Journal of Statistical Planning and Inference 23 (1989) 163-179
North-Holland

CONSISTENCY FOR LEAST SQUARES REGRESSION ESTIMATORS WITH INFINITE VARIANCE DATA

Daren B.H. CLINE

Department of Statistics, Texas A&M University, College Station, TX 77843-3143, and Department of Statistics, Center for Stochastic Processes, University of North Carolina, Chapel Hill, NC 27514, U.S.A.

Received 12 November 1987; revised manuscript received 7 June 1988
Recommended by M. Rosenblatt

Abstract: The least squares estimators are discussed for the linear regression model with random predictors. Both predictors and errors may have infinite variance. Under the condition that the predictors are in a stable domain of attraction, we determine necessary and sufficient conditions for weak consistency of the least squares estimators in the simple linear model. The conditions vary, depending on whether the intercept parameter is included in the model. We also give sufficient conditions for consistency in a multiple regression setting.

AMS Subject Classification: 60F05, 62J05.

Key words: Infinite variance regression; consistency; regular variation.

1. Introduction

We are concerned with asymptotic results for the least squares estimators of regression models when both errors and regressors may have infinite variance. In particular, we wish to obtain as close to necessary and sufficient conditions for weak consistency of the estimators, as is possible. To obtain such precise conditions, we will concentrate on the two simple models:

$$Y=\beta_1 X+Z \qquad \text{(Model I)}$$

and

$$Y=\beta_0+\beta_1 X+Z. \qquad \text{(Model II)}$$

However, we will also provide sufficient conditions for weak consistency of estimators for the multiple regression model:

$$Y=X'\beta+Z, \quad X,\beta\in\mathbb{R}^k. \qquad \text{(Model III)}$$

* Research supported by the Air Force Office of Scientific Research Grant Number F49620 85 C 0144, and partially supported by National Science Foundation Research Grant Number DMS-8711924.

Asymptotic distributions for the estimators will be discussed in a subsequent paper since the techniques are different for that theory.

Various authors have studied practical and theoretical aspects of this problem when either the errors or the regressors have infinite variance. Blattberg and Sargent (1971) and Smith (1973) offered initial work for models in which the errors followed stable laws. Press (1975) discusses general inference procedures for stable data. Andrews (1987a, b) provides a complete list of references for applications of infinite variance regression and time series models, particularly in economics.

Among those who have studied asymptotic results are Kanter and Steiger (1974) and Maller (1981). Kanter and Steiger limited their work to the special case where both X and Z have symmetric distributions with asymptotically Pareto tails of the same index. They showed consistency of the estimators for Model I (and actually get a bound on the convergence rate). Maller provided a general result in case both X and Z are in a normal domain of attraction. Andrews (1987a, b) was concerned with sufficient, rather than necessary, conditions for strong, rather than weak, consistency. Carroll and Cline (1988) considered consistency and rates of convergence of weighted least squares estimators, when weights are determined by sample variances. They showed that slow rates and inconsistency can occur, even with normal data.

The limitation that errors and regressors have equivalent tails, which Kanter and Steiger imposed, was apparently due their greater interest in autoregressive processes. Others that have investigated least squares estimation in autoregressive processes include Hannan and Kanter (1977), Yohai and Maronna (1977) and Davis and Resnick (1985a, b, 1986). Indeed, when the errors have regularly varying tail probabilities, such processes must have probability tails similar to the errors (Cline, 1983). In our case, however, we do not wish to suppose that regressors and errors have similar tails, or even that they are in similar domains of attraction.

Like Kanter and Steiger, Maller, and Andrews, our motivation is to avoid specific distributional assumptions. Ideally, we would like to dispence with all assumptions. But with only an assumption on the distribution tails of X, we will determine the precise requirements on the errors to obtain consistency. Of course, the adverse effect of heavy-tailed errors in regression is well documented and numerous alternatives to least squares have been proposed (e.g. Huber, 1981, Maronna and Yohai, 1981, and Bierens, 1981). However, we are here less interested in efficiency than in establishing the scope of least squares estimation. Stochastic regressors with infinite variance actually moderate the effect of large errors. Thus, as Kanter and Steiger, Maller, and Andrews have shown, consistency is possible if the regressors have heavy enough tails.

Our focus, then, is to describe how consistency depends on the exact relationship between regressors' and errors' distributions, and thereby to determine necessary conditions for consistency in least squares regression. The precise relationship turns out to be relatively simple to describe. Besides extending the mathematical understanding, we hope this work will also provide practical insight into the range of situations in which least squares is at least appropriate.

Before stating our theorems, we first set forth our notation. After the theorem statements, in the second section, we will provide several preliminary results and in the third section we prove the theorems.

Consider Models I and II, where X is real valued. We will assume $X \sim F$ and $Z \sim G$, independently. The pairs (X_j, Z_j) will be independent copies of (X, Z). The distribution of XZ will be denoted with H. Let F_1, G_1 and H_1 be the distributions of $|X|$, $|Z|$ and $|XZ|$, respectively. Define the truncated moment functions

$$\mu(t) = \int_{-t}^{t} uF(\mathrm{d}u),$$

$$\mu_1(t) = \int_{-t}^{t} |u|F(\mathrm{d}u) \quad \text{and} \quad \mu_2(t) = \int_{-t}^{t} u^2F(\mathrm{d}u).$$

The functions ν, ν_1 and ν_2 are defined similarly for G, as are λ, λ_1 and λ_2 for H.

The class of functions which vary regularly with index ϱ will be denoted RV_ϱ.

Generally speaking, consistency in Model I depends on an asymptotic relationship between the tails of Z and X^2. This is stated in Theorem 1.1.

Theorem 1.1. *Suppose $\mu_2 \in \mathrm{RV}_{2-\alpha}$ for some $\alpha > 0$. Then the least squares estimator $\hat\beta_{1,n}$ for Model* I *is weakly consistent if and only if each of the following hold*:

(i) $\lim_{t\to\infty} t(1-G_1(t))/\mu_2(t^{1/2}) = 0$, *and*

(ii) $\lim_{t\to\infty} \lambda(t)/\mu_2(t^{1/2}) = 0$ *whenever* $\alpha = 2$.

In particular, if $\nu_1 \in \mathrm{RV}_{1-\gamma}$, consistency occurs if $\gamma > \frac{1}{2}\alpha$ and only if $\gamma \geq \frac{1}{2}\alpha$.

Aside from centering considerations in special cases, the precise condition for consistency is roughly that the probability tails for Z are dominated by those for X^2. That this should be the case can be argued heuristically as follows. The estimator $\hat\beta_{1,n}$ for Model I satisfies

$$\hat\beta_{1,n} - \beta_1 = \frac{\sum_{j=1}^{n} X_j Z_j}{\sum_{j=1}^{n} X_j^2},$$

where (X_j, Z_j) are independent copies of (X, Z). Letting H be the distribution of XZ and F_2 be the distribution of X^2, we see that $\hat\beta_{1,n} - \beta_1 \to 0$ when the tails of F_2 dominate those of H. Since the tails of H dominate both those of F and those of G, and the tails of F_2 dominate those of F, then it is at least sufficient that the tails of F_2 dominate those of G. This argument, of course, will be made precise. It relies on the fact that, for $a_n^2 = \inf\{t\colon n\mu_2(t^{1/2}) \leq t\}$, $a_n^{-2} \sum_{j=1}^{n} X_j^2$ is stochastically bounded.

The conditions in Theorem 1.1 are necessary under the assumption $\mu_2 \in \mathrm{RV}_{2-\alpha}$, $\alpha > 0$. However, one might wonder whether the regular variation of μ_2 is required. Under strict dominated variation requirements, similar arguments could be valid. On the other hand, the general condition may not be so nicely expressed in terms of a comparison between G_1 and μ_2. Also, if $\mu_2 \in \mathrm{RV}_2$ then $1 - F_1$ is slowly varying. In this case $a_n^{-2} \sum_{j=1}^{n} X_j^2$ is unbounded and it is sufficient that $a_n^{-2} \sum_{j=1}^{n} X_j Z_j$ be stochastically bounded.

The consistency problem for Model II differs from Model I in that the location parameter β_0 is of overriding importance. Indeed, it must be meaningful, at least in the sense of the ordinary weak law of large numbers.

Theorem 1.2. *Assume $\mu_2 \in \mathrm{RV}_{2-\alpha}$. The least squares estimators $(\hat\beta_{0,n}, \hat\beta_{1,n})$ for Model* II *are weakly consistent if $\overline{Z}_n \to 0$ in probability. Furthermore, this condition is necessary if one of the following hold.*

(i) $\alpha > 1$,

(ii) $E|Z|^\gamma < \infty$ *for some* $\gamma > \alpha/(\alpha+1)$, *or*

(iii) $v_1 \in \mathrm{RV}_{1-\gamma}$ *for some* $\gamma > 0$.

(*Note that* $\bar{Z}_n \xrightarrow{\mathrm{P}} 0$ *implies* $v_1 \in \mathrm{RV}_0$.)

Theorem 1.2 shows that the regressor with the lightest tails, in this case the constant 1, dictates the necessary behavior of the error probability tails. This concept carries over to multiple regression. On the other hand, $\hat\beta_{1,n}$ can be consistent even when $\hat\beta_{0,n}$ is not. The precise condition is exactly that for Model I.

We turn now the problem of consistency for multiple regression. Suppose that $X, X_1, X_2, \ldots$ are independent and have multivariate distribution F ($X \in \mathbb{R}^k$) and that $Z, Z_1, Z_2, \ldots$ are independent with distribution G. As before, the two sequences are independent. For $\beta \in \mathbb{R}^k$, the usual multiple regression model is

$$Y_j = X_j'\beta + Z_j. \qquad \text{(Model III)}$$

To give general conditions we first define the marginal distributions F_i (for X_{ij}) and H_i (for $X_{ij}Z_j$). Let

$$\mu_{i2}(t) = \int_{-t}^{t} u^2 F_i(\mathrm{d}u) \quad \text{and} \quad \lambda_i(t) = \int_{-t}^{t} u H_i(\mathrm{d}u).$$

Taking our cue from Theorems 1.1 and 1.2, we have the following.

Theorem 1.3. *Suppose $\mu_{i2} \in \mathrm{RV}_{2-\alpha_i}, \alpha_i > 0$, for each $i = 1, 2, \ldots, k$. Define $a_{in} = \inf\{t\colon n\mu_{i2}(t) \le t^2\}$. Let Q_n be the matrix*

$$Q_n = \left(\frac{1}{a_{in}a_{hn}} \sum_{j=1}^{n} X_{ij}X_{hj}\right)_{i,h=1}^{k}.$$

Define also $\mu_2^*(t) = \min_i(\mu_{i2}(t))$. The least squares estimator $\hat\beta_n$ for Model III *is consistent if each the following hold:*

(i) *The sequence $\{Q_n^{-1}\}$ is stochastically bounded.*

(ii) $\lim_{t\to\infty} t(1 - G_1(t))/\mu_2^*(t^{1/2}) = 0$

(iii) $\lim_{t\to\infty} \lambda_i(t)/\mu_2^*(t^{1/2}) = 0$ *whenever* $\alpha_i = 2$.

In particular, if (i) *holds, consistency occurs in any of the following cases:*

(a) $v_1 \in \mathrm{RV}_{1-\gamma}$, $\gamma > \frac{1}{2}\alpha^*$, *where* $\alpha^* = \max_i(\alpha_i)$.

(b) *Model* III *includes an intercept and* $\bar{Z}_n \xrightarrow{\mathrm{P}} 0$ *in probability.*

(c) *For some* i, $\alpha_i > \max_{h\ne i}(\alpha_h)$ *and* $a_{in}^{-2} \sum_{j=1}^{n} X_{ij}Z_j \xrightarrow{\mathrm{P}} 0$.

The necessity of conditions (ii) and (iii) in Theorem 1.3 is also true, we surmise, at least when $v_1 \in \mathrm{RV}_{1-\gamma}$, $\gamma > 0$. Proof of this would involve some knowledge of the joint distribution for $((X'X)^{-1}, X'Z)$. And that will require techniques using multivariable regular variation. Such techniques would also illustrate when condition (i) holds. One situation when (i) holds, however, is the following.

Lemma 1.4. *Let μ_{i2}, a_{in} and Q_n be as in Theorem 1.3. The sequence $\{Q_n^{-1}\}$ is stochastically bounded if*

$$\lim_{n\to\infty} \frac{E[|X_{ij}X_{hj}| 1_{|X_{ij}X_{hj}| \le a_{in}a_{hn}}]}{a_{in}a_{hn}} = 0,$$

whenever $EX_{ij}^2 = \infty$ or $EX_{hj}^2 = \infty$.

2. Preliminaries

This section provides preliminary lemmas; the next section will investigate consistency for estimators in Models I and II. The most important results in this section are the known results, Lemmas 2.2 and 2.3, and the new result Corollary 2.7. The remainder help establish Corollary 2.7, but are interesting in their own right as parallels to familiar results for regularly varying functions.

First, we make note of a basic property of regularly varying functions. If $s \in \mathrm{RV}_\varrho$, then for any $\varepsilon > 0$ and $K > 1$, there exists t_0 such that

$$\frac{1}{K} u^{\varrho - \varepsilon} \le \frac{s(tu)}{s(t)} \le K u^{\varrho + \varepsilon} \quad \text{for all } u \ge 1,\ t \ge t_0. \tag{2.1}$$

This well known result (Potter's Theorem, cf. Bingham, Goldie and Teugels, 1987, p. 25) is primarily used for uniform and/or dominated convergence arguments. It can be strengthened to the following lemma, enhancing its use and allowing one to avoid arguments which would otherwise be piecemeal.

Lemma 2.1. *Assume $s(t) \in \mathrm{RV}_\varrho$.*

(i) *Suppose $s(t)t^{-\delta}$ is bounded on $[0, 1]$ for some real δ. Then for each $\varepsilon > 0$ and $K > 1$ there exists t_0 such that*

$$\sup_{t \ge t_0} \frac{s(ut)}{s(t)} \le K \max(u^{\varrho+\varepsilon}, u^{\varrho-\varepsilon}, u^\delta 1_{u \le \varepsilon}) \quad \textit{for all } u > 0.$$

(ii) *Suppose $s(t)t^{-\delta}$ is bounded away from 0 on $[0, 1]$ for some real δ. Then for each $\varepsilon > 0$ and $K > 1$ there exists t_0 such that*

$$\inf_{t \ge t_0} \frac{s(ut)}{s(t)} \ge \frac{1}{K} \min(u^{\varrho+\varepsilon}, u^{\varrho-\varepsilon}, (u^{-\delta} 1_{u \le \varepsilon})^{-1}) \quad \textit{for all } u > 0.$$

Proof. (i) When $u \ge 1$ we can choose t_1 so that (2.1) holds, i.e.

$$\frac{1}{K}u^{\varrho-\varepsilon} \le \frac{s(ut)}{s(t)} \le Ku^{\varrho+\varepsilon} \quad \text{for all } u \ge 1,\ t \ge t_1. \tag{2.2}$$

The left inequality in (2.2) gives

$$\frac{s(ut)}{s(t)} = \frac{s(ut)}{s(ut(1/u))} \le Ku^{\varrho-\varepsilon} \quad \text{if } u \le 1,\ ut \ge t_1. \tag{2.3}$$

Now choose M_1 so that $s(t) \le M_1 t^{\varrho+\varepsilon}$ for all $t \ge 1$. Choose t_2 so that

$$\frac{s(t)}{t^{\varrho-\varepsilon}} \ge \frac{M_1}{K} t_1^{2\varepsilon} \quad \text{for all } t \ge t_2.$$

Thus, if $1 \le ut \le t_1$, $t \ge t_2$,

$$\begin{aligned} \frac{s(ut)}{s(t)} &\le M_1 (ut)^{\varrho+\varepsilon} \frac{Kt_1^{-2\varepsilon}}{M_1 t^{\varrho-\varepsilon}} \\ &= Ku^{\varrho+\varepsilon}\left(\frac{t}{t_1}\right)^{2\varepsilon} \le Ku^{\varrho-\varepsilon}. \end{aligned} \tag{2.4}$$

Now choose M_2 so that $s(t) \le M_2 t^{\delta}$ for all $t \le 1$. Choose $t_3 \ge 1/\varepsilon$ so that

$$\frac{s(t)}{t^{\varrho-\varepsilon}} \ge \frac{M_2}{K} \quad \text{for all } t \ge t_3.$$

Then if $ut \le 1$, $t \ge t_3$, it follows that $u \le \varepsilon$ and

$$\begin{aligned} \frac{s(ut)}{s(t)} &\le M_2 (ut)^{\delta} \frac{K}{M_2} t^{\varepsilon-\varrho} \\ &= Ku^{\delta} t^{\delta+\varepsilon-\varrho} \le K \max(u^{\delta}, u^{\varrho-\varepsilon}). \end{aligned} \tag{2.5}$$

The result follows from (2.2)–(2.5) by letting $t_0 = \max(t_1, t_2, t_3)$.

(ii) This follows by a similar argument. □

We start by recalling the general form for the weak law of large numbers.

Lemma 2.2 (Gnedenko and Kolmogorov, 1968, p. 134). *There exists a sequence* a_n *such that*

$$\frac{1}{a_n} \sum_{j=1}^{n} Z_j \xrightarrow{\mathrm{P}} z$$

if and only if each of the following hold for some monotone function $s(t)$:

$$\lim_{t\to\infty} s(t)(1 - G_1(t)) = 0, \tag{2.6}$$

$$\lim_{t\to\infty} \frac{s(t)v(t)}{t} = z \tag{2.7}$$

and

$$\lim_{t\to\infty} \frac{s(t)v_2(t)}{t^2} = 0. \tag{2.8}$$

In this case we may choose $a_n = \inf\{t : s(t) \ge n\}$. □

Note that if $(1-G_1) \in \mathrm{RV}_{-\gamma}$, $0<\gamma<2$, then

$$\frac{t^2(1-G_1(t))}{v_2(t)} \to \frac{2-\gamma}{\gamma},$$

rendering (2.6) and (2.8) equivalent. Another known result is the following.

Lemma 2.3 (Feller, 1971, pp. 236 and 448). *There exist a_n such that*

$$\frac{1}{a_n} \sum_{j=1}^{n} |X_j| \xrightarrow{\mathrm{w}} T,$$

where T is almost surely finite and positive, if and only if $\mu_1 \in \mathrm{RV}_{1-\alpha}$ for some $\alpha>0$. In this case we may choose $a_n = \inf\{t : n\mu_1(t) \le t\}$. Two possibilities may occur:

(i) $T=1$ *almost surely, if* $\alpha=1$.
(ii) T *has positive stable* (α) *law, if* $\alpha<1$. □

For the regression problem, we intend to apply Lemma 2.2 to the partial sums $\sum_{j=1}^{n} X_j Z_j$, that is, to H and its truncated moment functions. This could mean, for example, examining the tail behavior of H in terms of F and G as is done in Cline (1986). Actually, we will not require so much, although a few results are necessary.

Lemma 2.4. *Let $\alpha \ge 0$, $\gamma \ge 0$ and $\delta = \min(\alpha, \gamma)$.*

(i) *If* $(1-F_1) \in \mathrm{RV}_{-\alpha}$, $(1-G_1) \in \mathrm{RV}_{-\gamma}$, *then* $(1-H_1) \in \mathrm{RV}_{-\delta}$.
(ii) *If* $\mu_1 \in \mathrm{RV}_{1-\alpha}$, $v_1 \in \mathrm{RV}_{1-\gamma}$, *then* $\lambda_1 \in \mathrm{RV}_{1-\delta}$.
(iii) *If* $\mu_2 \in \mathrm{RV}_{2-\alpha}$, $v_2 \in \mathrm{RV}_{2-\gamma}$, *then* $\lambda_2 \in \mathrm{RV}_{2-\delta}$.

Proof. (i) When $\alpha>\gamma$, we have in fact

$$\lim_{t\to\infty} \frac{1-H_1(t)}{1-G_1(t)} = E|X|^{\gamma} < \infty, \tag{2.9}$$

since $E|X|^{\gamma+\varepsilon} < \infty$ for some $\varepsilon>0$. This is due to Breiman (1965, Proposition 3) for the case $\gamma=1$ (and hence for any $\gamma>0$). The proof is implicit in the proof of Lemma 2.5(i) below. The result is similar when $\gamma>\alpha$.

When $\gamma=\alpha$, the result holds by Embrechts and Goldie (1980, Corollary 3).

(ii) This is equivalent to (iii) by the transformation $x \to x^2$.

(iii) When $\gamma<2$, then $1-G_1(t) \in \mathrm{RV}_{-\gamma}$. So $1-H_1(t) \in \mathrm{RV}_{-\delta}$ for the same reasons as in (i). Thus $\lambda_2 \in \mathrm{RV}_{2-\delta}$. If $\gamma=\alpha=2$, the result is given by Maller (1981, Theorem 3). □

The next two results are completely new in that we do not require regular variation of the distribution tails, but rather of a dominating function.

Lemma 2.5. *Let s be a function in* RV_ϱ, $\varrho \geq 0$. *Assume F is not degenerate at zero and* $E|X|^\gamma < \infty$ *for some* $\gamma > \varrho$. *Then*

(i) $$\lim_{t\to\infty} s(t)(1-H_1(t)) = 0 \quad \Leftrightarrow \quad \lim_{t\to\infty} s(t)(1-G_1(t)) = 0.$$

(ii) $$\lim_{t\to\infty} \frac{s(t)\lambda_1(t)}{t} = 0 \quad \Leftrightarrow \quad \lim_{t\to\infty} \frac{s(t)\nu_1(t)}{t} = 0.$$

(iii) $$\lim_{t\to\infty} \frac{s(t)\lambda_2(t)}{t^2} = 0 \quad \Leftrightarrow \quad \lim_{t\to\infty} \frac{s(t)\nu_2(t)}{t^2} = 0.$$

(iv) $$\limsup_{t\to\infty} \left|\frac{s(t)\lambda(t)}{t}\right| \leq E|X|^\varrho \limsup_{t\to\infty} \left|\frac{s(t)\nu(t)}{t}\right|.$$

And, if $\lim_{t\to\infty} s(t)\nu(t)/t = c$, *then*

$$\lim_{t\to\infty} \frac{s(t)\lambda(t)}{t} = cE[\operatorname{sgn}(X)|X|^\varrho].$$

Proof. We prove only (i), since each of (ii)–(iv) may be proven in a similar manner.

(i) There is no loss in assuming $s(t)=1$ for all $t\leq 1$. We apply Lemma 2.1(ii) with $\varepsilon = \gamma - \varrho$ and $\delta = 0$. Then for some t_0,

$$\sup_{t\geq t_0} \frac{s(t)}{s(t/u)} \leq K\max(u^\gamma, 1) \quad \text{for all } u>0. \tag{2.10}$$

If $s(t)(1-G_1(t))\to c$, then Lebesgue convergence with (2.10) gives

$$\lim_{t\to\infty} s(t)(1-H_1(t)) = \int_0^\infty \lim_{t\to\infty}\left(\frac{s(t)}{s(t/u)}\, s(t/u)(1-G_1(t/u))\right)F_1(\mathrm{d}u)$$
$$= c\int_0^\infty u^\varrho F_1(\mathrm{d}u).$$

In particular, this holds with $c=0$.

Suppose instead that $s(t)(1-H_1(t))\to 0$. Choose $\delta>0$ so that $1-F_1(\delta)>0$. Then for every t,

$$(1-H_1(\delta t)) \geq \int_\delta^\infty (1-G_1(\delta t/u))F_1(\mathrm{d}u)$$
$$\geq (1-G_1(t))(1-F_1(\delta)).$$

Therefore,

$$\limsup_{t\to\infty} s(t)(1-G_1(t)) \leq \lim_{t\to\infty} \frac{s(t)}{s(\delta t)}\, \frac{s(\delta t)(1-H_1(\delta t))}{1-F_1(\delta)} = 0. \qquad \square$$

The various limits in (i), (ii) and (iii) of Lemma 2.5 will in fact imply each other, depending on the value of ϱ. This is explained in the next lemma.

Lemma 2.6. *Define for $t>0$,*

$$v_\delta(t)=\begin{cases}\displaystyle\int_0^t u^\delta G_1(\mathrm{d}u) & \text{if } \delta>0,\\ \displaystyle\int_t^\infty u^\delta G_1(\mathrm{d}u) & \text{if } \delta\le 0.\end{cases}$$

Suppose $s(t)\in \mathrm{RV}_\varrho$, $\varrho\ge 0$. If, for some real δ,

$$\lim_{t\to\infty} s(t)t^{-\delta}v_\delta(t)=c<\infty, \tag{2.11}$$

then $\delta\ge\varrho$ or $\delta\le 0$. Furthermore, if (2.11) *holds, then for any $\gamma>\varrho$ or $\gamma\le 0$ ($\gamma<0$, if $\varrho=0$),*

$$\lim_{t\to\infty} s(t)t^{-\gamma}v_\gamma(t)=c\left|\frac{\delta-\varrho}{\gamma-\varrho}\right|. \tag{2.12}$$

Remark. When $c>0$, then $v_\delta\in\mathrm{RV}_{\delta-\varrho}$ and this is the familiar result for regular variation (cf. Feller, 1971, p. 283). We will apply the lemma, however, with $c=0$.

Proof. Suppose $s(t)t^{-\delta}v_\delta(t)\to c<\infty$ for some δ. There is no loss in assuming $s(t)=1$ for all $t\le 1$. Note that for any δ, $t^{-\delta}\gamma_\delta(t)\to 0$. Thus we may as well assume s is bounded away from zero. Furthermore, if $\delta>0$ then v_δ is nondecreasing and $s(t)t^{-\delta}$ must be bounded. Thus either $\delta\ge\varrho$ or $\delta\le 0$.

Assume first that $\delta\ge\varrho$. Let $\gamma\le 0$, $\gamma<\varrho$. Note that

$$\begin{aligned}t^{-\gamma}v_\gamma(t)&=t^{-\gamma}\int_t^\infty u^{\gamma-\delta}v_\delta(\mathrm{d}u)\\ &=(\delta-\gamma)t^{-\gamma}\int_t^\infty u^{\gamma-\delta-1}v_\delta(u)\mathrm{d}u-t^{-\delta}v_\delta(t)\\ &=(\delta-\gamma)\int_1^\infty (x^{\gamma-1}(xt)^{-\delta}v_\delta(xt))\mathrm{d}x-t^{-\delta}v_\delta(t).\end{aligned}$$

We apply (2.1) with $0<\varepsilon<\varrho-\gamma$. There exists t_0 such that

$$\sup_{t\ge t_0}\frac{s(t)}{s(xt)}\le Kx^{\varepsilon-\varrho}\quad\text{for all } x\ge 1.$$

Since $x^{\gamma+\varepsilon-\varrho-1}$ is integrable on $[1,\infty)$, the Lebesgue convergence theorem yields

$$\begin{aligned}&\lim_{t\to\infty} s(t)t^{-\gamma}v_\gamma(t)\\ &\quad=(\delta-\gamma)\int_1^\infty \lim_{t\to\infty}\left(\frac{x^{\gamma-1}s(t)}{s(xt)}\, s(xt)(xt)^{-\delta}v_\delta(xt)\right)\mathrm{d}x-c\end{aligned}$$

$$= \left((\delta - \gamma) \int_1^\infty x^{\gamma-\varrho-1}\, dx - 1 \right) c$$

$$= \frac{\delta - \varrho}{\varrho - \gamma}\, c. \tag{2.13}$$

Now assume $\delta \le 0$. Let $\gamma > \varrho$. With $0 < \varepsilon < \gamma - \varrho$ and Lemma 2.1(ii), there is some t_0 such that

$$\sup_{t \ge t_0} \frac{s(t)}{s(xt)} \le K x^{-\varrho-\varepsilon} \quad \text{for all } x \le 1.$$

Again we use Lebesgue convergence since $x^{\gamma-\varrho-\varepsilon-1}$ is integrable on $[0, 1]$ and since $s(t)t^{-\delta} v_\delta(t)$ is bounded for all $t > 0$,

$$\lim_{t\to\infty} s(t)t^{-\gamma} v_\gamma(t)$$

$$= (\gamma - \delta) \int_0^1 \lim_{t\to\infty} \left(\frac{x^{\gamma-1} s(t)}{s(xt)}\, s(xt)(xt)^{-\delta} v_\delta(xt)\, dx \right) - c$$

$$= \left((\gamma - \delta) \int_0^1 x^{\gamma-\varrho-1}\, dx - 1 \right) c$$

$$= \frac{\varrho - \delta}{\gamma - \varrho}\, c. \tag{2.14}$$

Two cases remain to be resolved. First, assume again that $\delta \ge \varrho$, but choose $\gamma > 0$. By the first argument, (2.13) holds (with γ_1 replacing γ) for some $\gamma_1 \le 0$. By the second argument, which leads to (2.14),

$$\lim_{t\to\infty} s(t)t^{-\gamma} v_\gamma(t) = \frac{\varrho - \gamma_1}{\gamma - \varrho}\, \frac{\delta - \varrho}{\varrho - \gamma_1}\, c.$$

This resolves the first case. The second case, $\delta \le 0$, $\gamma \le 0$ (with $\gamma < 0$, if $\varrho = 0$) is likewise resolved using first (2.14), then (2.13). $\square$

The above lemmas lead to the following.

Corollary 2.7. *Suppose $s(t)$ is unbounded and in* RV_ϱ, $\varrho \ge 0$, *and define* $a_n = \inf\{t\colon s(t) \ge n\}$. *Assume* $E|X|^\gamma < \infty$ *for some* $\gamma > \varrho$ *and F is not degenerate at zero.*

(i) *If* $\varrho < 1$ *the following are equivalent:*

(a) $$\frac{1}{a_n} \sum_{j=1}^{n} X_j Z_j \xrightarrow{\mathrm{P}} 0.$$

(b) $$\frac{1}{a_n} \sum_{j=1}^{n} Z_j \xrightarrow{\mathrm{P}} 0.$$

(c) $$\lim_{t\to\infty} s(t)(1 - G_1(t)) = 0.$$

(d) $$\lim_{t\to\infty} s(t)(1 - H_1(t)) = 0.$$

(ii) *If* $\varrho=1$, (a) *is equivalent to* (c) *combined with*

(e) $$\lim_{t\to\infty}\frac{s(t)\lambda(t)}{t}=0.$$

Proof. (i) Statement (b) implies (c) by Lemma 2.2. By Lemma 2.6, (c) implies each of the three conditions (2.6)–(2.8). Thus, by Lemma 2.2, (c) implies (b). By Lemma 2.5(i), (c) is equivalent to (d). Statement (a) is equivalent to (d) for the same reason that (b) is equivalent to (c).

(ii) Again, (c) and (d) are equivalent by Lemma 2.5(i) and each implies

$$\lim_{t\to\infty}\frac{s(t)\lambda_2(t)}{t^2}=0.$$

Applying Lemma 2.2, (a) is thus equivalent to (c) with (e). □

Remark. In situation (i) of Corollary 2.7, one has in fact that (a) and (b) are each equivalent to each of

(a′) $$\frac{1}{a_n}\sum_{j=1}^{n}|X_jZ_j|\xrightarrow{\mathrm{P}}0,$$

and

(b′) $$\frac{1}{a_n}\sum_{j=1}^{n}|Z_j|\xrightarrow{\mathrm{P}}0.$$

This is not generally the case for situation (ii) of the corollary.

3. Consistency proofs

We begin by noting that the least squares estimator $\hat\beta_{1,n}$ for Model I satisfies

$$\hat\beta_{1,n}-\beta_1=\frac{\sum_{j=1}^n X_jZ_j}{\sum_{j=1}^n X_j^2}. \tag{3.1}$$

The least squares estimators $\hat\beta_{0,n}$ and $\hat\beta_{1,n}$ for Model II satisfy

$$\begin{aligned}\hat\beta_{0,n}-\beta_0&=\bar Z_n-(\hat\beta_{1,n}-\beta_1)\bar X_n\\&=\frac{\bar Z_n\sum_{j=1}^n X_j^2-\bar X_n\sum_{j=1}^n Z_jX_j}{\sum_{j=1}^n X_j^2-n\bar X_n^2}\end{aligned} \tag{3.2}$$

and

$$\hat\beta_{1,n}-\beta_1=\frac{\sum_{j=1}^n X_jZ_j-n\bar X_n\bar Z_n}{\sum_{j=1}^n X_j^2-n\bar X_n^2}. \tag{3.3}$$

We will first prove conditions for consistency for the estimator of Model I. Note that we make no assumptions of symmetry. Kanter and Steiger (1974) prove a special case.

Proof of Theorem 1.1. Let $s(t)=t/\mu_2(t^{1/2})$ and $a_n^2=\inf\{t\colon s(t)\ge n\}$. From Lemma 2.3, applied to the distribution of X^2,

$$\frac{1}{a_n^2}\sum_{j=1}^n X_j^2 \xrightarrow{\mathrm{w}} T,$$

where T is almost surely finite and positive. We see therefore that

$$\hat{\beta}_{1,n}-\beta_1 = \frac{\sum_{j=1}^n X_j Z_j}{\sum_{j=1}^n X_j^2} \xrightarrow{\mathrm{P}} 0,$$

if and only if $a_n^{-2}\sum_{j=1}^n X_j Z_j \xrightarrow{\mathrm{P}} 0$. But $s\in \mathrm{RV}_{\alpha/2}$ and $E|X|^\delta<\infty$ for $\frac{1}{2}\alpha<\delta<\alpha$. By Corollary 2.7, therefore, consistency is equivalent to

$$\lim_{t\to\infty} s(t)(1-G_1(t))=0 \tag{3.4}$$

and, if $\alpha=2$,

$$\lim_{t\to\infty}\frac{s(t)\lambda(t)}{t}=0. \tag{3.5}$$

(Note that (3.4) implies (3.5) if $\alpha<2$ by Lemma 2.6.) This is the equivalence to be proved.

Finally, if $v_1\in \mathrm{RV}_{1-\gamma}$ with $\gamma<\frac{1}{2}\alpha$ (so $\gamma<1$) then $1-G_1(t)\in \mathrm{RV}_{-\gamma}$. Since $s(t)\in \mathrm{RV}_{\alpha/2}$, (3.4) fails. On the other hand if $1\geq\gamma>\frac{1}{2}\alpha$, similar reasoning assures that (3.4) holds. □

Unfortunately, it is not possible to replace (3.5) with a version involving v instead of λ, when $\alpha=2$. One condition implying (3.5) is suggested by Lemma 2.4(iv):

$$\left(\lim_{t\to\infty}\frac{s(t)v(t)}{t}\right)EX=0,$$

where the limit is assumed to exist, finite.

Proof of Theorem 1.2. We will handle the three cases as follows. First, we will show necessity and sufficiency of $\bar{Z}_n \xrightarrow{\mathrm{P}} 0$ in case $\alpha>1$. Then we will show sufficiency when $\alpha\leq 1$. Finally, we will demonstrate necessity for cases (ii) and (iii).

Again define $a_n^2=\inf\{t\colon n\mu_2(t^{1/2})\leq t\}$. By Lemma 2.3, $a_n^{-2}\sum_{j=1}^n X_j^2$ converges weakly to a positive stable law or to 1.

(i) Assume $\alpha>1$. Then $E|X|<\infty$ and

$$\operatorname*{plim}_{n\to\infty}\frac{n\bar{X}_n^2}{a_n^2}=\operatorname*{plim}_{n\to\infty}\frac{\bar{X}_n^2}{\mu_2(a_n)}=\frac{(EX)^2}{EX^2}.$$

Thus

$$\frac{1}{a_n^2}\left(\sum_{j=1}^n X_j^2-n\bar{X}_n^2\right)\xrightarrow{\mathrm{w}}\begin{cases}\dfrac{\mathrm{Var}(X)}{EX^2} & \text{if } EX^2<\infty,\\ 1 & \text{if } EX^2=\infty, \alpha=2,\\ \text{positive stable} & \text{if } 1<\alpha<2.\end{cases}$$

From (3.3) we have therefore that $\hat{\beta}_{1,n} \xrightarrow{P} \beta_1$ if and only if

$$\frac{1}{a_n^2}\left(\sum_{j=1}^{n} X_j Z_j - n\bar{X}_n \bar{Z}_n\right) \xrightarrow{P} 0. \tag{3.6}$$

Similarly, $\hat{\beta}_{0,n} \xrightarrow{P} \beta_0$ if and only if

$$\frac{1}{a_n^2}\left(\bar{Z}_n \sum_{j=1}^{n} X_j^2 - \bar{X}_n \sum_{j=1}^{n} X_j Z_j\right) \xrightarrow{P} 0. \tag{3.7}$$

Since $E|X| < \infty$, we can reparametrize the model as

$$Y_j = (\beta_0 + \beta_1(EX-1)) + \beta_1(X_j - EX + 1) + Z_j.$$

Thus we see there is no loss in assuming $EX=1$. Then $\bar{X}_n \xrightarrow{P} 1$ and (3.6), (3.7) become equivalent to

$$\left(\frac{1}{a_n^2}\sum_{j=1}^{n} (X_j - \bar{X}_n)^2\right)(\bar{Z}_n) \xrightarrow{P} 0,$$

$$\left(\frac{1}{a_n^2}\sum_{j=1}^{n} (X_j - \bar{X}_n)^2\right)\left(\frac{1}{n}\sum_{j=1}^{n} X_j Z_j\right) \xrightarrow{P} 0.$$

Thus, consistency occurs if and only if $\bar{Z}_n \xrightarrow{P} 0$ and $n^{-1}\sum_{j=1}^{n} X_j Z_j \xrightarrow{P} 0$. Furthermore, Corollary 2.7(ii) and Lemma 2.5(iv) show that $\bar{Z}_n \xrightarrow{P} 0$ implies

$$\lim_{t\to\infty} t(1-H_1(t)) = \lim_{t\to\infty} \lambda(t) = 0.$$

And this is sufficient for $n^{-1}\sum_{j=1}^{n} X_j Z_j \xrightarrow{P} 0$.

Suppose instead $\alpha \le 1$. We will show that $\bar{Z}_n \xrightarrow{P} 0$ is sufficient for consistency. We know $1-F_1(t) \in \mathrm{RV}_{-\alpha}$ and $\mu_1 \in \mathrm{RV}_{1-\alpha}$. In this case Lemma 2.3 says that if $b_n = \inf\{t:\ n\mu_1(t) \le t\}$, then

$$\frac{1}{b_n}\sum_{j=1}^{n} |X_j| \xrightarrow{w} \begin{cases} \text{positive stable law} & \text{if } \alpha < 1, \\ 1 & \text{if } \alpha = 1. \end{cases}$$

However, b_n/a_n is slowly varying. Hence,

$$\operatorname*{plim\,sup}_{n\to\infty} \frac{n\bar{X}_n^2}{a_n^2} \le \operatorname*{plim}_{n\to\infty} \frac{b_n^2}{na_n^2}\left(\frac{1}{b_n}\sum_{j=1}^{n} |X_j|\right)^2 = 0. \tag{3.8}$$

Therefore, $a_n^{-2}(\sum_{j=1}^{n} X_j^2 - n\bar{X}_n^2) \xrightarrow{w}$ a positive stable law. We again see that consistency is equivalent to (3.6) and (3.7).

If $\bar{Z}_n \xrightarrow{P} 0$, then $\nu_1(xt) - \nu_1(t) \to 0$ for any $x>0$. Clearly, ν_1 is thus slowly varying. By Lemma 2.4(ii), $\lambda_1 \in \mathrm{RV}_{1-\alpha}$. Again applying Lemma 2.3, there exist d_n such that $d_n^{-1}\sum_{j=1}^{n} |X_j Z_j|$ converges weakly and d_n/a_n is slowly varying. Include the facts that n/b_n is bounded and b_n/a_n is slowly varying and we have

$$\operatorname*{plim\,sup}_{n\to\infty} \left(\frac{1+|\bar{X}_n|}{a_n^2}\sum_{j=1}^{n} |X_j Z_j|\right)$$

$$\le \operatorname*{plim}_{n\to\infty}\left(\frac{b_n d_n}{na_n^2}\right)\left(\frac{n}{b_n}+\frac{1}{b_n}\sum_{j=1}^{n}|X_j|\right)\left(\frac{1}{d_n}\sum_{j=1}^{n}|X_jZ_j|\right)=0. \tag{3.9}$$

Thus, (3.7) holds since $a_n^{-2}\sum_{j=1}^n X_j^2$ converges weakly to a positive stable and $\bar{Z}_n \xrightarrow{\text{P}} 0$. Likewise, (3.6) follows from (3.8) and (3.9) and $\bar{Z}_n \xrightarrow{\text{P}} 0$. (In fact, (3.7) and $\alpha \le 1$ imply (3.6).)

We now show that under (ii) or (iii), $\bar{Z}_n \xrightarrow{\text{P}} 0$ is necessary for consistency.

(ii) Suppose $E|Z|^\gamma<\infty$ for some $\gamma>\alpha/(\alpha+1)$. There is no loss in taking $\gamma<\alpha\le 1$. Thus $E|XZ|^\gamma<\infty$. Then

$$\frac{1}{n^{1/\gamma}}\sum_{j=1}^{n}|X_jZ_j| \le \left(\frac{1}{n}\sum_{j=1}^{n}|X_jZ_j|^\gamma\right)^{1/\gamma}$$

and is stochastically bounded. Since $na_n^2/n^{1/\gamma}b_n\to\infty$, it then follows that

$$\frac{b_n}{na_n^2}\left|\sum_{j=1}^{n}X_jZ_j\right| \xrightarrow{\text{P}} 0.$$

Therefore, in this case, (3.7) is equivalent to

$$\frac{\bar{Z}_n}{a_n^2}\sum_{j=1}^{n}X_j^2 \xrightarrow{\text{P}} 0,$$

which is equivalent to $\bar{Z}_n \xrightarrow{\text{P}} 0$.

(iii) Finally, suppose $v_1\in \mathrm{RV}_{1-\gamma}$ for some $\gamma>0$. In view of the above we can assume $\gamma<\alpha$. Let $\{Z_j'\}$ be a sequence of independent, G-distributed random variables and independent of $\{Z_j\}$ and $\{X_j\}$. If (3.7) holds, then certainly

$$\frac{1}{na_n^2}\left(\sum_{j=1}^{n}(Z_j-Z_j')\sum_{j=1}^{n}X_j^2-n\bar{X}_n\sum_{j=1}^{n}X_j(Z_j-Z_j')\right)\xrightarrow{\text{P}} 0. \tag{3.10}$$

Define $W_j=Z_j-Z_j'$ and let G_2 and H_2 be the distributions of $|W_j|$ and $|X_jW_j|$, respectively.

By (2.9),

$$\lim_{t\to\infty}\frac{1-H_2(t)}{1-G_2(t)}=E|X|^\gamma,$$

which says there exists c_n such that both

$$\frac{1}{c_n}\sum_{j=1}^{n}W_j \quad\text{and}\quad \frac{1}{c_n}\sum_{j=1}^{n}X_jW_j$$

converge weakly to symmetric stable (γ) laws. In fact (cf. Cline, 1988),

$$\left(\frac{1}{c_n}\sum_{j=1}^{n}W_j,\frac{1}{c_n}\sum_{j=1}^{n}X_jW_j,\frac{1}{a_n^2}\sum_{j=1}^{n}|X_j|,\frac{1}{a_n^2}\sum_{j=1}^{n}X_j^2\right)\xrightarrow{\text{w}}(S_1,S_2,0,T),$$

where (S_1,S_2) is independent of T. It therefore follows that

$$\frac{1}{c_na_n^2}\left(\sum_{j=1}^{n}W_j\sum_{j=1}^{n}X_j^2-\sum_{j=1}^{n}X_j\sum_{j=1}^{n}X_jW_j\right)\xrightarrow{\text{w}} S_1T$$

and thus neither (3.10) nor (3.7) can hold when $\gamma<\alpha$. $\square$

The second term in (3.7) certainly seems to be relatively negligible, but in the cases where both terms are unbounded we found that we needed to impose some regularity on G. Thus, Theorem 3.2 does not quite show that $\bar{Z}_n \xrightarrow{P} 0$ is necessary in all cases for consistency in Model II.

Before proving Theorem 1.3, consistency for multiple regression, the following remark is helpful. An obvious, but limited, approach to the consistency problem is to determine a sequence c_n such that $c_n(X'X)^{-1}$ is stochastically bounded and

$$\frac{1}{c_n}X'Z = \left(\frac{1}{c_n}\sum_{j=1}^{n} X_{ij}Z_j\right)_{i=1}^{k} \xrightarrow{P} 0.$$

This approach is limited because $\hat{\beta}_n$ may be consistent when no such sequence exists (and generally this is the case). For example, consider Model II. Consistency occurs if $\bar{Z}_n \xrightarrow{P} 0$, but it is not necessary that $n^{-1}\sum_{j=1}^{n} X_jZ_j \xrightarrow{P} 0$. The terms involving $\sum_{j=1}^{n} X_jZ_j$ are sufficiently dominated to allow consistency.

Proof of Theorem 1.3. The least squares estimator $\hat{\beta}_n$ satisfies

$$(\hat{\beta}_n - \beta) = (X'X)^{-1}X'Z,$$

where $X = (X_j')_{j=1}^{n}$, $Z = (Z_j)_{j=1}^{n}$. We will assume $X'X$ is almost surely of full rank when $n \geq k$.

Note that $\min_i(a_{in}) = \inf\{t\colon n\mu_2^*(t) \leq t^2\}$. As before,

$$\frac{1}{a_{in}^2}\sum_{j=1}^{n} X_{ij}^2 \xrightarrow{w} \begin{cases} \text{positive stable } (\tfrac{1}{2}\alpha_i) \text{ law} & \text{if } \alpha_i < 2, \\ 1 & \text{if } \alpha_i = 2. \end{cases}$$

Let R_{ihn} and W_{ihn} be defined by

$$Q_n^{-1} = (R_{ihn})_{i,h} \quad \text{and} \quad W_n = (X'X)^{-1} = (W_{ihn})_{i,h}.$$

By a little algebra, $W_{ihn} = (1/a_{in}a_{hn})R_{ihn}$. Since

$$(\hat{\beta}_{in} - \beta_i) = \sum_{h=1}^{k}\left(W_{ihn}\sum_{j=1}^{n} X_{hj}Z_j\right),$$

$\hat{\beta}_n$ is consistent if

$$\frac{1}{a_{in}a_{hn}}\sum_{j=1}^{n} X_{hj}Z_j \xrightarrow{P} 0 \quad \text{for all } i,h. \tag{3.11}$$

To show that (3.11) in fact holds, first set

$$s_{ih}(t) = \inf\{n\colon a_{in}a_{hn} \geq t\}.$$

Note that s_{ih} is regularly varying with index $\varrho_{ih} = \alpha_i\alpha_h/(\alpha_i + \alpha_h) < \alpha_h$. Thus $E|X_{hj}|^{\delta} < \infty$ for some $\delta > \varrho_{ih}$. Furthermore,

$$s_{ih}(t) \leq \inf\left\{n\colon \min_i(a_{in}^2) \geq t\right\} = t/\mu_2^*(t^{1/2}).$$

Therefore, by (ii),

$$\lim_{t\to 0} s_{ih}(t)(1-G_1(t)) = 0. \tag{3.12}$$

Additionally, $\varrho_{ih} \le 1$ with $\varrho_{ih}=1$ only if $\alpha_i=\alpha_h=2$. In the latter case, condition (iii) implies

$$\lim_{t\to\infty} \frac{s_{ih}(t)}{t}\lambda_h(t) = 0 \quad \text{when } \varrho_{ih}=1. \tag{3.13}$$

By Corollary 2.7, (3.12) and (3.13) imply (3.11) and hence consistency holds.

(a) In the particular case $v_1 \in \mathrm{RV}_{1-\gamma}$, $\gamma > \frac{1}{2}\alpha^*$, condition (ii) holds automatically and (iii) is trivially true.

(b) In case Model III has an intercept, we can reparametrize so that $EX_{hj}=0$ whenever $\alpha_h=2$ and X_{hj} is not degenerate. In this case $\mu_2^*(t)=1$, for all $t\ge 1$. Thus (ii) and (iii) become

$$\lim_{t\to\infty} t(1-G_1(t)) = \lim_{t\to\infty} v(t) = 0,$$

i.e., $\bar{Z}_n \xrightarrow{\mathrm{P}} 0$.

(c) A similar argument holds when there is some i for which α_i is the unique maximum. □

Finally, we have the proof of Lemma 1.4.

Proof of Lemma 1.4. The condition guarantees that Q_n converges to a matrix whose only random components are on the diagonal and whose nonzero degenerate components form a positive definite matrix. □

Acknowledgement

The author greatly appreciates the hospitality of the Center for Stochastic Processes at the University of North Carolina.

References

Andrews, D.W.K. (1987a). Least squares regression with integrated or dynamic regressors under weak error conditions. *Econom. Theory* 3, 98–116.

Andrews, D.W.K. (1987b). On the performance of least squares in linear regression with nonexistent error means. Cowles Foundation Tech. Rpt., Yale University.

Bierens, H.J. (1981). Robust methods and asymptotic theory in nonlinear econometrics. *Lect. Notes Econ. Math. Systems No.* 192. Springer-Verlag, Berlin.

Bingham, N.H., C.M. Goldie and J.L. Teugels (1987). *Regular Variation.* Cambridge University Press, Cambridge.

Blattberg, R. and T. Sargent (1971). Regression with non-Gaussian stable disturbances: some sampling results. *Econometrica* **39**, 501–510.

Breiman, L. (1965). On some limit theorems similar to the arcsine law. *Theory Probab. Appl.* **10**, 323-331.

Carroll, R.J. and D.B.H. Cline (1988). An asymptotic theory for weighted least-squares with weights estimated by replication. *Biometrika* **75**, 35-43.

Cline, D.B.H. (1983). Infinite series of random variables with regularly varying tails. Tech. Rpt., Inst. Appl. Math. Statist., Univ. British Columbia.

Cline, D.B.H. (1986). Convolution tails, product tails and domains of attraction. *Probab. Theory Related Fields* **72**, 525-557.

Cline, D.B.H. (1988). Joint stable attraction of two sums of products. *J. Multivariate Anal.* **25**, 272-285.

Davis, R. and S. Resnick (1985a). Limit theory for moving averages of random variables with regularly varying tail probabilities. *Ann. Probab.* **13**, 179-197.

Davis, R. and S. Resnick (1985b). More limit theory for the sample correlation function of moving averages. *Stochastic Process. Appl.* **20**, 257-279.

Davis, R. and S. Resnick (1986). Limit theory for the sample covariance and correlation functions of moving averages. *Ann. Statist.* **14**, 533-558.

Embrechts, P. and C.M. Goldie (1980). On closure and factorization properties of subexponential and related distributions. *J. Austral. Math. Soc.* (*Ser. A*) **29**, 243-256.

Feller, W. (1971). *An Introduction to Probability Theory and its Applications*, 2nd ed., Wiley, New York.

Gnedenko, B.V. and A.N. Kolmogorov (1968). *Limit Distributions for Sums of Independent Random Variables*, transl. by K.L. Chung. Addison-Wesley, New York.

de Haan, L. (1970). On regular variation and its application to the weak convergence of sample extremes. *Math. Centre Tracts No.* 32, Mathematisch Centrum, Amsterdam.

Hannan, E.J. and M. Kanter (1977). Autoregressive processes with infinite variance. *J. Appl. Probab.* **14**, 411-415.

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

Kanter, M. and W.L. Steiger (1974). Regression and autoregression with infinite variance. *Adv. Appl. Probab.* **6**, 768-893.

Maller, R.A. (1981). A theorem on products of random variables, with application to regression. *Austral. J. Statist.* **23**, 117-185.

Maronna, R.A. and V.J. Yohai, (1981). Asymptotic behavior of general *M*-estimates for regression and scale with random carriers. *Z. Wahrsch. Verw. Geb.* **58**, 7-20.

Press, S.J. (1975). Stable distributions: probability, inference and applications in finance - a survey and a review of recent results. In: G.P. Patel et al., Eds., *Statistical Distributions in Scientific Work* 1, 87-1-2, Reidel, Dordrecht.

Smith, V.K. (1973). Least squares regression with Cauchy errors. *Bull. Oxford Univ. Inst. Econom. Statist.* **35**, 223-231.

Yohai, V.J. and R.A. Maronna, (1977). Asymptotic behavior of least squares estimators for autoregressive processes with infinite variance. *Ann. Statist.* **5**, 554-560.