots < d_k$ and $f(y) = dF(y)/dy$ and $f'(y) = d^2F(y)/dy^2$ are bounded in a neighbourhood of σQ_l, $0 \le l \le k$.

The asymptotic representations for μ_n involve

$$\gamma_1 = \sigma^{-1} \int_{-\infty}^{\infty} \{\psi_a'((z-\tau)/\sigma) + \psi_c'((z-\tau)/\sigma)\} dF(z)\,, \tag{2.2}$$

$$\gamma_2 = \sigma^{-1} \int_{-\infty}^{\infty} (z-\tau)\{\psi_a'((z-\tau)/\sigma) + \psi_c'((z-\tau)/\sigma)\} dF(z)\,, \tag{2.3}$$

$$\xi_1 = \sum_{l=1}^{k} (d_l - d_{l-1}) f(\sigma Q_l + \tau) \tag{2.4}$$

and

$$\xi_2 = \sum_{l=1}^{m} (d_l - d_{l-1})(\sigma Q_l + \tau) f(\sigma Q_l + \tau)\,. \tag{2.5}$$

We will assume that the integrals γ_1 and γ_2 are finite and that $\gamma_1 + \xi_1 > 0$.

Condition (M1) is the usual condition imposed when the scale is unknown. Two possible choices of σ_n are investigated by Welsh (1986). Condition (M3) is essentially a moment condition which holds if ψ_a'' is bounded and either

(a) $\psi_a'(z) = 0$ for $z < a$ or $z > b$, $-\infty < a < b < \infty$

or

(b) $\int |z|^{2+\varepsilon} dF(z) < \infty$, for some $\varepsilon > 0$

holds. We will not prove that (a) or (b) imply Condition (M3) because the proofs are straightforward. Conditions (M4) and (M5) show explicitly the tradeoff between the smoothness of ψ and the smoothness of F. The above conditions are chosen for the ease with which they can be reconciled with Conditions (L3) and (L4).

If σ is known or can be factored out of the estimating equations (1.4), we can omit Condition (M1) and replace (M3) by

(M3′) For some $\kappa > 1$ and $\delta_0 > 0$,

$$\int_{-\infty}^{\infty} \left\{ \sup_{|u|\le\delta} |\psi_a''((z-\tau+u)/\sigma)| \right\}^{\kappa} dF(z) < \infty ,$$

for all $0 < \delta \le \delta_0$.

For one-step *M*-estimators, we will impose Conditions (L1), (M1)–(M5) with $\tau = \tau_0$ and the additional condition

CONDITION. (S1) There is an estimator ξ_{1n} satisfying $n^{1/4}(\xi_{1n} - \xi_1) = O_p(1)$.

This condition is related to Condition (L5)(i) since it involves estimating the density at a finite number of fixed points.

3. Representations for one-step *L*-estimators

The representations will be derived from the asymptotic linearity results of Jurečková and Sen (1989). The following result is extracted from Theorems 2.1, 3.1 and 4.1 of Jurečková and Sen (1989).

LEMMA 3.1. (Jurečková and Sen (1989)) *Let ψ be a bounded function of the form*

$$\psi = \psi_a + \psi_c + \psi_s ,$$

where ψ_a is absolutely continuous with absolutely continuous derivative $\psi_a'(z) = d\psi_a(z)/dz$, ψ_c is a continuous, piecewise linear function which is constant in a neighbourhood of $\pm\infty$ and ψ_s is a monotone step function. Suppose that Conditions (i)–(iii) *hold and* $a_1' = (1, 0, \ldots, 0) \in \mathbb{R}^p$. *Then if Conditions* (M3′) *and* (M4) *hold with $\tau = 0$ and $\psi_s = 0$,*

$$\sup_{|t|\le B} \left| n^{-1} \sum_{j=1}^{n} x_j\{\psi((Z_j - n^{-1/2}x_j't)/\sigma) - \psi(Z_j/\sigma)\} + n^{-3/2}\gamma_1 X'Xt \right| = O_p(n^{-1})$$

for each fixed $B < \infty$, where γ_1 is defined in (2.2) *with $\tau = 0$. If Condition* (M5) *holds with $\tau = 0$ and $\psi_a = \psi_c = 0$,*

$$\sup_{|t|\le B} \left| n^{-1} \sum_{j=1}^{n} x_j\{\psi((Z_j - n^{-1/2}x_j't)/\sigma) - \psi(Z_j/\sigma)\} + n^{-3/2}\xi_1 X'Xt \right| = O_p(n^{-3/4}) ,$$

for each fixed $B<\infty$, *where* ξ_1 *is defined in* (2.4) *with* $\tau=0$.

It is convenient and instructive to consider separately the cases $m=0$ and $h=0$ before considering the general class of one-step L-estimators. For the case $m=0$, we will require the following preliminary lemma which is an extension of Lemma 2 of Welsh (1987*b*);

LEMMA 3.2. *Suppose that Conditions* (i)–(iii) *and* (L2)–(L4) *hold. Then if* $m=0$,

$$\sup_{|t|\le B}\left| n^{-1}\sum_{j=1}^{n} x_j\int\{F(y+n^{-1/2}x_j't)-F(y)-n^{1/2}x_j'tf(y)\}\cdot h(F(y))dy\right| = O(n^{-1}),$$

and

$$\sup_{|t|\le B}\left| n^{-1}\sum_{j=1}^{n} x_j\int\{I(Z_j\le y+n^{-1/2}x_j't)-F(y+n^{-1/2}x_j't) - I(Z_j\le y)+F(y)\}h(F(y))dy\right| = O_p(n^{-1})$$

for any fixed $B>0$.

PROOF. The proof of the lemma depends on whether h is trimmed or not; we give the proof in the trimmed case and note that the proof in the untrimmed case is slightly simpler. Without loss of generality, suppose that h has a single discontinuity at s_0. Let $S_0=F^{-1}(s_0)$. Then fix $t\in\mathbb{R}^p$ and let $\eta_j=n^{-1/2}x_j't$. If $\eta_j>0$,

$$\begin{aligned}
&\left|\int_{-\infty}^{\infty}\{F(y+\eta_j)-F(y)-\eta_jf(y)\}h(F(y))dy\right|\\
&\le\left|\int_{-\infty}^{\infty}\int_0^{\eta_j}\{f(y+v)-f(y)\}h(F(y))dvdy\right|\\
&=\left|\int_{-\infty}^{\infty}\int_0^{\eta_j}\{h(F(y-v))-h(F(y))\}dvdF(y)\right|\\
&\le\int_{-\infty}^{S_0-n^{-1/2}}\int_0^{\eta_j}|h(F(y-v))-h(F(y))|dvdF(y)\\
&\quad+\int_{S_0-n^{-1/2}}^{S_0+n^{-1/2}}\int_0^{\eta_j}|h(F(y-v))-h(F(y))|dvdF(y)\\
&\quad+\int_{S_0+n^{-1/2}}^{\infty}\int_0^{\eta_j}|h(F(y-v))-h(F(y))|dvdF(y)\,.
\end{aligned}$$

Now for $K_1, K_2 < \infty$,

$$\int_{-\infty}^{S_0 - n^{-1/2}} \int_0^{\eta_j} |h(F(y-v)) - h(F(y))| dv dF(y)$$
$$= \int_{-\infty}^{S_0 - n^{-1/2}} \int_0^{\eta_j} |h_a(F(y-v)) - h_a(F(y))| dv dF(y)$$
$$= \int_{-\infty}^{S_0 - n^{-1/2}} \int_0^{\eta_j} \Big| I\{F(y-v) > F(y)\} \int_0^{F(y-v)-F(y)} h_a'(F(y)+u) du$$
$$+ I\{F(y-v) \leq F(y)\}$$
$$\cdot \int_0^{F(y)-F(y-v)} h_a'(F(y)+u) du \Big| dv dF(y)$$
$$\leq K_1 \int_{-\infty}^{S_0 - n^{-1/2}} \int_0^{\eta_j} |F(y-v) - F(y)| dv dF(y)$$
$$\leq K_2 \int_{-\infty}^{S_0 - n^{-1/2}} \int_0^{\eta_j} v dv dF(y)$$
$$\leq n^{-1} 2^{-1} K_2 |x_j|^2 |t|^2 .$$

Similarly,

$$\int_{S_0 + n^{-1/2}}^{\infty} \int_0^{\eta_j} |h(F(y-v)) - h(F(y))| dv dF(y) \leq n^{-1} 2^{-1} K_2 |x_j|^2 |t|^2 .$$

Also, since h is bounded, for $K < \infty$,

$$\int_{S_0 - n^{-1/2}}^{S_0 + n^{-1/2}} \int_0^{\eta_j} |h(F(y-v)) - h(F(y))| dv dF(y) \leq n^{-1} K |x_j| |t| .$$

Combining these bounds, we have that for $K < \infty$,

$$\left| \int_{-\infty}^{\infty} \{F(y+\eta_j) - F(y) - \eta_j f(y)\} h(F(y)) dy \right| \leq K n^{-1} (|x_j|^2 |t|^2 + |x_j| |t|) .$$

We obtain the same bound if $\eta_j < 0$ so

$$\sup_{|t| \leq B} \left| n^{-1} \sum_{j=1}^{n} x_j \int \{F(y + n^{-1/2} x_j' t) - F(y) - n^{-1/2} x_j' t f(y)\} h(F(y)) dy \right|$$
$$\leq n^{-1} K n^{-1} \sum_{j=1}^{n} (|x_j|^{\kappa+1} B^{\kappa+1} + |x_j| B)$$
$$= O(n^{-1})$$

by Condition (iii).

Now let ψ_L be defined by (2.1) with $m = 0$ and $\sigma = 1$. Then ψ_L is an absolutely continuous function which satisfies Conditions (M3′) and (M4) with $\tau = 0$. Also,

$$\int_{-\infty}^{\infty} \psi_L'(z)dF(z) = \int_{-\infty}^{\infty} h(F(z))dF(z) = H(1)$$

so by Lemma 3.1 and the first part of the present Lemma,

$$\begin{aligned}
&\sup_{|t|\leq B} \Big| n^{-1} \sum_{j=1}^{n} x_j \int \{I(Z_j \leq y + n^{-1/2}x_j't) - F(y + n^{-1/2}x_j't) \\
&\qquad\qquad - I(Z_j \leq y) + F(y)\}h(F(y))dy \Big| \\
&\leq \sup_{|t|\leq B} \Big| n^{-1} \sum_{j=1}^{n} x_j \int \{I(Z_j \leq y + n^{-1/2}x_j't) - I(Z_j \leq y)\}h(F(y))dy \\
&\qquad\qquad + n^{-3/2}H(1)X'Xt \Big| \\
&\quad + \sup_{|t|\leq B} \Big| n^{-1} \sum_{j=1}^{n} x_j \int \{F(y + n^{-1/2}x_j't) - F(y)\}h(F(y))dy \\
&\qquad\qquad - n^{-3/2}H(1)X'Xt \Big| \\
&\leq \sup_{|t|\leq B} \Big| n^{-1} \sum_{j=1}^{n} x_j \{\psi_L(Z_j - n^{-1/2}x_j't) - \psi_L(Z_j)\} + n^{-3/2}H(1)X'Xt \Big| \\
&\quad + \sup_{|t|\leq B} \Big| n^{-1} \sum_{j=1}^{n} x_j \int \{F(y + n^{-1/2}x_j't) - F(y) \\
&\qquad\qquad - n^{-1/2}x_j'tf(y)\}h(F(y))dy \Big| \\
&= O_p(n^{-1}) . \qquad \square
\end{aligned}$$

The following theorem generalises Theorem 1 of Welsh (1987*b*) and Theorem 2.1 of Jurečková (1986). The proof extends that of Theorem 1 of Welsh (1987*b*) using the technique of Jurečková (1986).

THEOREM 3.1. *Suppose that Conditions* (i)–(iii) *and* (L1)–(L4) *hold. Let* $a_1 = (1, 0, \ldots, 0)' \in \mathbb{R}^p$. *Then if* $m = 0$ *in* (1.3),

$$\lambda_n - \theta_0 - T_L(F)a_1 = (X'X)^{-1} \sum_{j=1}^{n} x_j\psi_L(Z_j/\sigma) + O_p(n^{-1}) ,$$

where $T_L(F)$ is defined in (1.2) *and ψ_L is defined in* (2.1).

PROOF. We prove the result for type II estimators; the proof for type I estimators is similar. Write

$$n^{-1}X'X(\lambda_n - \theta_0 - T_L(F)a_1) - n^{-1}\sum_{j=1}^{n} x_j\psi_L(Z_j/\sigma)$$

$$= n^{-1}\sum_{j=1}^{n} x_j\Big[x_j'(\theta_n - \theta_0) + T_L(G_n) - T_L(F) - \int\{I(r_j \le y) - G_n(y)\}h(G_n(y))dy + \int\{I(e_j \le y) - F(y)\}h(F(y))dy\Big].$$

Let d_j be any fixed component of x_j, $1 \le j \le n$, and, writing $d_j = d_j^+ - d_j^-$ if necessary, take $d_j \ge 0$, $1 \le j \le n$, without loss of generality. For $\bar{d} = n^{-1}\sum_{j=1}^{n} d_j > 0$, put

$$Q_n(y) = (n\bar{d})^{-1}\sum_{j=1}^{n} d_jI(r_j \le y), \qquad P_n(y) = (n\bar{d})^{-1}\sum_{j=1}^{n} d_jI(Z_j \le y)$$

and

$$\bar{P}_n(y) = (n\bar{d})^{-1}\sum_{j=1}^{n} d_jF(y + x_j'(\theta_n - \theta_0)) .$$

Arguing as in Welsh (1987*b*), we see that the result will hold if we can show that $R_i = O_p(n^{-1})$, $i = 1,\dots,4$, where

$$R_1 = \int_{-\infty}^{\infty} W_{G_n,F}(y)\{G_n(y) - F(y)\}dy ,$$

with

$$W_{G,F}(y) = \begin{cases} \{H(F(y)) - H(G(y))\}/\{G(y) - F(y)\} - h(F(y)) & \text{if } G(y) \ne F(y) \\ 0 & \text{otherwise} \end{cases}$$

$$R_2 = \int_{-\infty}^{\infty} \{Q_n(y) - \bar{P}_n(y) - P_n(y) + F(y)\}h(G_n(y))dy ,$$

$$R_3 = \int_{-\infty}^{\infty} \{\bar{P}_n(y) + P_n(y) - 2F(y)\}\{h(G_n(y)) - h(F(y))\}dy$$

and

$$R_4 = n^{-1}\sum_{j=1}^{m} x_j\int_{-\infty}^{\infty}\{F(y + x_j'(\theta_n - \theta_0)) - F(y) - x_j'(\theta_n - \theta_0)f(y)\}h(F(y))dy\,.$$

Suppose first that Conditions (L3) and (L4) hold so h is trimmed. Then there exists $-\infty < a < b < \infty$ such that for $\sup_y |G(y) - F(y)| < \min(\alpha, 1-\beta)$, we have $W_{G,F}(y) = h(G(y)) = h(F(y)) = 0$ for $y < a$ or $y > b$. That is, for n large enough, the range of integration of the integrals in R_1–R_4 can be restricted to $[a, b]$. Without loss of generality, suppose that h has a single jump discontinuity at $\alpha < s_0 < \beta$. Let $S_0 = F^{-1}(s_0)$ so $a < S_0 < b$.

For n sufficiently large,

$$\begin{aligned}|R_1| &\le \int_a^b |W_{G_n,F}(y)\{G_n(y) - F(y)\}|dy\\ &= \int_a^{S_0-n^{-1/2}} |W_{G_n,F}(y)\{G_n(y) - F(y)\}|dy\\ &\quad + \int_{S_0-n^{-1/2}}^{S_0+n^{-1/2}} |W_{G_n,F}(y)\{G_n(y) - F(y)\}|dy\\ &\quad + \int_{S_0+n^{-1/2}}^{b} |W_{G_n,F}(y)\{G_n(y) - F(y)\}|dy\,.\end{aligned}$$

Now with K a generic positive constant,

$$\begin{aligned}&\int_a^{S_0-n^{-1/2}} |W_{G_n,F}(y)\{G_n(y) - F(y)\}|dy\\ &\le \int_a^{S_0-n^{-1/2}}\Big[I(G_n(y) > F(y))\int_0^{G_n(y)-F(y)} |h(F(y)+u) - h(F(y))|du\\ &\qquad + I(G_n(y) < F(y))\int_{G_n(y)-F(y)}^{0} |h(F(y)+u) - h(F(y))|du\Big]dy\\ &\le \int_a^{S_0-n^{-1/2}}\Big[I(G_n(y) > F(y))\int_0^{G_n(y)-F(y)} \Big|\int_0^u h_a'(F(y)+v)dv\Big| du\\ &\qquad + I(G_n(y) < F(y))\int_{G_n(y)-F(y)}^{0} \Big|\int_{-u}^{0} h(F(y)+v)dv\Big| du\Big]dy\\ &\le K\int_a^{S_0-n^{-1/2}}\Big[I(G_n(y) > F(y))\int_0^{G_n(y)-F(y)} udu\\ &\qquad + I(G_n(y) < F(y))\int_{G_n(y)-F(y)}^{0} (-u)du\Big]dy\end{aligned}$$

$$\leq K \sup_y |G_n(y) - F(y)|^2 |S_0 - n^{-1/2} - a|$$

$$= O_p(n^{-1}),$$

by a result of Koul (1969) and Bickel (1973) (see (3.4) of Welsh (1987*b*)). Similarly,

$$\int_{S_0+n^{-1/2}}^{b} |W_{G_n,F}(y)\{G_n(y) - F(y)\}|dy = O_p(n^{-1}).$$

Also, $W_{G,F}(\cdot)$ is bounded so

$$\int_{S_0-n^{-1/2}}^{S_0+n^{-1/2}} |W_{G_n,F}(y)\{G_n(y) - F(y)\}|dy$$

$$\leq Kn^{-1/2} \sup_y |G_n(y) - F(y)| = O_p(n^{-1}),$$

so

$$|R_1| = O_p(n^{-1}).$$

Similarly,

$$|R_3| \leq O_p(n^{-1/2})\int_a^b |h(G_n(y)) - h(F(y))|dy$$

$$\leq O_p(n^{-1/2})\int_a^{S_0-n^{-1/2}} |G_n(y) - F(y)|dy + O_p(n^{-1})$$

$$+ O_p(n^{-1/2})\int_{S_0+n^{-1/2}}^{b} |G_n(y) - F(y)|dy$$

$$= O_p(n^{-1}).$$

Finally, $R_2 = O_p(n^{-1})$ and $R_4 = O_p(n^{-1})$ by Lemma 3.1.

Now suppose that the second part of Condition (L3) holds. Then

$$|R_1| \leq \sup_y |W_{G_n,F}(y)| \int |G_n(y) - F(y)|dy$$

$$\leq \sup_y |G_n(y) - F(y)| O_p(n^{-1/2})$$

$$= O_p(n^{-1})$$

by Lemma 1 of Welsh (1987*b*). Similarly, R_3 is $O_p(n^{-1})$. As before, R_2 and R_4 are $O_p(n^{-1})$ by Lemma 3.2 and Theorem 3.1 is obtained. □

The quantile component of the one-step *L*-estimator may now be treated in a straightforward manner. We have the following result.

THEOREM 3.2. *Suppose that Conditions* (i)–(iii), (L1), (L2) *and* (L5) *hold. Let* $a_1 = (1,0,\dots,0)' \in \mathbb{R}^p$. *Then if* $h = 0$ *in* (1.3),

$$\lambda_n - \theta_0 - T_L(F)a_1 = (X'X)^{-1}\sum_{j=1}^{n} x_j\psi_L(z_j/\sigma) + O_p(n^{-3/4}),$$

where $T_L(F)$ *is defined in* (1.2) *and* ψ_L *is defined in* (2.1).

PROOF. The result will follow if we can show that for a fixed q, $0 < q < 1$,

$$\begin{aligned}(3.1)\quad n^{-1}\sum_{j=1}^{n} x_j[&G_n^{-1}(q) - F^{-1}(q) + x_j'(\theta_n - \theta_0)\\ &- \{I(r_j \le G_n^{-1}(q)) - q\}/\phi_n(q) - \{I(Z_j \le F^{-1}(q)) - q\}/\phi(q)]\\ = &-n^{-1}\sum_{j=1}^{n} x_j[I(r_j \le G_n^{-1}(q)) - I(Z_j \le F^{-1}(q))\\ &\qquad - \phi(q)\{G_n^{-1}(q) - F^{-1}(q) + x_j'(\theta_n - \theta_0)\}]/\phi(q)\\ &+ \{\phi(q)^{-1} - \phi_n(q)^{-1}\}n^{-1}\sum_{j=1}^{n} x_j\{I(r_j \le G_n^{-1}(q)) - q\}\\ = &\ O_p(n^{-3/4}).\end{aligned}$$

Now, by Theorem 1 of Welsh (1986) and a result of Koul (1969) and Bickel (1973), we have that

$$n^{-1}\sum_{j=1}^{n} x_j\{I(r_j \le G_n^{-1}(q)) - q\} = O_p(n^{-1/2})$$

so the second term in (3.1) is $O_p(n^{-3/4})$. To complete the proof of Theorem 3.2, we need to show that the first term in (3.1) is $O_p(n^{-3/4})$ and this will follow if we can show that for any fixed $y \in \mathbb{R}$,

$$\sup_{|t| \le B}\left| n^{-1}\sum_{j=1}^{n} x_j\{I(Z_j \le y + n^{-1/2}x_j't) - I(Z_j \le y) - n^{-1/2}x_j'tf(y)\}\right| = O_p(n^{-3/4}),$$

for any fixed $B > 0$. For fixed $y \in \mathbb{R}$, let $\psi(z) = I(z \le y)$. Then Condition (M5) holds and $\xi_1 = f(y)$ so the result follows from Lemma 3.1. □

Combining Theorems 3.1 and 3.2, we immediately obtain the following result for the general class of one-step L-estimators.

THEOREM 3.3. *Suppose that Conditions* (i)–(iii) *and* (L1)–(L5) *hold. Let* $a_1 = (1, 0, \ldots, 0)' \in \mathbb{R}^p$. *Then*

$$\lambda_n - \theta_0 - T_L(F)a_1 = (X'X)^{-1} \sum_{j=1}^{n} x_j \psi_L(z_j/\sigma) + O_p(n^{-s})$$

with

$$s = \begin{cases} 1 & \text{if } m = 0 \\ 3/4 & \text{else}, \end{cases}$$

where $T_L(F)$ *is defined in* (1.2) *and* ψ_L *is defined in* (2.1).

4. Representations for *M*-estimators

It is again convenient to consider the cases in which $\psi_s = 0$ and $\psi_a = \psi_c = 0$ before considering the general case $\psi = \psi_a + \psi_c + \psi_s$. The representations will be derived from the asymptotic linearity results of Jurečková and Sen (1989). The following result is extracted from Theorems 2.2, 3.2 and 4.2 of Jurečková and Sen (1989).

LEMMA 4.1. (Jurečková and Sen (1989)) *Suppose that Conditions* (i)–(iii) *hold*, $T_M(F)$ *is defined in* (1.5) *and* $a_1' = (1, 0, \ldots, 0) \in \mathbb{R}^p$. *Then if Conditions* (M2)–(M4) *hold with* $\tau = T_M(F)$ *and* $\psi_s = 0$,

$$\sup_{|t| \le B_1} \sup_{|u| \le B_2} \Bigg| n^{-1} \sum_{j=1}^{n} x_j \{\psi(e^{-n^{-1/2}u}(Z_j - T_M(F) - n^{-1/2}x_j't)/\sigma) - \psi((Z_j - T_M(F))/\sigma)\} + n^{-1/2}(n^{-1}\gamma_1 X'Xt + \sigma\gamma_2 u a_1) \Bigg| = O_p(n^{-1})$$

for each fixed B_1, $B_2 < \infty$, *where* γ_1 *and* γ_2 *are defined in* (2.2) *and* (2.3), *respectively, with* $\tau = T_M(F)$. *If Condition* (M5) *holds with* $\tau = T_M(F)$ *and* $\psi_a = \psi_c = 0$,

$$\sup_{|t| \le B_1} \sup_{|u| \le B_2} \Bigg| n^{-1} \sum_{j=1}^{n} x_j \{\psi(e^{-n^{-1/2}u}(Z_j - T_M(F) - n^{-1/2}x_j't)/\sigma) - \psi((Z_j - T_M(F))/\sigma)\} + n^{-1/2}(n^{-1}\xi_1 X'Xt + \sigma\xi_2 u a_1) \Bigg| = O_p(n^{-3/4}),$$

for each fixed B_1, $B_2 < \infty$, *where* ξ_1 *and* ξ_2 *are defined in* (2.4) *and* (2.5), *respectively, with* $\tau = T_M(F)$.

We are able to prove the following theorem for *M*-estimators.

THEOREM 4.1. *Suppose that Conditions* (i)-(iii) *and* (M1)-(M4) *hold with* $\tau = T_M(F)$ *defined in* (1.5). *Let* $a_1' = (1,0,\ldots,0) \in \mathbb{R}^p$. *Then if* $\psi_s = 0$, *there exists a solution* μ_n *of* (1.6) *satisfying*

$$n^{1/2}|\mu_n - \theta_0 - T_M(F)a_1| = O_p(1)\,. \tag{4.1}$$

Moreover, for any solution of (1.6) *satisfying* (4.1),

$$\begin{aligned}\mu_n - \theta_0 - T_M(F)a_1 = {} & \gamma_1^{-1}(X'X)^{-1}\sum_{j=1}^{m} x_j\psi((Z_j - T_M(F))/\sigma) \\ & - \gamma_1^{-1}\gamma_2 a_1(\sigma_n - \sigma) + O_p(n^{-1})\,,\end{aligned} \tag{4.2}$$

where γ_1 *and* γ_2 *are defined in* (2.2) *and* (2.3), *respectively, with* $\tau = T_M(F)$.

PROOF. Let $\mathscr{B} = \{t \in \mathbb{R}^p\colon |t| = B\}$ for some fixed $B > 0$ and define

$$E_n(t,\sigma) = n^{-1/2}\sum_{j=1}^{n} x_j\psi((Z_j - T_M(F) - n^{-1/2}x_j't)/\sigma), \quad t \in \mathbb{R}^p, \quad \sigma > 0\,.$$

Notice that

$$E_n(n^{1/2}(\mu_n - \theta_0 - T_M(F)a_1), \sigma_n) = n^{-1/2}\sum_{j=1}^{n} x_j\psi((Y_j - x_j'\mu_n)/\sigma_n) \tag{4.3}$$

so that $n^{1/2}(\mu_n - \theta_0 - T_M(F)a_1)$ is a solution of $E(t,\sigma_n) = 0$ if and only if μ_n is a solution of (1.4). The result (4.1) will follow from result 6.3.4 of Ortega and Rheinboldt ((1973), p.163) if we can show that $t'E_n(t,\sigma_n) < 0$ for $t \in \mathscr{B}$ in probability.

Now for $K > 0$,

$$\begin{aligned} & P\{t'E_n(t,\sigma_n) < 0 \text{ for all } t \in \mathscr{B}\} \\ & \quad \geq P\{t'E_n(t,\sigma_n) < -K/2 \text{ for all } t \in \mathscr{B}\} \\ & \quad \geq P\{t'E_n(0,\sigma) - \gamma_1 t'\Gamma_n t - \gamma_2 n^{1/2}(\sigma_n - \sigma)t'a_1 < -K \text{ for all } t \in \mathscr{B}\} \\ & \qquad - P\{t'E_n(t,\sigma_n) - t'E_n(0,\sigma) + \gamma_1 t'\Gamma_n t + \gamma_2 n^{1/2}(\sigma_n - \sigma)t'a_1 > K/2 \\ & \qquad\qquad \text{for all } t \in \mathscr{B}\}\,,\end{aligned}$$

where $\Gamma_n = n^{-1}X'X$. Let $\varepsilon > 0$ be given. Then if $\lambda_{\max}(\Gamma_n)$ is the largest eigenvalue of Γ_n,

$$\begin{aligned} & P\{t'E_n(0,\sigma) - \gamma_1 t'\Gamma_n t - \gamma_2 n^{1/2}(\sigma_n - \sigma)t'a_1 < -K \text{ for all } t \in \mathscr{B}\} \\ & \quad \geq P\{t'E_n(0,\sigma) - \gamma_2 n^{1/2}(\sigma_n - \sigma)t'a_1 < \gamma_1 B^2\lambda_{\max}(\Gamma_n) - K \text{ for all } t \in \mathscr{B}\}\end{aligned}$$

$$\geq P\{B|E_n(0,\sigma) - \gamma_2 n^{1/2}(\sigma_n - \sigma)t'a_1| < \gamma_1 B^2 \lambda_{\max}(\Gamma_n) - K\}$$
$$\geq P\{|E_n(0,\sigma)| < (\gamma_1 B\lambda_{\max}(\Gamma_n) - K/B)/2\}$$
$$- P\{|\gamma_2 n^{1/2}(\sigma_n - \sigma)| > (\gamma_1 B\lambda_{\max}(\Gamma_n) - K/B)/2\}$$
$$\geq 1 - 4E|E_n(0,\sigma)|^2(\gamma_1 B\lambda_{\max}(\Gamma_n) - K/B)^2$$
$$- P\{|\gamma_2 n^{1/2}(\sigma_n - \sigma)| > (\gamma_1 B\lambda_{\max}(\Gamma_n) - K/B)/2\}$$
$$\geq 1 - \varepsilon/2$$

for K large enough, and

$$P\{t'E_n(t,\sigma_n) - t'E_n(0,\sigma) + \gamma_1 t'\Gamma_n t + \gamma_2 n^{1/2}(\sigma_n - \sigma)t'a_1 > K/2 \text{ for all } t \in \mathscr{B}\}$$
$$\leq P\{|E_n(t,\sigma_n) - E_n(0,\sigma) + \gamma_1\Gamma_n t + \gamma_2 n^{1/2}(\sigma_n - \sigma)a_1| > K/2 \text{ for all } t \in \mathscr{B}\}$$
$$\leq P\left\{\sup_{|t|\leq B} |E_n(t,\sigma_n) - E_n(0,\sigma) + \gamma_1\Gamma_n t + \gamma_2 n^{1/2}(\sigma_n - \sigma)a_1| > K/2B\right\}$$
$$< \varepsilon/2$$

by Lemma 4.1, so

$$P\{t'E_n(t,\sigma_n) < 0 \text{ for all } t \in \mathscr{B}\} \geq 1 - \varepsilon$$

and (4.1) obtains.

To complete the proof of Theorem 4.1, replace t by $n^{1/2}(\mu_n - \theta_0 - T_M(F)a_1)$ and u by $n^{1/2}\log(\sigma_n/\sigma)$ in Lemma 4.1, apply (4.3) and rearrange the remaining terms. □

Now consider M-estimators based on ψ_s. We are able to prove the following analogue of Theorem 4.1 for the case that ψ_a and ψ_c vanish. Since ψ_s is not continuous, the proof uses an argument of Jurečková (1977) which exploits the monotonicity of ψ_s.

THEOREM 4.2. *Suppose that Conditions* (i)–(iii), (M1), (M2) *and* (M5) *hold with* $\tau = T_M(F)$ *defined in* (1.5). *Let* $a_1' = (1,0,\ldots,0) \in \mathbb{R}^p$. *Then if* $\psi_a = \psi_c = 0$, *it follows that for* μ_n *satisfying* (1.4),

$$n^{1/2}(\mu_n - \theta_0 - T_M(F)a_1) = O_p(1)\,. \tag{4.4}$$

Moreover, for any μ_n *satisfying* (1.4) *and* (4.4), (4.5) $$\mu_n - \theta_0 - T_M(F)a_1 = \xi_1^{-1}(X'X)^{-1} \sum_{j=1}^{n} x_j\psi((Z_j - T_M(F))/\sigma) - \xi_1^{-1}\xi_2 a_1(\sigma_n - \sigma) + O_p(n^{-3/4}),$$

where ξ_1 and ξ_2 and defined in (2.4) *and* (2.5), *respectively, with $\tau = T_M(F)$.*

PROOF. As in the proof of Theorem 4.1, define

$$E_n(t, \sigma) = n^{-1/2} \sum_{j=1}^{n} x_j\psi((Z_j - T_M(F) - n^{-1/2}x_j't)/\sigma), \quad t \in \mathbb{R}^p, \quad \sigma > 0 .$$

Then it follows from (1.4) that for any ε, $K > 0$, there is an $n_0 > 0$ such that for $n > n_0$,

$$P\{|E_n(n^{1/2}(\mu_n - \theta_0 - T_M(F)a_1), \sigma_n)| \geq K\} < \varepsilon$$

and hence for B, $K > 0$ that,

$$\begin{aligned} &P\{n^{1/2}|\mu_n - \theta_0 - T_M(F)a_1| \geq B\} \\ &\quad \leq P\left\{ \inf_{|t| \geq B} |E_n(t, \sigma_n)| < K \right\} \\ &\qquad + P\left\{ n^{1/2}|\mu_n - \theta_0 - T_M(F)a_1| \geq B, \inf_{|t| \geq B} |E_n(t, \sigma_n)| \geq K \right\} \\ &\quad = P\left\{ \inf_{|t| \geq B} |E_n(t, \sigma_n)| < K \right\} + \varepsilon . \end{aligned}$$

If $|t| \geq B$, put $v = Bt/|t|$ so that $|v| = B$ and $t = \tau v$ with $\tau = |t|/B \geq 1$. Then

$$\begin{aligned} |E_n(t, \sigma_n)| &\geq - t'E_n(t, \sigma_n)/|t| \\ &= - v'E_n(\tau v, \sigma_n)/B \\ &\geq - v'E_n(v, \sigma_n)/B \end{aligned}$$

as $- v'E_n(\tau v, \sigma_n)$ is non-decreasing in τ. Consequently, with $\Gamma_n = n^{-1}X'X$,

$$\begin{aligned} &P\left\{ \inf_{|t| \geq B} |E_n(t, \sigma_n)| < K \right\} \\ &\quad \leq P\left\{ \inf_{|v| = B} - v'E_n(v, \sigma_n) < KB \right\} \end{aligned}$$

$$\leq P\left\{\inf_{|v|=B} - v'E_n(0,\sigma) + \xi_1 v'\Gamma_n v + \xi_2 n^{1/2}(\sigma_n - \sigma)v'a_1 < 2KB\right\}$$

$$+ P\left\{\inf_{|v|=B} - v'E_n(v,\sigma_n) < KB,\right.$$

$$\left.\inf_{|v|=B} - v'E_n(0,\sigma) + \xi_1 v'\Gamma_n v + \xi_2 n^{1/2}(\sigma_n - \sigma)v'a_1 > 2KB\right\}.$$

If $\lambda_{\min}(\Gamma_n)$ is the smallest eigenvalue of Γ_n,

$$P\left\{\inf_{|t|\geq B} |E_n(t,\sigma_n)| < K\right\}$$

$$\leq P\{-B|E_n(0,\sigma) - \xi_2 n^{1/2}(\sigma_n - \sigma)a_1| < 2KB - \xi_1 B^2 \lambda_{\min}(\Gamma_n)\}$$

$$+ P\left\{\sup_{|v|=B} v'E_n(v,\sigma_n) - v'E_n(0,\sigma)\right.$$

$$\left.+ \xi_1 v'\Gamma_h v - \xi_2 n^{1/2}(\sigma_n - \sigma)v'a_1 > KB\right\}$$

$$\leq P\{|E_n(0,\sigma) - \xi_2 n^{1/2}(\sigma_n - \sigma)a_1| > B\xi_1\lambda_{\min}(\Gamma) - 2K\}$$

$$+ P\left\{\sup_{|v|=B} |E_n(v,\sigma_n) - E_n(0,\sigma) + \xi_1 v'\Gamma_n v + \xi_2 n^{1/2}(\sigma_n - \sigma)a_1| > K\right\}$$

$$\leq \varepsilon$$

for B large enough, for any $\varepsilon > 0$, by Chebychev's inequality and Lemma 4.1.

The representation (4.5) follows from (4.4), (4.3) and Lemma 4.1. □

Combining the above results, we immediately obtain the following result for the general class of M-estimators.

THEOREM 4.3. *Suppose that Conditions* (i)–(iii) *and* (M1)–(M5) *hold with* $\tau = T_M(F)$ *defined in* (1.5). *Let* $a_1' = (1,0,\dots,0) \in \mathbb{R}^p$. *Then if* ψ *is either continuous or monotone, it follows that for* μ_n *satisfying* (1.4)

$$n^{1/2}(\mu_n - \theta_0 - T_M(F)a_1) = O_p(1)\,. \tag{4.6}$$

Moreover, for any μ_n *satisfying* (1.4) *and* (4.6),

$$\mu_n - \theta_0 - T_M(F)a_1 = (\gamma_1 + \xi_1)^{-1}(X'X)^{-1}\sum_{j=1}^{m} x_j\psi((Z_j - T_M(F))/\sigma)$$

$$-(\gamma_1 + \xi_1)^{-1}(\gamma_2 + \xi_2)a_1(\sigma_n - \sigma) + O_p(n^{-s})$$

with

$$s = \begin{cases} 1 & \text{if } \psi_s = 0 \\ 3/4 & \text{else}, \end{cases}$$

where γ_1, γ_2, ξ_1 *and* ξ_2 *are defined in* (2.2)–(2.5), *respectively, with* $\tau = T_M(F)$.

5. Representations for one-step M-estimators

Asymptotic representations for one-step M-estimators can be derived from the asymptotic linearity results of the previous section. For the reasons noted in the introduction, we will restrict attention to type II estimators and obtain analogues of Theorems 4.1–4.3 for these estimators.

THEOREM 5.1. *Suppose that Conditions* (i)–(iii), (L1) *and* (M1)–(M4) *hold with* $\tau = \tau_0$. *Let* $a_1' = (1, 0, \ldots, 0) \in \mathbb{R}^p$. *Then if* $\tilde{\mu}_n$ *is a type* II *one-step M-estimator and* $\psi_s = 0$,

$$\begin{aligned} \tilde{\mu}_n - \theta_0 - \{\tau_0 + \gamma_1^{-1} E\psi((Z - \tau_0)/\sigma)\} a_1 &= \gamma_1^{-1}(X'X)^{-1} \sum_{j=1}^{n} x_j \{\psi((Z_j - \tau_0)/\sigma) - E\psi((Z - \tau_0)/\sigma)\} \\ &\quad - \gamma_1^{-1}\gamma_2(\sigma_n - \sigma) a_1 - \gamma_1^{-2} E\psi((Z - \tau_0)/\sigma)(\gamma_{1n} - \gamma_1) a_1 + O_p(n^{-1}), \end{aligned}$$

where γ_1 *and* γ_2 *are defined by* (2.2) *and* (2.3), *respectively, with* $\tau = \tau_0$.

PROOF. Notice that

$$\begin{aligned} \tilde{\mu}_n - \theta_0 - \{\tau_0 + \gamma_1^{-1} E\psi((Z - \tau_0)/\sigma)\} a_1 &= \tau_n - \theta_0 - \tau_0 a_1 \\ &\quad + \gamma_{1n}^{-1}(X'X)^{-1} \sum_{j=1}^{n} x_j \{\psi((r_j - \tau_{n1})/\sigma_n) - E\psi((Z - \tau_0)/\sigma)\} \\ &\quad + (\gamma_{1n}^{-1} - \gamma_1^{-1}) a_1 E\psi((Z - \tau_0)/\sigma) \\ &= \gamma_{1n}^{-1}(X'X)^{-1} \Big[\gamma_1 X'X(\tau_n - \theta_0 - \tau_0 a_1) + \sum_{j=1}^{n} x_j \{\psi((r_j - \tau_{n1})/\sigma) \\ &\qquad - E\psi((Z - \tau_0)/\sigma) - E\psi((Z - \tau_0)/\sigma)\} \Big] \\ &\quad + \gamma_{1n}^{-1}(\gamma_{1n} - \gamma_1)(\tau_n - \theta_0 - \tau_0 a_1) \\ &\quad - (\gamma_{1n}\gamma_1)^{-1}(\gamma_{1n} - \gamma_1) a_1 E\psi((Z - \tau_0)/\sigma). \end{aligned}$$

Now

$$n^{-1}\sum_{j=1}^{n} x_j\{\psi((r_j-\tau_{n1})/\sigma_n) - E\psi((Z-\tau_0)/\sigma)\} + n^{-1}\gamma_1 X'X(\tau_n-\theta_0-\tau_0 a_1)$$
$$= n^{-1}\sum_{j=1}^{n} x_j\{\psi((Z_j-\tau_0)/\sigma) - E\psi((Z-\tau_0)/\sigma)\}$$
$$-\gamma_2(\sigma_n-\sigma)a_1 + O_p(n^{-1}),$$

by Lemma 4.1, so the result will follow if we can show that

$$\gamma_{1n}-\gamma_1 = O_p(n^{-1/2}).$$

But

$$\begin{aligned}\gamma_{1n}-\gamma_1 &= \frac{1}{2}n^{-1/2}\sum_{j=1}^{n}\{\psi((r_j-\tau_{n1}+n^{-1/2})/\sigma_n)\\ &\qquad -\psi((r_j-\tau_{n1}-n^{-1/2})/\sigma_n)\}-\gamma_1\\ &= \frac{1}{2}n^{-1/2}\sum_{j=1}^{n}\{\psi((r_j-\tau_{n1}+n^{-1/2})/\sigma_n)-\psi((Z_j-\tau_0)/\sigma)\}\\ &\quad+\frac{1}{2}n^{1/2}\{\gamma_1\bar{x}'(\tau_n-\theta_0-\tau_0a_1-n^{-1/2}a_1)+\gamma_2(\sigma_n-\sigma)\}\\ &\quad-\frac{1}{2}n^{-1/2}\sum_{j=1}^{n}\{\psi((r_j-\tau_{n1}+n^{-1/2})/\sigma_n)-\psi((Z_j-\tau_0)/\sigma)\}\\ &\quad-\frac{1}{2}n^{1/2}\{\gamma_1\bar{x}'(\tau_n-\theta_0-\tau_0a_1-n^{-1/2}a_1)+\gamma_2(\sigma_n-\sigma)\}\\ &= O_p(n^{-1/2})\end{aligned}$$

by Lemma 4.1, so the result obtains. □

Condition (S1) ensures that we can estimate ξ_1 so by a very similar argument to that used to prove Theorem 5.1, we immediately obtain the following two theorems.

THEOREM 5.2. *Suppose that conditions* (i)–(iii), (L1), (M1), (M2), (M5) *and* (S1) *hold with* $\tau=\tau_0$. *Let* $a_1'=(1,0,\ldots,0)\in\mathbb{R}^p$. *Then if* $\tilde{\mu}_n$ *is a type II one-step M-estimator and* $\psi_a=\psi_c=0$,

$$\begin{aligned}\tilde{\mu}_n - \theta_0 - \{\tau_0 + \xi_1^{-1} E\psi((z-\tau_0)/\sigma)\}a_1 \\ = \xi_1^{-1}(X'X)^{-1} \sum_{j=1}^{n} x_j\{\psi((Z_j-\tau_0)/\sigma) - E\psi((Z-\tau_0)/\sigma)\} \\ - \xi_1^{-1}\xi_2(\sigma_n - \sigma)a_1 - \xi_1^{-2}E\psi((Z-\tau_0)/\sigma)(\xi_{1n} - \xi_1)a_1 + O_p(n^{-3/4}),\end{aligned}$$

where ξ_1 *and* ξ_2 *are defined in* (2.4) *and* (2.5), *respectively, with* $\tau = \tau_0$.

THEOREM 5.3. *Suppose that Conditions* (i)–(iii), (L1), (M1)–(M5) *and* (S1) *hold with* $\tau = \tau_0$. *Let* $a_1' = (1, 0, \ldots, 0) \in \mathbb{R}^p$. *Then if* $\tilde{\mu}_n$ *is a type* II *one-step M-estimator,*

$$\begin{aligned}\tilde{\mu}_n - \theta_0 - \{\tau_0 + (\gamma_1 + \xi_1)^{-1} E\psi((Z-\tau_0)/\sigma)\}a_1 \\ = (\gamma_1 + \xi_1)^{-1}(X'X)^{-1} \sum_{j=1}^{n} x_j\{\psi((Z_j-\tau_0)/\sigma) - E\psi((Z-\tau_0)/\sigma)\} \\ - (\gamma_1 + \xi_1)^{-1}(\gamma_2 + \xi_2)(\sigma_n - \sigma)a_1 \\ - (\gamma_1 + \xi_1)^{-2}E\psi((Z-\tau_0)/\sigma)(\gamma_{1n} + \xi_{1n} - \gamma_1 - \xi_1)a_1 + O_p(n^{-s})\end{aligned}$$

with

$$s = \begin{cases} 1 & \text{if } \psi_s = 0 \\ 3/4 & \text{else}, \end{cases}$$

where γ_1, γ_2, ξ_1 *and* ξ_2 *are defined in* (2.2)–(2.5), *respectively, with* $\tau = \tau_0$.

6. Asymptotic relations

In this section, we examine the asymptotic relationship between *L*-, *M*- and one-step *M*-estimators and, in particular, give conditions under which they are asymptotically equivalent. However, it is worth keeping in mind that there are important differences between the classes of estimators. In *L*-estimation, observations are weighted according to their position in the sample while in *M*-estimation, observations are weighted according to their magnitude. A consequence of this distinction is that while *L*-estimators are naturally scale equivariant, *M*-estimators usually require a concomittant scale estimate to achieve scale equivariance.

It follows immediately from the results of Sections 4 and 5 that *M*- and one-step *M*-estimators will be asymptotically equivalent under the conditions of Section 5 provided $\tau_0 = T_M(F)$ defined in (1.5). Except in the case that F is symmetric, this condition will usually only be satisfied if $\tau_n = \mu_n$! Notice that this is true for the slope as well as the intercept.

Now consider the relationship between *L*- and *M*-estimators. It is apparent from the results of Sections 3 and 4 that asymptotically equivalent L- and M-estimators will be related through h, $\{(w_i, q_i)\colon 1 \le i \le m\}$ and ψ_c, $\{(d_l, Q_l)\colon 1 \le l \le k\}$. In fact, it follows that given an L-estimator λ_n based on h, $\{(w_i, q_i)\colon 1 \le i \le m\}$ we can construct a related M-estimator μ_n by taking

$$\psi_c(z) = -\int \{I(\sigma z \le y) - F(y)\} h(F(y)) dy ,$$

$$d_l = \begin{cases} \sum_{i=1}^{m} \dfrac{w_i q_i}{\phi(q_i)} - \sum_{i=l+1}^{m} \dfrac{w_i}{\phi(q_i)}, & 0 \le l \le m-1 \\ \sum_{i=1}^{m} \dfrac{w_i q_i}{\phi(q_i)}, & l = m \end{cases}$$

and

$$Q_l = F^{-1}(q_l)/\sigma \qquad 1 \le l \le m .$$

Conversely, given an M-estimator μ_n based on ψ_c, $\{(d_l, Q_l)\colon 1 \le l \le k\}$, we can construct a related L-estimator λ_n by taking

$$h(u) = \sigma^{-1} \psi_c'(F^{-1}(u)/\sigma) ,$$
$$w_i = (d_i - d_{i-1}) f(\sigma Q_i), \qquad 1 \le i \le k$$

and

$$q_i = F(\sigma Q_i), \qquad 1 \le i \le k .$$

For simplicity, we will restrict attention to the case that h, $\{(w_i, q_i)\colon 1 \le i \le m\}$ is given. The results below give conditions under which the slope components of related estimators are asymptotically equivalent. For the intercept components of related estimators to be asymptotically equivalent, we will additionally require that σ is known or that $\gamma_2 = \xi_2 = 0$, where γ_2 and ξ_2 are defined in (2.3) and (2.5), respectively. The first condition is often unrealistic and the second is satisfied when F is symmetric, ψ_c is antisymmetric and $\{(d_l, Q_l)\colon 1 \le l \le k\}$ are appropriately chosen.

It is interesting that the tail conditions for M-estimators are different from those for L-estimators. For untrimmed L-estimators, we require Condition (L3), that is

$$\int_{-\infty}^{\infty} \left\{ \sup_{|u| \le \delta} |h_a'(z+u)| \right\}^{\kappa} dF(z) < \infty ,$$

and

$$\int_{-\infty}^{\infty}[F(y+\varepsilon)\{1-F(y-\varepsilon)\}]^{1/2}dy<\infty \quad \text{for} \quad \varepsilon>0 .$$

However, for related M-estimators, we require

(E) $\int_{-\infty}^{\infty}\psi_L(z/\sigma)dF(z)<\infty$, and for some $\nu>1$ and $\delta_0>0$,

$$\int_{-\infty}^{\infty}\left\{|z| \sup_{|v|\le\delta}\sup_{|u|\le\delta}|h_a'(e^{-v}(z+u))|\right\}^{\kappa}dF(z)<\infty ,$$

and

$$\int_{-\infty}^{\infty}\left\{z^2\sup_{|u|\le\delta}|h_a'(z+u)|\right\}^{\kappa}dF(z)<\infty, \quad \text{for all} \quad 0<\delta\le\delta_0 .$$

As noted in the introduction, these conditions hold if h_a' is bounded and trimmed or bounded and $\int_{-\infty}^{\infty}|z|^{2+\varepsilon}dF(z)<\infty$, for some $\varepsilon>0$, holds. This moment condition and the tail condition for L-estimators are closely related (compare them when F is regularly varying) but are not in general identical.

The following result is for the case that $m=0$.

THEOREM 6.1. *Suppose that Conditions* (i)–(iii), (L1)–(L3), (M1) *and* (E) *hold. Let* $a_1=(1,0,\ldots,0)'\in\mathbb{R}^p$. *Then if* $m=0$ *in* (1.3), *there exists a related M-estimator* μ_n *such that*

$$\lambda_n-\mu_n-T_L(F)a_1=\gamma_1^{-1}\gamma_2a_1(\sigma_n-\sigma)+O_p(n^{-1}) ,$$

where $T_L(F)$ *is defined in* (1.2) *and* γ_1 *and* γ_2 *are defined in* (2.2) *and* (2.3), *respectively.*

PROOF. The representation for λ_n obtained in Theorem 3.1 holds so the result will follow if we can show that the conditions of Theorem 4.1 hold with $\psi=\psi_L$. Now

$$E\psi_L(Z/\sigma)=0$$

so $T_M(F)=0$. It is straightforward to show that these conditions do in fact hold. □

For the case that h vanishes, we have the following result.

THEOREM 6.2. *Suppose that Conditions* (i)–(iii), (L1), (L2), (L5) *and* (M1) *hold. Then if* $\psi_c=0$ *in* (1.3), *it follows that for the related M-*

estimator μ_n *satisfying* (1.4),

$$\lambda_n - \mu_n T_L(F) a_1 = \xi_1^{-1} \xi_2 a_1 (\sigma_n - \sigma) + O_p(n^{-3/4}) \,,$$

where $T_L(F)$ *is defined in* (1.2), ξ_1 *and* ξ_2 *are defined in* (2.4) *and* (2.5) *with* $\tau = 0$, *respectively, and* $a_1' = (1, 0, \ldots, 0) \in \mathbb{R}^p$.

PROOF. The result will follow from Theorems 3.2 and 4.2 if we can show that Conditions (M2) and (M5) hold with $\tau = 0$. It is straightforward to show that these conditions do in fact hold. □

Finally, we note that we can combine Theorems 6.1 and 6.2 to obtain asymptotic equivalence results for *L*- and *M*-estimators with both a discrete and a continuous component and that results in which σ is treated as known can be derived from the present results in a straightforward manner.

Acknowledgements

The authors would like to thank both referees for their careful reading of our paper.

REFERENCES

Bickel, P. J. (1973). On some analogues to linear combinations of order statistics in the linear model, *Ann. Statist.*, **1**, 597–616.

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

Huber, P. J. (1973). Robust regression: Asymptotics, conjectures and Monte Carlo, *Ann. Statist.*, **1**, 799–821.

Jaeckel, L. A. (1971). Robust estimates of location: Symmetry and asymmetric contamination, *Ann. Math. Statist.*, **42**, 1020–1034.

Jurečková, J. (1977). Asymptotic relations of *M*-estimates and *R*-estimates in linear regression model, *Ann. Statist.*, **5**, 464–472.

Jurečková, J. (1983*a*). Robust estimators of location and regression parameters and their second order asymptotic relations, *Proc. 9th Prague Conf. on Inform. Theory, Random Processes and Statist. Decision Function*, (ed. J. A. Višek), 19–32, Reidel, Dordrecht.

Jurečková, J. (1983*b*). Winsorised least squares estimator and its *M*-estimator counterpart, *Contributions to Statistics: Essays in honour of Norman L. Johnson*, (ed. P. K. Sen), 237–245, North Holland, Amsterdam.

Jurečková, J. (1983*c*). Trimmed polynomial regression, *Comment. Math. Univ. Carolin.*, **24**, 597–607.

Jurečková, J. (1984). Regression quantiles and trimmed least squares estimator under a general design, *Kybernetika* (*Prague*), **20**, 345–357.

Jurečková, J. (1986). Asymptotic representations of *L*-estimators and their relations to *M*-estimators, *Sequential Anal.*, **5**, 317–338.

Jurečková, J. and Portnoy, S. (1987). Asymptotics for one-step *M*-estimators in regression with application to combining efficiency and high breakdown point, *Comm. Statist. A—Theory Methods*, **16**, 2187–2199.

Jurečková, J. and Sen, P. K. (1984). On adaptive scale-equivariant *M*-estimators in linear models, *Statist. Decisions*, **2** (Suppl. Issue No. 1), 31–46.

Jurečková, J. and Sen, P. K. (1989). Uniform second order asymptotic linearity of *M*-statistics in linear models, *Statist. Decisions*, **7**, 263–276.

Koenker, R. and Bassett, G. (1978). Regression quantiles, *Econometrica*, **46**, 33–50.

Koenker, R. and Portnoy, S. (1987). *L*-estimation for linear models, *J. Amer. Statist. Assoc.*, **82**, 851–857.

Koul, H. L. (1969). Asymptotic behavior of Wilcoxon type confidence regions, *Ann. Math. Statist.*, **40**, 1950–1979.

Ortega, J. M. and Rheinboldt, W. C. (1973). *Iterative Solution of Nonlinear Equations in Several Variables*, Academic Press, New York.

Relles, D. (1968). Robust regression by modified least squares, Ph. D. Thesis, Yale University.

Rivest, L. P. (1982). Some asymptotic distributions in the location-scale model, *Ann. Inst. Statist. Math.*, **34**, 225–239.

Ruppert, D. and Carroll, R. J. (1980). Trimmed least squares estimation in the linear model, *J. Amer. Statist. Assoc.*, **75**, 828–838.

van Eeden, C. (1983). On the asymptotic relation between *L*-estimators and *M*-estimators and their asymptotic efficiency relative to the Cramér-Rao lower bound, *Ann. Statist.*, **11**, 674–690.

Welsh, A. H. (1986). Bahadur representations for robust scale estimators based on regression residuals, *Ann. Statist.*, **14**, 1246–1251.

Welsh, A. H. (1987*a*). The trimmed mean in the linear model (with discussion), *Ann. Statist.*, **15**, 20–36 (correction: *ibid.* (1988). **16**, 480).

Welsh, A. H. (1987*b*). One-step *L*-estimators for the linear model, *Ann. Statist.*, **15**, 626–641 (correction: *ibid.* (1988). **16**, 481).

Welsh, A. H. (1987*c*). Kernel estimates of the sparsity function, *Statistical Data Analysis Based on the L_1-norm and Related Methods*, (ed. Y. Dodge), 369–378, Elsevier Science Publishers B. V., Amsterdam.

Yohai, V. J. and Maronna, R. A. (1979). Asymptotic behaviour of *M*-estimators for the linear model, *Ann. Statist.*, **7**, 258–268.