The Annals of Statistics
1987, Vol. 15, No. 4, 1548–1562

CONSISTENT ESTIMATION OF THE INFLUENCE FUNCTION OF LOCALLY ASYMPTOTICALLY LINEAR ESTIMATORS[1]

BY CHRIS A. J. KLAASSEN

University of Leiden

Consider estimators which behave locally asymptotically like an average of some function taken at the observations. This function is called the influence function and one calls such estimators locally asymptotically linear. It is shown that the influence function of a locally asymptotically linear estimator can be estimated consistently and conversely, that, given a consistent estimator of the influence function, estimators can be constructed which are locally asymptotically linear in that influence function. With the help of these results an adaptive estimator is constructed for a partially irregular model.

1. Introduction. Let $X_1, \dots, X_n$ be independent and identically distributed random variables with distribution $P_{\vartheta, g}$ on some measurable space $(\mathbf{X}, \mathbf{A})$. The parameter (ϑ, g) of $P_{\vartheta, g}$ is known to belong to a set $\Theta \times G$, with $\Theta \subset \mathbf{R}^d$ open and G some set of functions, say, but otherwise (ϑ, g) is unknown. For every $g \in G$, let $J(\cdot;\cdot, g)$: $\mathbf{X} \times \Theta \to \mathbf{R}^d$ be a measurable function satisfying

$$\int J(x; \vartheta, g)\, dP_{\vartheta, g}(x) = 0, \qquad \vartheta \in \Theta, \tag{1.1}$$

and

$$\int |J(x; \vartheta, g)|^2\, dP_{\vartheta, g}(x) < \infty, \qquad \vartheta \in \Theta, \tag{1.2}$$

where $|\cdot|$ denotes the Euclidean norm.

In this semiparametric model we shall study estimator sequences $\{T_n\}$ of ϑ, $T_n = t_n(X_1, \dots, X_n)$ and t_n: $\mathbf{X}^n \to \mathbf{R}^d$ measurable, which are locally asymptotically linear in the following sense. For every $(\vartheta, g) \in \Theta \times G$ and for every sequence $\{\vartheta_n\}$ with $|\vartheta_n - \vartheta| = O(n^{-1/2})$,

$$n^{1/2}\left(T_n - \vartheta_n - n^{-1} \sum_{i=1}^{n} J(X_i; \vartheta_n, g)\right) = o_{\vartheta_n, g}(1), \quad \text{as } n \to \infty, \tag{1.3}$$

holds. Here $o_{\vartheta_n, g}(1)$ is shorthand for "tending to zero in $P_{\vartheta_n, g}$-probability." In view of (1.3), $J(\cdot; \vartheta, g)$ is called the influence function of $\{T_n\}$ under (ϑ, g). Estimators, which satisfy (1.3) for $\vartheta_n = \vartheta$ and may violate (1.3) for other sequences $\{\vartheta_n\}$, are called asymptotically linear. Of course, locally asymptotically linear estimators are the more desirable ones, but in sufficiently smooth models asymptotically linear estimators are locally asymptotically linear automatically [see the discussion between formulas (2.2) and (2.3)].

Received May 1985; revised February 1987.
[1]Research partially supported by ONR Contract N00014-80-C-0163.
AMS 1980 *subject classifications*. Primary 62G05; secondary 62F12, 62F35.
Key words and phrases. Influence function, asymptotically linear estimators, consistency, adaptive estimation, irregular model.

In models without nuisance parameters, i.e., with g known and $G = \{g\}$, typically estimators of ϑ exist, which satisfy (1.3) with $J(\cdot\,; \vartheta, g)$ being the so-called efficient influence function and which are, therefore, efficient. Note that it follows from Theorem 4.2 of Hájek (1972) that under local asymptotic normality, estimators of which all coordinates are locally asymptotically minimax have to be asymptotically linear in the efficient influence function.

Efficient estimators satisfying (1.3) also exist in some semiparametric models in which g is unknown and ranges over some "big" set G. Such so-called adaptive estimators have been presented in the symmetric location problem by many authors. To mention but a few: van Eeden (1970), Stone (1975) and Beran (1978). Bickel (1982) has given a construction of adaptive estimators which works in more general situations too.

In Theorem 6.3.1 and Corollary 6.3.2 of Huber (1981), conditions are given which guarantee asymptotic linearity of M-estimators. Typically, the influence functions are not efficient here.

In Section 2, we will present an estimator which is locally asymptotically linear under smoothness conditions on the model and the influence function, under the assumption of existence of an $n^{1/2}$ consistent estimator of ϑ and under the assumption that the influence function can be estimated consistently in the following sense. There exists an estimator sequence $\{\hat{J}_n(\cdot\,;\cdot\,; X_1,\ldots, X_n)\}$, based on $X_1,\ldots, X_n$, such that for every $(\vartheta, g) \in \Theta \times G$ and for every sequence $\{\vartheta_n\}$ with $|\vartheta_n - \vartheta| = O(n^{-1/2})$,

$$(1.4) \qquad \int \left| \hat{J}_n(x; \vartheta_n; X_1,\ldots, X_n) - J(x; \vartheta_n, g)\right|^2 dP_{\vartheta_n, g}(x) = o_{\vartheta_n, g}(1)$$

and

$$(1.5) \qquad n^{1/2} \int \hat{J}_n(x, \vartheta_n; X_1,\ldots, X_n)\, dP_{\vartheta_n, g}(x) = o_{\vartheta_n, g}(1)$$

hold. Note that Condition H of Bickel (1982) is stronger in that it forces the left-hand side of (1.5) to vanish. The sufficiency of conditions (1.4) and (1.5) for the existence of efficient asymptotically linear estimators has already been shown by Schick (1986) in a class of semiparametric models, which is slightly more restricted than ours. In Section 3, we shall prove under a mild regularity condition that the existence of a consistent estimator of the influence function is also necessary for the existence of locally asymptotically linear estimators, i.e., that (1.3) implies (1.4) and (1.5).

The results of Sections 2 and 3 thus imply that, under regularity conditions on the model and the desired influence function and provided an $n^{1/2}$ consistent estimator of ϑ exists, an estimator of ϑ which is locally asymptotically linear in this influence function exists, iff this influence function can be estimated consistently in the sense of (1.4) and (1.5). Consequently, if a locally asymptotically linear estimator can be constructed, it can be done by first estimating the influence function consistently and then applying one of the methods of Section 2. Along this line of approach most adaptive estimators have been constructed.

Some examples and a discussion of these results are given in Section 4.

2. Constructions for locally asymptotically linear estimators. To the assumptions made in the first paragraph of the introduction, in this section we add contiguity:

$$\{P^n_{\vartheta_n, g}\} \triangleleft\triangleright \{P^n_{\vartheta, g}\} \tag{2.1}$$

and

$$n^{1/2}\left(\vartheta_n - \vartheta + n^{-1} \sum_{i=1}^{n} \{J(X_i; \vartheta_n, g) - J(X_i; \vartheta, g)\} \right) = o_{\vartheta, g}(1), \tag{2.2}$$

for every $(\vartheta, g) \in \Theta \times G$ and for every sequence $\{\vartheta_n\}$, with $|\vartheta_n - \vartheta| = O(n^{-1/2})$. In regular cases, contiguity is present and often (2.2) holds too [cf. (6.43) of Bickel (1982), Theorem 5.1.1 of van Eeden (1983) and (2.5) of Schick (1986)]. Moreover, if there exists an estimator which is locally asymptotically linear in the influence function J and if the contiguity (2.1) holds, then subtracting (1.3) from (1.3) with $\vartheta_n = \vartheta$ yields (2,2). In fact under (2.1) and (1.3) with $\vartheta_n = \vartheta$, the convergence relations (1.3) and (2.2) are equivalent. Consequently, in this situation asymptotic linearity and local asymptotic linearity are the same. Indeed, twice (1.3); once as it stands, once with $\vartheta_n = \vartheta$.

Furthermore, we assume the existence of a (preliminary) estimator sequence $\{\tilde{T}_n\}$ of ϑ, $\tilde{T}_n = \tilde{t}_n(X_1, \dots, X_n)$, with

$$n^{1/2}(\tilde{T}_n - \vartheta) = O_{\vartheta, g}(1), \qquad (\vartheta, g) \in \Theta \times G. \tag{2.3}$$

Note that if (1.1)–(1.3) are fulfilled, then the $n^{1/2}$ consistency (2.3) of $\tilde{T}_n = T_n$ is a simple consequence of the central limit theorem.

If in this situation the parameter g is known, then Le Cam's discretization trick immediately yields a locally asymptotically linear estimator as follows. Define $\bar{T}_{n0}$ as a point in $n^{-1/2}\mathbf{Z}^d$ closest to $\tilde{T}_n$. Now

$$\bar{T}_n = \bar{T}_{n0} + n^{-1} \sum_{i=1}^{n} J(X_i; \bar{T}_{n0}, g) \tag{2.4}$$

satisfies (1.3) [cf. the argument after (3.7) of Bickel (1982)].

Another method would be "splitting" as follows. Let $\{\mu_n\}$ be a sequence of positive integers with $\mu_n n^{-1} \to \mu$, $0 < \mu < 1$. Define

$$T^*_{n1} = \tilde{t}_{\mu_n}(X_1, \dots, X_{\mu_n}), \qquad T^*_{n2} = \tilde{t}_{n-\mu_n}(X_{\mu_n+1}, \dots, X_n). \tag{2.5}$$

Note that because of independence,

$$n^{1/2}\left(T^*_{n2} - \vartheta + \mu_n^{-1} \sum_{i=1}^{\mu_n} \{J(X_i; T^*_{n2}, g) - J(X_i; \vartheta, g)\} \right) = o_{\vartheta, g}(1) \tag{2.6}$$

holds. Consequently,

$$\begin{aligned} T_n^* = {} & \mu_n n^{-1}\left\{ T^*_{n2} + \mu_n^{-1} \sum_{i=1}^{\mu_n} J(X_i; T^*_{n2}, g) \right\} \\ & + (n - \mu_n) n^{-1}\left\{ T^*_{n1} + (n - \mu_n)^{-1} \sum_{i=\mu_n+1}^{n} J(X_i; T^*_{n1}, g) \right\} \end{aligned} \tag{2.7}$$

satisfies (1.3).

If J and, hence, g are unknown, then we would like to replace J in (2.7) by an estimator of it. To make this work we push the "splitting" trick leading to (2.7) one step further by splitting the sample $X_1, \ldots, X_{\mu_n}$, on which T_{n1}^* is based, into two parts: one part that produces a preliminary estimate of ϑ and the other part that yields an estimate of J. In the same way the sample on which T_{n2}^* is based will be split. To be more precise, let $\{\lambda_n\}$, $\{\mu_n\}$ and $\{\nu_n\}$ be sequences of positive integers with

$$\lambda_n n^{-1} \to \lambda, \qquad \mu_n n^{-1} \to \mu, \qquad \nu_n n^{-1} \to \nu, \qquad 0 < \lambda < \mu < \nu < 1. \tag{2.8}$$

Define [cf. (1.4) and (1.5)]

$$\tilde{T}_{n1} = \tilde{t}_{\lambda_n}(X_1, \ldots, X_{\lambda_n}), \qquad \tilde{T}_{n2} = \tilde{t}_{\nu_n - \mu_n}(X_{\mu_n+1}, \ldots, X_{\nu_n}), \tag{2.9}$$

$$\begin{aligned} \hat{J}_{n1}(x; \theta) &= \hat{J}_{\mu_n - \lambda_n}(x; \theta; X_{\lambda_n+1}, \ldots, X_{\mu_n}), \\ \hat{J}_{n2}(x; \theta) &= \hat{J}_{n-\nu_n}(x; \theta; X_{\nu_n+1}, \ldots, X_n). \end{aligned} \tag{2.10}$$

In the proof of Theorem 2.1 it will be shown that, indeed,

$$\begin{aligned} T_n = {} & \mu_n n^{-1}\left\{\tilde{T}_{n2} + \mu_n^{-1} \sum_{i=1}^{\mu_n} \hat{J}_{n2}(X_i; \tilde{T}_{n2})\right\} \\ & + (n - \mu_n)n^{-1}\left\{\tilde{T}_{n1} + (n - \mu_n)^{-1} \sum_{i=\mu_n+1}^{n} \hat{J}_{n1}(X_i; \tilde{T}_{n1})\right\} \end{aligned} \tag{2.11}$$

satisfies (1.3).

Schick (1986) has used both the "discretization" and the "splitting" trick and has suggested in his formula (2.7)

$$\begin{aligned} \hat{T}_n = {} & \mu_n n^{-1}\left\{\overline{T}_{n0} + \mu_n^{-1} \sum_{i=1}^{\mu_n} \hat{J}_{n-\mu_n}(X_i; \overline{T}_{n0}; X_{\mu_n+1}, \ldots, X_n)\right\} \\ & + (n - \mu_n)n^{-1}\left\{\overline{T}_{n0} + (n - \mu_n)^{-1} \sum_{i=\mu_n+1}^{n} \hat{J}_{\mu_n}(X_i; \overline{T}_{n0}; X_1, \ldots, X_{\mu_n})\right\} \end{aligned} \tag{2.12}$$

as an asymptotically linear estimator [cf. (2.4) and (2.7)]. Note that Schick (1986) considers efficient influence functions in smooth, but still quite general semiparametric models. His conditions (2.1) and (A.1) on the underlying densities imply our (1.1), (1.2), (2.1) and (2.2).

Bickel (1982) has used a vanishingly small part of the observations to estimate the influence function. Such a construction cannot work here unless condition (1.5) is strengthened. Schick (1987) gives conditions stronger than (1.4) and (1.5) under which asymptotically linear estimators can be constructed using an estimator of the influence function based on the entire sample and not on just a part of it as in (2.11) and (2.12). Finally, note that (2.11) avoids the discretization as used in (2.12) and in (3.6) of Bickel (1982). In particular, this is important if one is interested in equivariant estimators in problems with a transformation group.

THEOREM 2.1. *Let* (1.1), (1.2) *and* (2.1)–(2.3) *hold. If there exists an estimator sequence* $\{\hat{J}_n(\cdot;\cdot;X_1,\dots,X_n)\}$ *of the influence function satisfying* (1.4) *and* (1.5), *then there exists an estimator sequence of* ϑ *satisfying* (1.3).

PROOF. In view of (2.1) and (2.2) it suffices to show that $\{T_n\}$ defined by (2.11) satisfies (1.3) with $\vartheta_n = \vartheta$. Therefore, we shall prove

$$(n-\mu_n)n^{-1/2}\left(\tilde{T}_{n1} - \vartheta + (n-\mu_n)^{-1} \times \sum_{i=\mu_n+1}^{n} \left\{\hat{J}_{n1}(X_i;\tilde{T}_{n1}) - J(X_i;\vartheta,g)\right\}\right) = o_{\vartheta,g}(1). \tag{2.13}$$

Because of (2.3) and the independence of $\tilde{T}_{n1}$ and $(X_{\lambda_n+1},\dots,X_n)$, this holds if for all $\{\vartheta_n\}$ with $|\vartheta_n - \vartheta| = O(n^{-1/2})$, we have

$$n^{1/2}\left(\vartheta_n - \vartheta + (n-\mu_n)^{-1} \times \sum_{i=\mu_n+1}^{n} \left\{\hat{J}_{n1}(X_i;\vartheta_n) - J(X_i;\vartheta,g)\right\}\right) = o_{\vartheta,g}(1). \tag{2.14}$$

Taking the conditional expectation of the Euclidean norm squared of the left-hand side of (2.15), given $X_{\lambda_n+1},\dots,X_{\mu_n}$, we see that relations (1.1), (1.4) and (1.5) yield

$$(n-\mu_n)^{-1/2}\sum_{i=\mu_n+1}^{n}\left\{\hat{J}_{n1}(X_i;\vartheta_n) - J(X_i;\vartheta_n,g)\right\} = o_{\vartheta_n,g}(1). \tag{2.15}$$

Contiguity, (2.15), (2.14) and (2.2) complete the proof. □

3. The necessity of the existence of consistent estimators of the influence function. In the situation of the first paragraph of the introduction we assume for a moment $d = 1$ and that the following strengthening of (1.3) holds:

$$\operatorname{var}_{\vartheta_n,g}\left\{n^{1/2}\left(T_n - \vartheta_n - n^{-1}\sum_{i=1}^{n} J(X_i;\vartheta_n,g)\right)\right\} = o(1). \tag{3.1}$$

By Hájek's projection lemma, this implies

$$\operatorname{var}_{\vartheta_n,g}\left\{\sum_{j=1}^{n} E_{\vartheta_n,g}\left(n^{1/2}(T_n-\vartheta_n) - n^{-1/2}\sum_{i=1}^{n} J(X_i;\vartheta_n,g)|X_j\right)\right\} = o(1) \tag{3.2}$$

and, hence, assuming t_n is symmetric in its arguments,

$$\operatorname{var}_{\vartheta_n,g}\left\{n\left(E_{\vartheta_n,g}(T_n|X_1) - \vartheta_n\right) - J(X_1;\vartheta_n,g)\right\} = o(1). \tag{3.3}$$

Since $n(E_{\vartheta_n,g}(T_n|X_1) - \vartheta_n)$ can be estimated very accurately, without knowledge of J, by taking k_n samples of size n [cf. (3.12)], it is intuitively clear from (3.3) that J can be estimated consistently in the sense of (1.4).

This argument can be made precise for general d without the symmetry assumption on t_n and (3.1) can be replaced by (1.3) by using a truncation trick. However, we need a regularity condition on the influence function.

THEOREM 3.1. *If* (1.1)–(1.3) *are fulfilled and, if for* $(\vartheta, g) \in \Theta \times G$,

$$\lim_{a\to\infty} \limsup_{\vartheta'\to\vartheta} \int |J(x;\vartheta',g)|^2 1_{\{|J(x;\vartheta',g)|\geq a\}}\, dP_{\vartheta',g}(x) = 0 \tag{3.4}$$

holds, then there exists an estimator sequence $\{\hat{J}_n(\cdot;\cdot;X_1,\dots,X_n)\}$ *satisfying* (1.4) *and* (1.5).

PROOF. In this proof we will suppress in the notation the dependence of var, E, $dP(x)$ and $o_P(1)$ on (ϑ_n, g). Furthermore, we will use the notation

$$\Delta_n = n^{1/2}(T_n - \vartheta_n), \qquad J_n = n^{-1/2}\sum_{i=1}^n J(X_i;\vartheta_n,g).$$

Since it suffices to consider (1.1)–(1.5) componentwise, we assume $d = 1$ without loss of generality.

Let $\psi\colon \mathbf{R}\to\mathbf{R}$ be an odd twice differentiable function with derivatives ψ' and ψ'' satisfying $|\psi| \leq 1$, $0 < \psi' \leq 1$ and $|\psi''| \leq 2$. Furthermore, denote $\psi_{nj}(x) = E(\psi(\Delta_n)|X_j = x)$ and $\psi_J(x) = E(\psi(J_n)|X_1 = x)$. Finally, let X_0 be a random variable such that $X_0, X_1,\dots,X_n$ are i.i.d. and let $\{\vartheta_n\}$ be such that $|\vartheta_n - \vartheta| = O(n^{-1/2})$.

By Hájek's projection lemma we have

$$\begin{aligned}
&\operatorname{var}\{\psi(\Delta_n) - \psi(J_n)\}\\
&\quad\geq \sum_{j=1}^n \operatorname{var}\{E(\psi(\Delta_n) - \psi(J_n)|X_j)\}\\
&\quad= \sum_{j=1}^n E[\psi_{nj}(X_0) - \psi_J(X_0) - E\psi(\Delta_n) + E\psi(J_n)]^2 \qquad (3.5)\\
&\quad\geq E\left\{n^{-1/2}\sum_{j=1}^n [\psi_{nj}(X_0) - \psi_J(X_0) - E\psi(\Delta_n) + E\psi(J_n)]\right\}^2,
\end{aligned}$$

which, together with (1.3) and the boundedness of ψ' and ψ, yields

$$\operatorname{var}\left\{n^{-1/2}\sum_{j=1}^n \psi_{nj}(X_0) - n^{1/2}\psi_J(X_0)\right\} = o(1). \tag{3.6}$$

With the notation

$$J_n^* = n^{-1/2}\sum_{i=2}^n J(X_i;\vartheta_n,g), \qquad \gamma_n = E\psi'(J_n^*), \tag{3.7}$$

the boundedness of ψ' and ψ'' implies

$$\begin{aligned}&\left|\gamma_n J(x;\vartheta_n,g) - n^{1/2}\psi_J(x) + E\big(n^{1/2}\psi(J_n^*)\big)\right| \\ (3.8)\qquad &= \left|E\int_0^{n^{-1/2}J(x;\vartheta_n,g)} n^{1/2}[\psi'(J_n^*) - \psi'(J_n^*+z)]\,dz\right| \\ &\le |J(x;\vartheta_n,g)| \wedge \big(n^{-1/2}J^2(x;\vartheta_n,g)\big)\end{aligned}$$

and, hence, (3.4) yields

$$\begin{aligned}&\limsup_{n\to\infty}\operatorname{var}\{\gamma_n J(X_0;\vartheta_n,g) - n^{1/2}\psi_J(X_0)\} \\ (3.9)\qquad &\le \lim_{a\to\infty}\limsup_{n\to\infty}\left\{EJ^2(X_0;\vartheta_n,g)\mathbf{1}_{\{|J(X_0;\vartheta_n,g)|\ge a\}} + n^{-1}a^4\right\} = 0.\end{aligned}$$

Together with (3.6), this implies

$$(3.10)\qquad \operatorname{var}\left\{n^{-1/2}\sum_{j=1}^n \psi_{nj}(X_0) - \gamma_n J(X_0;\vartheta_n,g)\right\} = o(1).$$

Let $\{k_n\}$ be a sequence of positive integers satisfying

$$(3.11)\qquad nk_n^{-1} = o(1)$$

and let Y_{ij}, $i = 1,\dots,k_n$, $j = 1,\dots,n$, be random variables such that $X_0, X_1,\dots,X_n, Y_{11},\dots,Y_{k_n n}$ are i.i.d. We define the random function J_n^ψ by

$$\begin{aligned}J_n^\psi(x;\theta) = k_n^{-1}\sum_{i=1}^{k_n}\Bigg\{&n^{-1/2}\sum_{j=1}^n \psi\big(n^{1/2}(t_n(Y_{i1},\dots,Y_{ij-1},x,Y_{ij+1},\dots,Y_{in}) - \theta)\big) \\ (3.12)\qquad &-n^{1/2}\psi\big(n^{1/2}(t_n(Y_{i1},\dots,Y_{in}) - \theta)\big)\Bigg\}, \\ &x\in\mathbf{X},\ \theta\in\Theta,\end{aligned}$$

and we note that

$$\begin{aligned}(3.13)\qquad &\operatorname{var}\left\{J_n^\psi(X_0;\vartheta_n) - n^{-1/2}\sum_{j=1}^n \psi_{nj}(X_0)\right\} \\ &= E\operatorname{var}\big(J_n^\psi(X_0;\vartheta_n)\big|X_0\big) \le 4nk_n^{-1}.\end{aligned}$$

Combining (3.13), (3.11), (3.10), $EJ_n^\psi(X_0;\vartheta_n) = 0$ and (1.1), we obtain

$$(3.14)\qquad E\{J_n^\psi(X_0;\vartheta_n) - \gamma_n J(X_0;\vartheta_n,g)\}^2 = o(1).$$

Let Z_{ij}, $i = 1,\dots,n$, $j = 1,\dots,n$, be random variables such that $X_0, X_1,\dots,X_n, Y_{11},\dots,Y_{k_n n}, Z_{11},\dots,Z_{nn}$ are i.i.d. and define

$$(3.15)\qquad \hat\gamma_n(\theta) = n^{-1/5} + n^{-1}\sum_{i=1}^n \psi'\big(n^{1/2}(t_n(Z_{i1},\dots,Z_{in}) - \theta)\big),\qquad \theta\in\Theta,$$

and

$$(3.16)\qquad \tilde\gamma_n = E\hat\gamma_n(\vartheta_n) = n^{-1/5} + E\psi'(\Delta_n).$$

By (1.1), (3.4), the central limit theorem and the symmetry and positivity of ψ', the limit points of $\{\gamma_n\}$ are positive and finite and by (1.3) and the continuity and boundedness of ψ',

$$\tilde{\gamma}_n - \gamma_n = o(1) \tag{3.17}$$

holds. Consequently, (3.14) implies

$$E\{\tilde{\gamma}_n^{-1} J_n^{\psi}(X_0; \vartheta_n) - J(X_0; \vartheta_n, g)\}^2 = o(1). \tag{3.18}$$

Furthermore, we have

$$E\big(\tilde{\gamma}_n^{-1} - \hat{\gamma}_n^{-1}(\vartheta_n)\big)^2 = E\left(\frac{\hat{\gamma}_n(\vartheta_n) - \tilde{\gamma}_n}{\tilde{\gamma}_n \hat{\gamma}_n(\vartheta_n)}\right)^2 \le n^{4/5} \operatorname{var} \hat{\gamma}_n(\vartheta_n) \le n^{-1/5} \tag{3.19}$$

and, hence, the independence of $\hat{\gamma}_n(\vartheta_n)$ and $J_n^{\psi}(X_0; \vartheta_n)$, the finiteness of the limit points of $\{\gamma_n\}$ and the formulas (3.14) and (3.4) yield

$$E\{\big(\tilde{\gamma}_n^{-1} - \hat{\gamma}_n^{-1}(\vartheta_n)\big) J_n^{\psi}(X_0; \vartheta_n)\}^2 = o(1). \tag{3.20}$$

Combining (3.18) and (3.20) we see that we may define the estimator sequence $\{\hat{\gamma}_n^{-1}(\cdot) J_n^{\psi}(\cdot;\cdot)\}$ of $J(\cdot;\cdot, g)$ satisfying (1.4). This estimator sequence is based on samples of size $nk_n + n^2$ and, consequently, in the preceding part of the proof it would have sufficed to consider $\{\vartheta_n\}$ with $|\vartheta_n - \vartheta| = O(n^{-1/2} k_n^{-1/2})$ instead of $O(n^{-1/2})$.

From the preceding, we conclude that there exists an estimator sequence $\{\tilde{J}_n(\cdot;\cdot)\}$ based on $X_1, \dots, X_n$ satisfying (1.4). Let $\{\mu_n\}$ be as in (2.8) and define

$$\begin{aligned} \hat{J}_n(x; \theta) = {} & \tilde{J}_{\mu_n}(x; \theta) + t_{n-\mu_n}\big(X_{\mu_n+1}, \dots, X_n\big) - \theta \\ & -(n - \mu_n)^{-1} \sum_{i=\mu_n+1}^{n} \tilde{J}_{\mu_n}(X_i; \theta). \end{aligned} \tag{3.21}$$

With the help of (1.3), (1.1) and (1.4) and by taking a conditional second moment, given $X_1, \dots, X_{\mu_n}$, we obtain

$$\begin{aligned} \hat{J}_n(x; \vartheta_n) = {} & \tilde{J}_{\mu_n}(x; \vartheta_n) - \int \tilde{J}_{\mu_n}(y; \vartheta_n)\, dP(y) \\ & + (n - \mu_n)^{-1} \sum_{i=\mu_n+1}^{n} \Big\{ J(X_i; \vartheta_n, g) - \tilde{J}_{\mu_n}(X_i; \vartheta_n) \\ & \qquad + \int \tilde{J}_{\mu_n}(y; \vartheta_n)\, dP(y) \Big\} + o_P(n^{-1/2}) \\ = {} & \tilde{J}_{\mu_n}(x; \vartheta_n) - \int \tilde{J}_{\mu_n}(y; \vartheta_n)\, dP(y) + o_P(n^{-1/2}), \end{aligned} \tag{3.22}$$

where the $o_P(n^{-1/2})$ term does not depend on x. Consequently, $\{\hat{J}_n(\cdot;\cdot)\}$ satisfies both (1.4) and (1.5). □

Theorem 3.1 is also valid if the uniform integrability condition (3.4) is replaced by the weaker conditions

$$\limsup_{n\to\infty} \int |J(x;\vartheta_n,g)|^2 1_{\{|J(x;\vartheta_n,g)|\geq \varepsilon n^{1/2}\}}\, dP_{\vartheta_n,g}(x) = 0, \quad \text{for all } \varepsilon > 0, \tag{3.23}$$

$$\limsup_{n\to\infty} \int |J(x;\vartheta_n,g)|^2\, dP_{\vartheta_n,g}(x) < \infty. \tag{3.24}$$

Since the right-hand side of (3.8) can be estimated by

$$|J(x;\vartheta_n,g)|\left[\mathbf{1}_{\{|J(x;\vartheta_n,g)|\geq \varepsilon n^{1/2}\}} + \varepsilon\right],$$

the left-hand side of (3.9) can be bounded with the help of (3.23) by ε^2 times the left-hand side of (3.24). Furthermore, (3.23) and (3.24) yield the natural (Lindeberg) condition in the argument immediately after (3.16).

Note that in Theorems 2.1 and 3.1 we have constructed the sequence $\{\hat{J}_n\}$ from the sequence $\{T_n\}$ and vice versa, and *not* $\hat{J}_n$ from T_n and T_n from $\hat{J}_n$ for n fixed. A finite sample inequality, which exhibits a relation between $\hat{J}_n$ and T_n, is given in Klaassen and van Zwet (1985) for situations in which, given ϑ, there exists a nontrivial sufficient statistic for g.

4. Discussion and examples. Theorem 3.1 shows in particular that for locally asymptotically linear adaptive estimators to exist, there should be estimators of the efficient influence function consistent in the sense of (1.4) and (1.5). For simplicity take $d = 1$ and consider a parametric submodel with densities $f(\cdot;\vartheta,\eta)$, $g = \eta \in \mathbf{R}$, with respect to some σ-finite measure μ. Arguing heuristically, we see that (1.5) means that there should exist an estimator $\hat{\eta}_n$ of η satisfying

$$n^{1/2}\int\left[\frac{\partial}{\partial\vartheta}\log f(x;\vartheta,\hat{\eta}_n)\right] f(x;\vartheta,\eta)\, d\mu(x) = o_P(1). \tag{4.1}$$

If $f(\cdot;\vartheta,\eta)$ is smooth in η, this becomes

$$n^{1/2}(\hat{\eta}_n - \eta)\int\left[\frac{\partial}{\partial\vartheta}\log f(x;\vartheta,\eta)\right]\left[\frac{\partial}{\partial\eta}\log f(x;\vartheta,\eta)\right] f(x;\vartheta,\eta)\, d\mu(x) = o_P(1). \tag{4.2}$$

Since typically $n^{1/2}(\hat{\eta}_n - \eta) \neq o_P(1)$ in smooth models, this shows that the last factor of the left-hand side of (4.2) has to vanish. But this is exactly Stein's necessary condition for adaptive estimation [see Stein (1956) and (3.1) of Bickel (1982)].

However, also if $f(\cdot;\vartheta,\eta)$ is not smooth in η, (4.1) might hold. As a simple example consider

$$f(x;\vartheta,\eta) = \vartheta^{-1}e^{-\vartheta^{-1}(x-\eta)}\mathbf{1}_{(\eta,\infty)}(x), \qquad \vartheta > 0.$$

Here, the left-hand side of (4.1) equals

$$n^{1/2}\vartheta^{-2}\int_{\eta}^{\infty}(x-\hat{\eta}_n-\vartheta)\mathbf{1}_{(\hat{\eta}_n,\infty)}(x)\vartheta^{-1}e^{-\vartheta^{-1}(x-\eta)}\,dx = n^{1/2}\vartheta^{-2}(\eta-\hat{\eta}_n)\vee 0$$

and, with $\hat{\eta}_n$ the smallest of $X_1,\dots,X_n$, (4.1) is satisfied.

Usually, in situations in which the density $f(\cdot;\vartheta,\eta)$ is not smooth in η, estimating η is less difficult than in the smooth case and there exist estimators $\hat{\eta}_n$ with $n^{1/2}(\hat{\eta}_n-\eta)=o_P(1)$. Then, approximating $(\partial/\partial\vartheta)\log f(\cdot;\vartheta,\eta)$ by a function $l_n(\cdot;\vartheta,\eta)$, which is differentiable in η with derivative bounded in absolute value by k_n, say ($k_n\to\infty$), (4.1) becomes

$$n^{1/2}\int l_n(x;\vartheta,\hat{\eta}_n)f(x;\vartheta,\eta)\,d\mu(x)=o_P(1), \tag{4.3}$$

which yields the condition

$$n^{1/2}k_n(\hat{\eta}_n-\eta)=o_P(1), \tag{4.4}$$

provided $n^{1/2}\int l_n(x;\vartheta,\eta)f(x;\vartheta,\eta)\,d\mu(x)=o(1)$.

A semiparametric estimation problem in which (1.5) is satisfied because of the phenomenon (4.4) and for which, in fact, adaptive estimators exist, is the following [cf. (4.18)]. Fix $\alpha>\frac{1}{2}$. Let $Y_1,\dots,Y_n$ be i.i.d. one-dimensional random variables with distribution function H and density h, which is symmetric about 0, has finite Fisher information $I(h)=\int(h'/h)^2h$ and has heavy, but not too heavy, tails in the sense that

$$\lim_{z\to\infty} zH(-\alpha\log z)=\infty,\qquad \lim_{z\to\infty} zH(-z)=0 \tag{4.5}$$

hold. Loosely speaking, H has heavier tails than Laplace with density $\frac{1}{2}\alpha^{-1}e^{-|\cdot|/\alpha}$ and less heavy tails than Cauchy. The observations $X_1,\dots,X_n$ are structurally defined by

$$X_i=\eta+e^{Y_i+\vartheta},\qquad i=1,\dots,n,\ \vartheta,\eta\in\mathbf{R}, \tag{4.6}$$

and one is interested in estimating ϑ with $g=(\eta,h)$ as a nuisance parameter. Note that η is identifiable via the support of X_1 and that ϑ is the point of symmetry of the distribution of $\log(X_1-\eta)$. Applying both Theorems 2.1 and 3.1 we shall prove

Theorem 4.1. *In model* (4.6) *adaptive estimation of* ϑ *is possible, i.e., there exists an estimator sequence* $\{T_n\}$ *of* ϑ *satisfying* (1.3) *for the efficient influence function*

$$J(\cdot;\vartheta,g)=-I^{-1}(h)h'/h(\log(\cdot-\eta)-\vartheta)\mathbf{1}_{(\eta,\infty)}(\cdot). \tag{4.7}$$

Proof. Denote the order statistics of $X_1,\dots,X_n$ by $X_{(1)}\le\cdots\le X_{(n)}$ and let $\psi\colon\mathbf{R}\to\mathbf{R}$ be an odd strictly increasing and bounded function with a bounded derivative. Define $\tilde{T}_n$ as the unique solution of

$$(4.8)\qquad n^{-1/2}\sum_{i=[n/3]}^{[2n/3]}\psi\big(\log(X_{(i)}-X_{(1)})-\theta\big)=0.$$

Because of

$$(4.9)\qquad X_{(1)}-\eta=o_P(n^{-\alpha}),$$

it can be shown that $\tilde{T}_n$ satisfies (2.3). The finiteness of $I(h)$ implies local asymptotic normality [see, e.g., Hájek (1972)] and, hence, contiguity of the probability measures induced by $(Y_1+\vartheta,\dots,Y_n+\vartheta)$ and $(Y_1+\vartheta_n,\dots,Y_n+\vartheta_n)$, respectively. Consequently, also the contiguity (2.1) holds and (2.2) may be deduced from contiguity and the existence of translation equivariant adaptive estimators satisfying (1.3) for the symmetric location problem [cf. (4.7) and the discussion after (2.2)].

In view of all this and Theorem 2.1, it suffices to show the existence of an estimator of the influence function (4.7) satisfying (1.4) and (1.5). Let η_n satisfy $\eta_n\ge\eta$, $\eta_n-\eta=o(n^{-\alpha})$. Denote the total variation distance between the distributions of $\log(X_1-\eta)$ and $\log|X_1-\eta_n|$, both under (ϑ_n,η,h), by $d(\eta,\eta_n)$. With $a_n=-\alpha\log(n/\log n)$, (4.5) yields

$$\begin{aligned}(4.10)\qquad d(\eta,\eta_n)&\le P(\log(X_1-\eta)\le a_n)+P(\log|X_1-\eta_n|\le a_n)\\&\quad+\int_{a_n}^{\infty}\Big|h(z-\vartheta_n)-(e^z+\eta_n-\eta)^{-1}e^zh\big(\log(e^z+\eta_n-\eta)-\vartheta_n\big)\Big|\,dz\\&\le 2H\big(\log(e^{a_n}+\eta_n-\eta)-\vartheta_n\big)\\&\quad+\int_{a_n}^{\infty}\int_0^{\eta_n-\eta}\big|(h-h')(\log(e^z+\zeta)-\vartheta_n)\big|(e^z+\zeta)^{-2}e^z\,d\zeta\,dz\\&\le o(1/\log n)+\int_0^{\eta_n-\eta}\int_{\log(e^{a_n}+\zeta)-\vartheta_n}^{\infty}|h(y)-h'(y)|e^{-y-\vartheta_n}\,dy\,d\zeta\\&\le o(1/\log n)+(\eta_n-\eta)e^{-a_n}(1+I^{1/2}(h))\\&=o\big((\log n)^{-1/2}\big).\end{aligned}$$

Let $m=m_n=[\log\log n]$. Again, from the existence of translation equivariant adaptive estimators satisfying (1.3) for the symmetric location problem, it follows by Theorem 3.1 that for η known there exists an estimator sequence $\{\tilde{J}_n(z;\theta)\}$ of $J_s(z;\theta,h)=-I^{-1}(h)h'/h(z-\theta)$ based on $\log(X_i-\eta)$, $i=1,\dots,m_n$, satisfying

$$(4.11)\qquad \int\big\{\tilde{J}_n(z;\vartheta_n)-J_s(z;\vartheta_n,h)\big\}^2h(z-\vartheta_n)\,dz=o_P(1).$$

Since η is unknown, we estimate it by $\hat{\eta}_n$, the smallest of $X_{m+1},\dots,X_n$ and we let $\tilde{J}_n(\cdot;\cdot)$ be based on $\log|X_i-\hat{\eta}_n|$, $i=1,\dots,m_n$. In view of the choice of m_n, (4.10) and the analogue of (4.9) for $\hat{\eta}_n$, a contiguity argument shows that

$\{\tilde{J}_n(\cdot;\cdot)\}$ still satisfies (4.11). Note that $\tilde{J}_n(\cdot;\cdot)$ is based on the complete sample $X_1,\ldots,X_n$ now.

Replacing $\tilde{J}_n$ by $(\tilde{J}_n \wedge m_n^{1/2}) \vee (-m_n^{1/2})$, we see that we may assume without loss of generality,

$$\sup_z \left|\tilde{J}_n(z;\vartheta_n)\right| = O_P\left(m_n^{1/2}\right). \tag{4.12}$$

Let k be the logistic density. By (4.12), (4.11), the continuity of translations in $\mathbf{L}_2$ and the finiteness of $I(h)$ and $\int u^2 k(u)\,du$, we have

$$\begin{aligned}
&\int\left\{\int \tilde{J}_n(z-m^{-1}u;\vartheta_n)k(u)\,du - J_s(z;\vartheta_n,h)\right\}^2 h(z-\vartheta_n)\,dz\\
&\quad\le \iint\Big\{\tilde{J}_n(z-m^{-1}u;\vartheta_n)\left[h^{1/2}(z-\vartheta_n)-h^{1/2}(z-\vartheta_n-m^{-1}u)\right]\\
&\qquad\quad+\left[\tilde{J}_n(z-m^{-1}u;\vartheta_n)-J_s(z-m^{-1}u;\vartheta_n,h)\right]h^{1/2}(z-\vartheta_n-m^{-1}u)\\
&\qquad\quad+\left[J_s(z-m^{-1}u;\vartheta_n,h)h^{1/2}(z-\vartheta_n-m^{-1}u)\right.\\
&\qquad\qquad\left.-J_s(z;\vartheta_n,h)h^{1/2}(z-\vartheta_n)\right]\Big\}^2 k(u)\,du\,dz \qquad (4.13)\\
&\quad\le m\iint\left[\int_0^{m^{-1}u}\tfrac{1}{2}h'h^{-1/2}(z-\zeta)\,d\zeta\right]^2 k(u)\,du\,dz + o_p(1)\\
&\quad= O(m^{-1}) + o_p(1).
\end{aligned}$$

Consequently, replacing $\tilde{J}_n(\cdot;\cdot)$ by $\int\tilde{J}_n(\cdot-m^{-1}u;\cdot)k(u)\,du$, we see that without loss of generality we may even assume

$$\sup_z\left|\frac{\partial^r}{\partial z^r}\tilde{J}_n(z;\vartheta_n)\right| = O_p\left(m_n^{r+1/2}\right), \qquad r=0,1 \tag{4.14}$$

[see also the proof of Lemma 2.1 of Bickel and Klaassen (1986)]. Replacing $\tilde{J}_n(z;\theta)$ by $\frac{1}{2}\tilde{J}_n(z;\theta)-\frac{1}{2}\tilde{J}_n(2\theta-z;\theta)$, we see that we may also assume that $\tilde{J}_n(z;\theta)$ is odd in z about θ.

Let $\chi\colon \mathbf{R}\to[0,1]$ be symmetric about 0, have a nonnegative bounded derivative on $(-\infty,0)$, be equal to 1 on $[-1,1]$ and vanish outside $[-2,2]$. Let $c_n = 2/\log\log n$ and define

$$J_n^*(z;\theta) = \tilde{J}_n(z;\theta)\chi(c_n(z-\theta)). \tag{4.15}$$

Also, $J_n^*(z;\theta)$ is odd in z about θ and satisfies (4.11) and (4.14); note that

$$\int\left\{J_n^*(z;\vartheta_n)-\tilde{J}_n(z;\vartheta_n)\right\}^2 h(z-\vartheta_n)\,dz = O_p\left(mH(-c_n^{-1})\right) = o_p(1). \tag{4.16}$$

Our estimate of $J(\cdot;\cdot,g)$ in (4.7) will be

$$\hat{J}_n(x;\theta) = \hat{J}_n(x;\theta;X_1,\ldots,X_n) = J_n^*\left(\log(x-X_{(1)});\theta\right)\mathbf{1}_{(X_{(1)},\infty)}(x). \tag{4.17}$$

Indeed, (4.17), (4.14), the symmetry of h, the properties of χ and $\mathscr{J}_n^*$ and (4.9) yield

$$n^{1/2}\left|\int \hat{\mathscr{J}}_n(x;\vartheta_n)\,dP_{\vartheta_n,g}(x)\right|$$

$$= n^{1/2}\left|\int \hat{\mathscr{J}}_n(\eta+e^z;\vartheta_n)h(z-\vartheta_n)\,dz\right|$$

$$\le n^{1/2}\left|\int \mathscr{J}_n^*(\log|e^z+\eta-X_{(1)}|;\vartheta_n)h(z-\vartheta_n)\,dz\right| + n^{1/2}O_p(m^{1/2})$$

$$\times\int_{-\infty}^{\log(X_{(1)}-\eta)}\chi\big(c_n(\log|e^z+\eta-X_{(1)}|-\vartheta_n)\big)h(z-\vartheta_n)\,dz$$

$$\le n^{1/2}\int\left|\int_0^{X_{(1)}-\eta}\left\{\frac{\partial}{\partial u}\mathscr{J}_n^*(u;\vartheta_n)\right\}\right. \tag{4.18}$$

$$\left.\times\mathbf{1}_{[0,2]}(c_n|u-\vartheta_n|)\Big|_{u=\log|e^z-\zeta|}(e^z-\zeta)^{-1}\,d\zeta\right|h(z-\vartheta_n)\,dz$$

$$+n^{1/2}O_p(m^{1/2})\chi\big(c_n(\log(X_{(1)}-\eta)-\vartheta_n)\big)$$

$$\le n^{1/2}O_p(m^{3/2})\int\int_0^{X_{(1)}-\eta}e^{-\vartheta_n+2/c_n}\,d\zeta\,h(z-\vartheta_n)\,dz+o_p(1)$$

$$=o_p(1),$$

i.e., $\hat{\mathscr{J}}_n(\cdot;\cdot)$ satisfies (1.5). Furthermore, (4.17), (4.11), (4.14), the properties of χ and (4.9) imply

$$\int\left\{\hat{\mathscr{J}}_n(x;\vartheta_n)-\mathscr{J}(x;\vartheta_n,g)\right\}^2 dP_{\vartheta_n,g}(x)$$

$$=\int_{\log(X_{(1)}-\eta)}^{\infty}\left\{\mathscr{J}_n^*(\log(e^z+\eta-X_{(1)});\vartheta_n)\right.$$

$$\left.-\mathscr{J}_s(z;\vartheta_n,h)\right\}^2h(z-\vartheta_n)\,dz$$

$$+\int_{-\infty}^{\log(X_{(1)}-\eta)-\vartheta_n}I^{-2}(h)(h')^2/h(y)\,dy$$

$$=\int_{\log(X_{(1)}-\eta)}^{\infty}\left\{\mathscr{J}_n^*(\log(e^z+\eta-X_{(1)});\vartheta_n)\right. \tag{4.19}$$

$$\left.-\mathscr{J}_n^*(z;\vartheta_n)\right\}^2h(z-\vartheta_n)\,dz+o_p(1)$$

$$\le\int_{\log(X_{(1)}-\eta)}^{\infty}\left\{\int_0^{X_{(1)}-\eta}O_p(m^{3/2})e^{-\vartheta_n+2/c_n}\,d\zeta\right\}^2$$

$$\times h(z-\vartheta_n)\,dz+o_p(1)$$

$$=o_p(1)$$

and, hence, (1.4). □

Let us consider now locally asymptotically linear estimators, which are not necessarily efficient (or adaptive). Arguing heuristically as in the beginning of this section, we see that (1.5) means that in smooth parametric submodels the following analogue of Stein's condition should hold:

$$\int J(x; \vartheta, \eta)\left[\frac{\partial}{\partial \eta}\log f(x; \vartheta, \eta)\right] f(x; \vartheta, \eta)\, d\mu(x) = 0. \tag{4.20}$$

In nonsmooth parametric submodels, it should be possible to estimate η consistently at a rate faster than $n^{1/2}$. As an example of (4.20), consider $X_1, \dots, X_n$ i.i.d. with density $g(\cdot - \vartheta)$ on $\mathbf{R}$; the unknown g has a unique median 0 with $g(0) > 0$. The influence function $J(x; \vartheta, g) = (2g(0))^{-1}\mathrm{sgn}(x - \vartheta)$ obviously satisfies (4.20). Estimating $1/g(0)$ by $\frac{1}{2}n^{1/2}\{X_{([n/2+n^{1/2}])} - X_{([n/2-n^{1/2}])}\}$, $X_{(1)} \le \cdots \le X_{(n)}$, the order statistics, we see that there exists an estimator $\hat{J}_n$ of J satisfying (1.4) and (1.5). Hence, it should be possible to apply Section 2. However, in this case the sample median is the natural locally asymptotically linear estimator.

In the symmetric location problem,

$$X_1, \dots, X_n \quad \text{i.i.d. density } h(\cdot - \vartheta),\ \vartheta \in \mathbf{R},\ h \text{ symmetric about } 0, \tag{4.21}$$

the M-estimator $\hat{\theta}_n$ defined by $\sum_{i=1}^n \psi(X_i - \hat{\theta}_n) = 0$, ψ odd and sufficiently smooth with derivative ψ', has influence function $(\int \psi' h)^{-1}\psi(x - \vartheta)$ [see (3.2.14) of Huber (1981)] and is asymptotically linear in this influence function [see Section 6.3 of Huber (1981)]. In this situation too, (4.20) is satisfied.

Of course, also the phenomenon (4.4) is not restricted to efficient estimation as the following generalization of Theorem 4.1 shows.

THEOREM 4.2. *Let $\mathscr{H} = \{h | h$ symmetric density with distribution function H satisfying* (4.5)$\}$. *If in the symmetric location model* (4.21) *with $h \in \mathscr{H}_0 \subset \mathscr{H}$, there exists an estimator sequence of ϑ satisfying* (1.3) *for the influence function $J(z; \vartheta, h) = J_0(z - \vartheta; h)$, which is odd in z about ϑ, then in model* (4.6) *with nuisance parameter $g = (\eta, h)$, $\eta \in \mathbf{R}$, $h \in \mathscr{H}_0$, there exists an estimator sequence of ϑ satisfying* (1.3) *for the influence function*

$$J(\cdot; \vartheta, g) = J_0(\log(\cdot - \eta) - \vartheta; h)\mathbf{1}_{(\eta, \infty)}(\cdot).$$

PROOF. Same as that of Theorem 4.1. □

Under (1.1), (1.2), (2.1), (3.23) and (3.24) it can be shown that (1.3) implies the possibility of constructing confidence sets for ϑ, based on T_n, which are asymptotically of a prescribed level. A proof of this statement can be based on Theorem 3.1. However, it is omitted since it should be possible to make the bootstrap work here by an appropriate strengthening of (1.3).

Acknowledgments. I would like to thank P. J. Bickel, A. Schick, A. W. van der Vaart and W. R. van Zwet for some useful discussions and the referees for suggesting improvements in the presentation.

REFERENCES

BERAN, R. (1978). An efficient and robust adaptive estimator of location. *Ann. Statist.* **6** 292–313.

BICKEL, P. J. (1982). On adaptive estimation. *Ann. Statist.* **10** 647–671.

BICKEL, P. J. and KLAASSEN, C. A. J. (1986). Empirical Bayes estimation in functional and structural models, and uniformly adaptive estimation of location. *Adv. in Appl. Math.* **7** 55–69.

HÁJEK, J. (1972). Local asymptotic minimax and admissibility in estimation. *Proc. Sixth Berkeley Symp. Math. Statist. Probab.* **1** 175–194. Univ. California Press.

HUBER, P. J. (1981). *Robust Statistics*. Wiley, New York.

KLAASSEN, C. A. J. and VAN ZWET, W. R. (1985). On estimating a parameter and its score function. In *Proc. of the Berkeley Conference in Honor of Jerzy Neyman and Jack Kiefer* (L. Le Cam and R. A. Olshen, eds.) **2** 827–839. Wadsworth, Monterey, Calif.

SCHICK, A. (1986). On asymptotically efficient estimation in semiparametric models. *Ann. Statist.* **14** 1139–1151.

SCHICK, A. (1987). A note on the construction of asymptotically linear estimators. *J. Statist. Plann. Inference* **16** 89–105.

STEIN, C. (1956). Efficient nonparametric testing and estimation. *Proc. Third Berkeley Symp. Math. Statist. Probab.* **1** 187–195. Univ. California Press.

STONE, C. J. (1975). Adaptive maximum likelihood estimators of a location parameter. *Ann. Statist.* **3** 267–284.

VAN EEDEN, C. (1970). Efficiency-robust estimation of location. *Ann. Math. Statist.* **41** 172–181.

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.

DEPARTMENT OF MATHEMATICS
UNIVERSITY OF LEIDEN
P.O. BOX 9512
2300 RA LEIDEN
NETHERLANDS