*The Annals of Statistics*
1978, Vol. 6, No. 1, 215–223

# ON CONSISTENCY IN TIME SERIES ANALYSIS

BY P. M. ROBINSON[1]

*Harvard University*

A number of statistics that arise in time series analysis can be represented as the sum of a partial realization of a possibly serially dependent and nonstationary discrete-parameter stochastic process. The almost sure and $L_p$, $p > 1$, convergence of such statistics is investigated, under various moment conditions. The results are applied to the least squares estimates of multiple regressions.

**1. Introduction.** In time series analysis one is sometimes concerned with the partial sums $S_N = \sum_{n=1}^N \xi_n$, $N \geqq 1$, of a realization of the discrete-parameter zero-mean stochastic process $\{\xi_n, n \geqq 1\}$, and wishes to determine for which increasing sequences $\{s_N, N \geqq 1\}$,

$$S_N/s_N \to 0\,, \quad \text{almost surely (a.s.)} \tag{1.1}$$

or

$$E|S_N/s_N|^\nu \to 0, \qquad \text{some} \quad \nu\,, \tag{1.2}$$

as $N \to \infty$. Both (1.1) and (1.2) imply that $S_N/s_N \to 0$ in probability, but neither (1.1) nor (1.2) imply the other.

Often $\{\xi_n, n \geqq 1\}$ is the output generated by passing a set of observable or unobservable discrete-parameter time series through a filter which, possibly, is nonlinear and time-varying. Some important situations of this type might be described by the following model, which we refer to as

CONDITIONS A.

$$\xi_n = \sum_{j_1=-\infty}^\infty \cdots \sum_{j_q=-\infty}^\infty \zeta_{\mathbf{j}n}\,, \tag{1.3}$$

where $\zeta_{\mathbf{j}n}$ is written in place of $\zeta_{j_1,\cdots,j_q n}$,

$$E(\zeta_{\mathbf{j}n} \,|\, \zeta_{\mathbf{j}m}, m < n) = 0\,, \quad \text{a.s.}, \tag{1.4}$$

all $-\infty < j_1, \cdots, j_q < \infty$ and all $n \geqq 2$, and for some $\nu > 1$ there exist nonnegative constants $\{\alpha_{kj}, -\infty < j < \infty, 1 \leqq k \leqq q\}$, $\{a_n, n \geqq 1\}$ such that

$$E|\zeta_{\mathbf{j}n}|^\nu < \alpha_\mathbf{j}{}^\nu a_n{}^\nu\,, \qquad \alpha_\mathbf{j} = \prod_{k=1}^q \alpha_{kj_k}\,, \tag{1.5}$$

$$\sum_{j=-\infty}^\infty \alpha_{kj} < \infty\,. \tag{1.6}$$

We take $\nu$ to be the largest value for which such relationships hold.

---

Received July 1975; revised March 1977.

[1] Part of this research was carried out while the author was at the University of California, Berkeley. I am grateful to the referee for comments on earlier versions. Support of NSF grant SOC75-13436 is acknowledged.

*AMS* 1970 *subject classifications*. Primary 60F15, 62M10; Secondary 60G45.

*Key words and phrases*. Almost sure convergence, $L_p$ convergence, martingale inequalities, time series regression.

Conditions A generalize a familiar representation for a time series,

$$\xi_n = \sum_{j=-\infty}^{\infty} \gamma_j \chi_{n-j}\,, \qquad \sum_{j=-\infty}^{\infty} |\gamma_j| < \infty\,, \tag{1.7}$$

where $\{\chi_n, -\infty < n < \infty\}$ is a sequence of independent and identically distributed random variables with zero means and finite variances. Then $\xi_n$ is strictly and second order stationary and ergodic. In Conditions A, stationarity is not implied by the homogeneity assumption (1.5); indeed, when $a_n \to \infty$ or when one can choose $a_n$ that $\to 0$, evolutionary behavior is implied. As in Hannan and Heyde [5], independence is replaced by a martingale assumption, (1.4). In Conditions A, the summable $\alpha_{kj}$ are not introduced until one bounds $E|\zeta_{jn}|^\nu$. Cases $\nu < 2$ and $\nu > 2$, as well as $\nu = 2$, seem important, and the range of possible $s_N$ depends on $\nu$. The multiple summation in (1.3) is motivated primarily to handle cases in which $\xi_n$ arises as the product of random variables, each of which may, for example, have a representation of the form (1.7). This is so when $S_N$ is a regression sum of squares, or when $S_N/N$ is the deviation from its expectation of a sample $r$th-order autocovariance of an $r$th-order stationary time series (see Robinson [8]). Also, a time-varying moving average of uncorrelated but heteroscedastic random variables may arise when certain continuous time stochastic processes are sampled at discrete but unequally spaced intervals (see Robinson [9]).

Define $b_{\nu 0} = 0$, $b_{\nu N} = \sum_1^N a_n^\nu$, $N > 1$. Throughout, $K$ represents a positive constant, not necessarily the same one.

THEOREM 1. *Let Conditions* A *hold with* $1 < \nu \leqq 2$ *and let*

$$b_{\nu[h^{k+1}]} < K b_{\nu[h^k]}\,, \qquad k \geqq 1 \tag{1.8}$$

*for some* $h > 1$. *Then in* (1.1) *we may choose*

$$s_N = (b_{\nu N} (\ln N)^{1+\phi+\nu} (\ln \ln N)^{1+\psi+\delta})^{1/\nu} \tag{1.9}$$

*when* $\phi + \nu > 0$, $\delta > 0$ *or when* $\phi + \nu = 0$, $\delta > \max(0, -\psi)$.

In most cases one would choose $\phi = \psi = 0$ in (1.9). However, when $a_n = n^{-\frac{1}{2}}$, $\nu = 2$, $h = 2$, one has $\phi = -1$, $\psi = 0$; when $a_n = (n \ln n)^{-\frac{1}{2}}$, $n \geqq 2$, $\nu = 2$, $h = 2$, one has $\phi = 0$, $\psi = -1$.

Martingale convergence theorems can yield stronger results, under generally stronger conditions. From Loève [6], we may be able to choose $s_N = b_{\nu N}^{1/\nu}$ for $1 < \nu < 2$, and from Neveu [7] we may be able to choose $s_N = b_{2N}^{\frac{1}{2}} (\ln b_{2N})^{\frac{1}{2}+\delta}$, $\delta > 0$, for $\nu = 2$.

This approach is not easily extended to handle a representation of the generality of Conditions A, however, and it cannot be used to improve the results of Theorems 2 and 3 below.

For $\nu > 2$, define by $\tau$ the largest power of 2 such that $\tau = 2^t < \nu$ and

$$E(\zeta_{jn}^\tau \mid \zeta_{jm}, m < n) < K \alpha_j^\tau a_n^\tau\,. \tag{1.10}$$

By virtue of (1.5), this is similar to $\{\zeta_{jn}, n \geqq 1\}$ being an independent sequence, up to $\tau$th moments.

THEOREM 2. *Let Conditions* A *hold with* $\nu > 2$ *and let*

$$b_{2[h^{k+1}]} < K b_{2[h^k]}, \qquad k > 1. \tag{1.11}$$

*When* $\nu > 2\tau$, *let*

$$(N-M)^{\nu/2\tau-1}(b_{\nu N} - b_{\nu M}) < K(b_{2N} - b_{2M})^{\nu/2}, \qquad 1 \leqq M \leqq N. \tag{1.12}$$

*Then in* (1.1), *we may choose*

$$s_N = b_{2N}^{\frac{1}{2}}(\ln N)^{1/\nu}(\ln\ln N)^{(1+\delta)/\nu}, \qquad any \quad \delta > 0. \tag{1.13}$$

We next consider $L_p$-convergence. Lemmas 1 and 2 below (established to prove Theorems 1 and 2) imply that in (1.2), $s_N$ may be chosen as

$$b_{\nu N}^{1/\nu} c_N, \quad 1 \leqq \nu \leqq 2; \qquad b_{2N}^{\frac{1}{2}} c_N, \quad \nu > 2,$$

where $\{c_n\}$ is any sequence such that $c_n \to \infty$ as $n \to \infty$. An improvement is possible when $1 < \nu < 2$.

CONDITIONS B. Define

$$\zeta'_{jn} = \zeta_{jn} I(|\zeta_{jn}|^\nu \leqq CE|\zeta_{jn}|^\nu), \qquad \zeta''_{jn} = \zeta_{jn} - \zeta'_{jn}.$$

For any $\varepsilon > 0$ and all $\mathbf{j}$, $n$, $C$ can be chosen such that

$$E|\zeta''_{jn}|^\nu < \varepsilon E|\zeta_{jn}|^\nu.$$

This is a type of uniform integrability condition.

THEOREM 3. *Let Conditions* A *and* B *hold with* $1 < \nu < 2$ *and let*

$$b_{2N}^\nu / b_{\nu N}^2 \to 0 \qquad as \quad N \to \infty. \tag{1.14}$$

*Then in* (1.2) *we may choose*

$$s_N = b_{\nu N}^{1/\nu}.$$

Each of the conditions (1.8), (1.11), (1.12), (1.14) requires that $a_n$ should not increase too fast, excluding $a_n = \theta^n$, $|\theta| > 1$, but not $a_n = n^\theta$.

The theorems are proved in Section 2. In Section 3, the results are applied to the convergence of least squares estimates of multiple time series regression.

**2. Proofs of theorems.** Our results rest heavily on calculations of the order of magnitude of the $\nu$th absolute moment of $S_{MN} = S_N - S_M$, $0 \leqq M < N$ ($S_0 = 0$).

LEMMA 1. *Let Conditions* A *hold, for* $1 < \nu \leqq 2$. *Then*

$$E|S_{MN}|^\nu < K(b_{\nu N} - b_{\nu M}).$$

PROOF. First we prove that

$$|S_{MN}|^\nu < K \sum'_{j_1} \cdots \sum'_{j_q} \alpha_{\mathbf{j}}^{1-\nu} |\sum_n \zeta_{\mathbf{j}n}|^\nu, \tag{2.1}$$

where the sum over $n$ is for $M < n \leqq N$ and, for $1 \leqq k \leqq q$, the sum over $j_k$ is over those $j$ for which $\alpha_{kj} \neq 0$. To obtain (2.1), one first writes

$$S_{MN} = \sum'_{j_1} \cdots \sum'_{j_q} \alpha_{1j_1}^{1-1/\nu} \alpha_{1j_1}^{1/\nu} \sum_n (\zeta_{\mathbf{j}n}/\alpha_{1j_1})$$

whence Hölder's inequality gives

$$|S_{MN}|^\nu < (\sum_j \alpha_{1j})^{\nu-1} \sum'_{j_1} \alpha_{1j_1}^{1-\nu} |\sum'_{j_2} \cdots \sum'_{j_q} \sum_n \zeta_{jn}|^\nu .$$

Similarly, one has

$$|\sum'_{j_2} \cdots \sum'_{j_q} \sum_n \zeta_{jn}|^\nu < (\sum_j \alpha_{2j})^{\nu-1} \sum'_{j_2} \alpha_{2j_2}^{1-\nu} |\sum'_{j_3} \cdots \sum'_{j_q} \sum_n \zeta_{jn}|^\nu .$$

After using Hölder's inequality a total of $q$ times, (2.1) follows by virtue of (1.6). Then from Von Bahr and Esseen [13],

$$E|\sum_n \zeta_{jn}|^\nu \leqq K \sum_n E|\zeta_{jn}|^\nu ,$$

because of (1.4). Then the lemma is established by using (1.5) and then (1.6). □

LEMMA 2. *Let Conditions* A *hold for* $\nu > 2$ *and, when* $\nu > 2\tau$, *let* (1.12) *hold. Then*

$$E|S_{MN}|^\nu < K(b_{2N} - b_{2M})^{\nu/2} .$$

PROOF. Because of (2.1) one has only to show that

$$E|\sum_n \zeta_{jn}|^\nu < K\alpha_j(b_{2N} - b_{2M})^{\nu/2} . \tag{2.2}$$

From (1.4), Theorem 9 of Burkholder [3] and Minkowski's inequality,

$$E|\sum_n \zeta_{jn}|^\nu < KE(\sum_n \zeta_{jn}^2)^{\nu/2} < K(N - M)^{\nu/2-1} \sum_n E|\zeta_{jn}|^\nu .$$

Thus when $t = 0$, (2.2) follows from use of (1.5). When $t > 0$, define

$$\eta_{jn}^{(0)} = \zeta_{jn} , \qquad \eta_{jn}^{(l)} = \eta_{jn}^{(l-1)2} - \bar{\eta}_{jn}^{(l)} , \qquad \bar{\eta}_{jn}^{(l)} = E(\eta_{jn}^{(l-1)2} | \zeta_{j,n-1}, \cdots, \zeta_{j1}) , \tag{2.3}$$

for $1 \leqq l \leqq t$. It follows from Theorem 9 of [3] that, for all $\theta$ and some $K$ depending only on $\theta$,

$$E|\sum_n \eta_{jn}^{(l)}|^\theta < KE(\sum_n \eta_{jn}^{(l)2})^{\theta/2} . \tag{2.4}$$

Repeated use of (2.3), (2.4) and the $c_r$-inequality produces

$$E|\sum_n \zeta_{jn}|^\nu < KE(\sum_n \eta_{jn}^{(t)2})^{\nu/2\tau} + K \sum_{l=1}^t E(\sum_n \bar{\eta}_{jn}^{(l)})^{\nu/2^l} . \tag{2.5}$$

Now by Jensen's inequality for conditional expectations and (1.10), (1.5),

$$E(\zeta_{jn}^r | \zeta_{j,n-1}, \cdots, \zeta_{j1}) < K\alpha_j^r a_n^r , \qquad r < \tau .$$

Then it is easily seen that

$$\bar{\eta}_{jn}^{(l)} < K\alpha_j^{2^l} a_n^{2^l} , \qquad 1 \leqq l \leqq t . \tag{2.6}$$

Thus,

$$\begin{aligned} \sum_{l=1}^t E(\sum_n \bar{\eta}_{jn}^{(l)})^{\nu/2^l} &< K\alpha_j^\nu \sum_{l=1}^t (b_{2^l N} - b_{2^l M})^{\nu/2^l} \\ &< K\alpha_j^\nu (b_{2N} - b_{2M})^{\nu/2} . \end{aligned} \tag{2.7}$$

Now on expanding $\eta_{jn}^{(t)}$ and using (1.5), (2.6), it is seen that

$$E\eta_{jn}^{(t)\nu/\tau} < Ka_n^\nu \alpha_j^\nu . \tag{2.8}$$

Thus

$$E(\textstyle\sum_n \eta_{jn}^{(t)2})^{\nu/2\tau} < K(N-M)^{\nu/2\tau-1} \sum_n E\eta_{jn}^{(t)\nu/\tau}$$
$$(2.9) \qquad < K(N-M)^{\nu/2\tau-1}\alpha_j^{\nu}(b_{\nu N}-b_{\nu M})$$
$$< K\alpha_j^{\nu}(b_{2N}-b_{2M})^{\nu/2},$$

the last inequality following from (1.12). Therefore the desired bound for (2.5) is provided by (2.7) and (2.9). □

PROOF OF THEOREM 1. The method of proof is similar to one used by Stout [12]. Write $\tilde{S}_k = S_{[h^k]}$, $\tilde{s}_k = s_{[h^k]}$. There are finitely many $k$ for which $h^k < 2$. For $k > 1$ such that $h^k \geqq 2$, and with $s_N$ given by (1.9),

$$P(|\tilde{S}_k| > \varepsilon\tilde{s}_k) \leqq E|\tilde{S}_k|^{\nu}\varepsilon^{-\nu}\tilde{s}_k^{-\nu}$$
$$\leqq K\varepsilon^{-\nu}(\ln[h^k])^{-1-\phi-\nu}(\ln\ln[h^k])^{-1-\psi-\delta}$$
$$< K\varepsilon^{-\nu}k^{-1-\phi-\nu}(\ln k)^{-1-\psi-\delta},$$

from Markov's inequality, Lemma 1 and $[h^k] \geqq Kh^k$. Thus, when $\phi+\nu > 0$ or when $\phi+\nu = 0$, $\psi+\delta > 0$ it follows from the Borel-Cantelli lemma that $\tilde{S}_k/\tilde{s}_k \to 0$, a.s. Now define $T_k = \max_{[h^k]<n\leqq[h^{k+1}]} |S_{[h^k]n}|$. We need consider only $k$ such that $[h^{k+1}] - [h^k] \geqq 1$. From Billingsley [2, page 102], Lemma 1 implies

$$(2.10) \qquad ET_k^{\nu} < K[\ln([h^{k+1}]-[h^k])]^{\nu}[b_{\nu[h^{k+1}]} - b_{\nu[h^k]}]$$
$$< Kk^{\nu}b_{\nu[h^k]}$$

using (1.8) and $[h^{k+1}] - [h^k] \leqq h^{k+1} - h^k + 1 \leqq Kh^k$. Thus, much as before

$$P(T_k > \tilde{s}_k) < ET_k^{\nu}\varepsilon^{-\nu}\tilde{s}_k^{-\nu} < K\varepsilon^{-\nu}k^{-1}(\ln k)^{-1-\delta}.$$

When $\delta > 0$, it follows that $T_k/\tilde{s}_k \to 0$, a.s. Finally, for $[h^k] \leqq N \leqq [h^{k+1}]$,

$$S_N/s_N \leqq |S_N - \tilde{S}_k|/\tilde{s}_k + |\tilde{S}_k|/\tilde{s}_k \leqq T_k/\tilde{s}_k + |\tilde{S}_k|/\tilde{s}_k. \qquad □$$

PROOF OF THEOREM 2. The proof is almost the same. For $k > 1$ such that $h^k \geqq 2$ and with $s_N$ given by (1.13), it is readily shown that

$$P(|\tilde{S}_k| > \varepsilon\tilde{s}_k) < K\varepsilon^{-\nu}k^{-1}(\ln k)^{-1-\delta},$$

so $\tilde{S}_k/\tilde{s}_k \to 0$, a.s., for $\delta > 0$. Now from [2, page 94], Lemma 2 implies

$$(2.11) \qquad P(T_k > \varepsilon) < K\varepsilon^{-\nu}(b_{2[h^{k+1}]} - b_{2[h^k]})^{\nu/2}.$$

Therefore, using (1.11), one has

$$P(T_k > \varepsilon\tilde{s}_k) < K\varepsilon^{-\nu}k^{-1}(\ln k)^{-1-\delta},$$

and so $T_k/s_k \to 0$, a.s., and the proof is completed as before. □

Serfling [10] proves inequalities similar to (2.10) and (2.11) for $\nu \geqq 2$. For $\nu > 2$, his assumed bound for $E|S_{MN}|^{\nu}$ (and derived bound for $E\max_{M<n\leqq N}|S_{Mn}|^{\nu}$), depend on $M$ and $N$ only through $N-M$, and he requires $2b_{2N} \leqq b_{2,2N}$. These assumptions exclude many cases where $a_n \to \infty$ or 0. Under Serfling's assumptions, there is no loss in generality in taking $h = 2$ in the above proof (see Stout [12, page 210]).

Serfling [11] lists several alternative $\{\xi_n\}$ for which results similar to those of Lemma 2 and Theorem 2 are available.

PROOF OF THEOREM 3. Because of (2.1) we need only show that for any $\varepsilon > 0$ which does not depend on **j**,

$$E|\textstyle\sum_{n=1}^{N} \zeta_{\mathbf{j}n}|^{\nu} < \varepsilon\alpha_{\mathbf{j}}{}^{\nu} b_{\nu N} \tag{2.12}$$

for sufficiently large $N$. Our method of proof then extends one given by Chow [4]. From [3, Theorem 9] and Conditions B,

$$\begin{aligned} E|\textstyle\sum_n \zeta_{\mathbf{j}n}|^{\nu} &< KE(\textstyle\sum_n \zeta_{\mathbf{j}n}^2)^{\nu/2} < EK(\textstyle\sum_n \zeta_{\mathbf{j}n}'^2)^{\nu/2} + KE(\textstyle\sum_n \zeta_{\mathbf{j}n}''^2)^{\nu/2} \\ &\leqq KE(C^2 \textstyle\sum_n E|\zeta_{\mathbf{j}n}'|^2)^{\nu/2} + K\textstyle\sum_n E|\zeta_{\mathbf{j}n}''|^{\nu} \\ &< \alpha_{\mathbf{j}} b_{\nu N}(Kb_{2N}^{\nu/2}/b_{\nu N} + K\varepsilon) \end{aligned}$$

so (2.12) follows from (1.14). □

**3. Multiple regression.** The preceding results may be applied to estimates of the model

$$y_n = \beta_1 z_{1n} + \cdots + \beta_p z_{pn} + x_n\,, \qquad n \geqq 1\,.$$

Define

$$\boldsymbol{\beta} = \begin{bmatrix} \beta_1 \\ \vdots \\ \beta_p \end{bmatrix}, \quad \mathbf{y}_N = \begin{bmatrix} y_1 \\ \vdots \\ y_N \end{bmatrix}, \quad \mathbf{x}_N = \begin{bmatrix} x_1 \\ \vdots \\ x_N \end{bmatrix}, \quad \mathbf{Z}_N = \begin{bmatrix} z_{11} & \cdots & z_{p1} \\ \vdots & & \vdots \\ z_{1N} & \cdots & z_{pN} \end{bmatrix}, \tag{3.1}$$

$$\hat{\boldsymbol{\beta}}_N = (\hat{\beta}_{jN}) = (\mathbf{Z}_N{}^T\mathbf{Z}_N)^{-1}\mathbf{Z}_N{}^T\mathbf{y}_N = \boldsymbol{\beta} + (\mathbf{Z}_N{}^T\mathbf{Z}_N)^{-1}\mathbf{Z}_N{}^T\mathbf{x}_N$$

which is the least squares estimate of $\boldsymbol{\beta}$. A variety of results may be generated from Theorems 1–3. Three examples will be given, under somewhat simplified conditions.

The first assumes the $z_{in}$ are random variables with representations

$$z_{in} = \textstyle\sum_{j=-\infty}^{\infty} u_{ijn}\,, \qquad n \geqq 1,\ 1 \leqq i \leqq p\,,$$

where there exist real constants $\alpha_{ij}$, $a_{in}$, $a_n$ such that $\sum_{j=-\infty}^{\infty} |\alpha_{ij}| < \infty$,

$$E(u_{ijn}\,|\,u_{ijm},\ m < n;\ x_m,\ m \leqq n) = 0\,, \quad \text{a.s.}, \tag{3.2}$$

$$E(u_{hj_1n}u_{ij_2n}\,|\,u_{hj_1m},\ u_{hj_2m},\ m < n) = \alpha_{hj}\alpha_{ij}a_{hn}a_{in}\,, \quad \text{a.s.}, \qquad j_1 = j_2 = j\,, \tag{3.3}$$

$$= 0\,, \quad \text{a.s.}, \qquad j_1 \neq j_2\,, \tag{3.4}$$

$$E(|u_{ijn}|^{\mu}\,|\,u_{hjm},\ m < n,\ 1 \leqq h \leqq p;\ x_m,\ m \leqq n) < K|\alpha_{ij}a_{in}|^{\mu} \tag{3.5}$$

$$E|x_n|^{\nu} < |a_n|^{\nu}\,. \tag{3.6}$$

These are Conditions C.

Define $\boldsymbol{\Omega}_N$ as the $p \times p$ matrix with $(h, i)$th element $\sum_{n=1}^{N} \omega_{hin}$, where

$$\omega_{hin} = Ez_{hn}z_{in} = (\textstyle\sum_{j=-\infty}^{\infty} \alpha_{hj}\alpha_{ij})a_{hn}a_{in}$$

and define

$$b_{\nu iN} = \textstyle\sum_{n=1}^{N} |a_{in}|^{\nu}\,, \qquad \mathbf{B}_N = \operatorname{diag}\{b_{21N}, \cdots, b_{2pN}\}\,. \tag{3.7}$$

THEOREM 4. *Let Conditions* C *hold with* $\mu > 4$, $\nu > 2$, $\mu \geqq \nu$, *and let*

$$\lim_{N\to\infty} \det \{\mathbf{B}_N^{-\frac{1}{2}}\mathbf{\Omega}_N \mathbf{B}_N^{-\frac{1}{2}}\} \neq 0 , \tag{3.8}$$

$$b_{2,h,2N} < K b_{2hN} , \qquad 1 \leqq h \leqq p,\ N \geqq 1 , \tag{3.9}$$

$$\limsup_{N\to\infty} (\ln N)^{4/\mu}(\ln \ln N)^{(4+\delta)/\mu} \max_{1\leqq n\leqq 2N} a_{hn}^2/b_{2hN} < \infty , \qquad 1 \leqq h \leqq p,\ \delta > 0 , \tag{3.10}$$

$$\limsup_{N\to\infty} (\ln N)^{2/\nu}(\ln \ln N)^{(2+\delta)/\nu} \max_{1\leqq n\leqq 2N} a_n{}^2/b_{2iN} < \infty , \qquad \delta > 0 . \tag{3.11}$$

*Then* $\hat\beta_{iN} \to \beta_i$, a.s. *as* $N \to \infty$.

PROOF. It is readily deduced from (3.1) that

$$|\hat\beta_{iN} - \beta_i| \leqq ||\mathbf{B}_N^{\frac{1}{2}}(\mathbf{Z}_N{}^T\mathbf{Z}_N)^{-1}\mathbf{B}_N^{\frac{1}{2}}|| \sum_{h=1}^{p} b_{2hN}^{-\frac{1}{2}} b_{2iN}^{-\frac{1}{2}} |\sum_{n=1}^{N} z_{hN} x_n| , \tag{3.12}$$

where $||\cdot||$ is the Euclidean norm. From (3.8) it suffices to prove

$$\mathbf{B}_N^{-\frac{1}{2}}(\mathbf{Z}_N{}^T\mathbf{Z}_N - \mathbf{\Omega}_N)\mathbf{B}_N^{-\frac{1}{2}} \to \mathbf{0} , \quad \text{a.s.}, \tag{3.13}$$

$$b_{2hN}^{-\frac{1}{2}} b_{2iN}^{-\frac{1}{2}} \sum_{n=1}^{N} z_{hn} x_n \to 0 , \quad \text{a.s.}, \qquad 1 \leqq h \leqq p . \tag{3.14}$$

Consider (3.13) first. Denote the $(k, l)$th element of $\mathbf{B}_N^{-\frac{1}{2}}(\mathbf{Z}_N{}^T\mathbf{Z}_N)^{-1}\mathbf{B}_N^{-\frac{1}{2}}$ by $b_{2kN}^{-\frac{1}{2}} b_{2lN}^{-\frac{1}{2}} \sum_1^N \xi_n$, where $\xi_n = \sum \sum_{j_1,j_2=-\infty}^{\infty} \zeta_{\mathbf{j}n}$,

$$\begin{aligned} \zeta_{\mathbf{j}n} = \zeta_{j_1 j_2 n} &= u_{kjn}u_{ljn} - a_{kn}a_{ln}\alpha_{kj}\alpha_{lj} , \qquad & j_1 = j_2 = j , \\ &= u_{kj_1n}u_{lj_2n} , & j_1 \neq j_2 . \end{aligned}$$

Then (1.4) is satisfied because of (3.3), (3.4), and (1.5) is satisfied with

$$\begin{aligned} E|\zeta_{\mathbf{j}n}|^{\mu/2} &< K(E|u_{kj_1n}|^\mu E|u_{lj_2n}|^\mu)^{\frac{1}{2}} + K|a_{kn}a_{ln}\alpha_{kj_1}\alpha_{lj_1}|^{\mu/2} \\ &< K|a_{kn}a_{ln}\alpha_{kj_1}\alpha_{lj_2}|^{\mu/2} , \end{aligned}$$

because of (3.3), (3.5). Since $\sum_j |\alpha_{lj}| < \infty$, $1 \leqq l \leqq p$, all of Conditions A are satisfied. Next, note that, because of (3.5), the result $E|\sum \xi_n|^{\mu/2} < K(\sum a_{kn}^2 a_{ln}^2)^{\mu/4}$ does not require a condition like (1.12). With regard to (1.11) note that

$$\sum_1^N a_{kn}^2 a_{ln}^2 \leqq \max_{1\leqq n\leqq 2N} a_{kn}^2 b_{2lN} , \qquad \sum_{N+1}^{2N} a_{kn}^2 a_{ln}^2 \leqq \max_{1\leqq n\leqq 2N} a_{kn}^2 b_{2lN} ,$$

using (3.9) in the second case. Therefore, from Theorem 2,

$$\sum_1^N (z_{kn}z_{ln} - \omega_{kln})/(\max_{1\leqq n\leqq 2N} |a_{kn}| b_{2lN}^{\frac{1}{2}} (\ln N)^{2/\mu}(\ln\ln N)^{(2+\delta)/\mu}) \to 0 , \quad \text{a.s.}$$

Then (3.12) is a consequence of (3.10). To establish (3.14) note that $\sum_1^N z_{hn}x_n$ is of the form $\sum_1^N \xi_n$, $\xi_n = \sum_{-\infty}^{\infty} \zeta_{jn}$, where $\zeta_{jn} = u_{hjn}x_n$. Then Conditions A are satisfied because of (3.2), (3.5), (3.6), with $E|\zeta_{jn}|^\nu < |\alpha_{hj}a_{hn}a_n|^\nu$, $\nu \leqq \mu$. Note now that $E|\sum \xi_n|^\nu < K(\sum a_n{}^2 a_{hn}^2)^{\nu/2}$ does not require a condition like (1.13), because of (3.5), (3.6) and $\mu \geqq \nu$. Note also that

$$\sum_1^N a_n{}^2 a_{hn}^2 \leqq \max_{1\leqq n\leqq 2N} a_n{}^2 b_{2hN} , \qquad \sum_{N+1}^{2N} a_n{}^2 a_{hn}^2 \leqq \max_{1\leqq n\leqq 2N} a_n{}^2 b_{2hN} .$$

Therefore, from Theorem 2,

$$\sum_1^N z_{hn}x_n/(\max_{1\leqq n\leqq 2N} |a_n| b_{2hN}^{\frac{1}{2}} (\ln N)^{1/\nu}(\ln\ln N)^{(1+\delta)/\nu}) \to 0 , \quad \text{a.s.},$$

and so (3.14) follows from (3.11). □

An example in which the regressors are "decreasing" while the residuals are "increasing" but Theorem 4 applies is when $a_{hn} = n^{-\frac{1}{2}}$, $a_n = \ln \ln n$.

Notice that some of the weak dependence assumptions on the $z_{hn}$, introduced to establish (3.13), are used to prove (3.14) without any weak dependence assumptions on $x_n$. By way of contrast, Anderson and Taylor [1, Theorem 2], for the case $p = 1$, assume $x_n$ is a martingale difference sequence and assume the a.s. divergency of $\sum_1^N z_n^2$, with no restrictions on its rate of increase.

Now let the $z_{hn}$ be fixed constants, and let $x_n = \sum_{j=-\infty}^{\infty} v_{jn}$, where $\alpha_j$, $a_n$ are such that $\sum_{j=-\infty}^{\infty} |\alpha_j| < \infty$, $E(v_{jn} \mid v_{jm}, m < n) = 0$, a.s., $E(|v_{jn}|^\nu \mid v_{jm}, m < n) < |\alpha_j a_n|^\nu$. For $\nu > 2$, let

$$E(|v_{jn}|^\nu \mid v_{jm}, m < n) < K|\alpha_j a_n|^\nu . \tag{3.15}$$

These are Conditions D. We replace $a_{hn}$ by $z_{hn}$ in the definitions of $b_{\nu hN}$ and $\mathbf{B}_N$ in (3.7).

THEOREM 5. *Let Conditions* D, (3.7) *and* (3.9) *hold. For* $1 < \nu \leqq 2$, *let*

$$\limsup_{N\to\infty} (\ln N)^{(2+2\nu)/\nu}(\ln \ln N)^{(2+\delta)/\nu} \max_{1\leqq n\leqq 2N} a_n^2/b_{2iN} < \infty$$

*and for* $\nu > 2$ *let* (3.11) *hold. Then* $\hat{\beta}_{iN} \to \beta_i$, a.s. *as* $N \to \infty$.

The proof applies Theorems 1 and 2 in a way that is similar to that of Theorem 4, so it is omitted. Anderson and Taylor [1, Theorem 1] proved that if the $x_n$ are i.i.d. normal variables, then a necessary and sufficient condition for $\hat{\boldsymbol{\beta}}_N \to \boldsymbol{\beta}$ a.s. is $(\mathbf{Z}_N{}^T\mathbf{Z}_N)^{-1} \to 0$. Our restriction (3.11) is relaxed as $\nu$ increases, however.

Finally, we give an $L_p$-convergence theorem.

THEOREM 6. *Let Conditions* D *hold. In case* $1 < \nu < 2$, *let*

$$\lim_{N\to\infty} \max_{1\leqq n\leqq N} a_n^2 b_{\nu hN}^{2/\nu}/(b_{2hN}\, b_{2iN}) = 0 , \qquad 1 \leqq h \leqq p , \tag{3.16}$$

*or let the* $|v_{jn}/\alpha_j a_n|^\nu$ *be uniformly integrable and*

$$\limsup_{N\to\infty} \max_{1\leqq n\leqq N} a_n^2 b_{\nu hN}^{2/\nu}/(b_{2hN}\, b_{2iN}) < \infty , \qquad 1 \leqq h \leqq p . \tag{3.17}$$

*In case* $\nu \geqq 2$, *let*

$$\lim_{N\to\infty} \max_{1\leqq n\leqq N} a_n^2/b_{2iN} = 0 . \tag{3.18}$$

*Then, for* $\nu > 1$, $E|\hat{\beta}_{iN} - \beta_i|^\nu \to 0$ *as* $N \to \infty$.

PROOF. We use (3.12) and thus have to investigate

$$E|\textstyle\sum_{n=1}^N z_{hn} x_n|^\nu . \tag{3.19}$$

From Lemma 1, this is bounded by

$$K \textstyle\sum_{n=1}^N |a_n a_{hn}|^\nu \leqq K \max_{1\leqq n\leqq N} |a_n|^\nu b_{\nu hN} , \qquad 1 < \nu < 2 . \tag{3.20}$$

Then the theorem is proved under (3.16). From Theorem 3, (3.19) is of smaller order than (3.20) as $N \to \infty$ under uniform integrability of $|v_{jn}/\alpha_j a_n|^\nu$ and (3.17), so that theorem is proved under these conditions. From (3.15) and Lemma 2, (3.19) is bounded by

$$K(\sum_{n=1}^{N} a_n{}^2 a_{hn}^2)^{\nu/2} < K \max_{1 \leqq n \leqq N} |a_n|^\nu b_{2hN}^{\nu/2} .$$

Then because of (3.18), the theorem is proved for $\nu \geqq 2$. □

## REFERENCES


[1] ANDERSON, T. W. and TAYLOR, J. B. (1976). Strong consistency of least squares estimates in normal linear regression. *Ann. Statist.* **4** 788–790.

[2] BILLINGSLEY, P. (1968). *Convergence of Probability Measures.* Wiley, New York.

[3] BURKHOLDER, D. L. (1966). Martingale transforms. *Ann. Math. Statist.* **37** 1494–1504.

[4] CHOW, Y. S. (1971). On the $L_p$-convergence for $n^{-1/p}S_n$, $0 < p < 2$. *Ann. Math. Statist.* **42** 393–394.

[5] HANNAN, E. J. and HEYDE, C. C. (1972). On limit theorems for quadratic functions of discrete time series. *Ann. Math. Statist.* **43** 2058–2066.

[6] LOÈVE, M. (1951). On almost sure convergence. *Proc. Second Berkeley Symp. Math. Statist. Prob.* **1** 279–304, Univ. of California Press.

[7] NEVEU, J. (1965). *Mathematical Foundations of the Calculus of Probability.* Holden-Day, San Francisco.

[8] ROBINSON, P. M. (1977). Distributed lag approximation to linear time invariant systems. *Ann. Statist.*, forthcoming.

[9] ROBINSON, P. M. (1977). Estimation of a time series model from unequally spaced data. *Stoch. Proc. Appl.*, forthcoming.

[10] SERFLING, R. J. (1970). Moment inequalities for the maximum cumulative sum. *Ann. Math. Statist.* **41** 1227–1234.

[11] SERFLING, R. J. (1970). Convergence properties of $S_N$ under moment restrictions. *Ann. Math. Statist.* **41** 1235–1248.

[12] STOUT, W. F. (1974). *Almost Sure Convergence.* Academic Press, New York.

[13] VON BAHR, B. and ESSEEN, C.-G. (1965). Inequalities for the $r$th absolute moment of a sum of random variables $1 \leqq r \leqq 2$. *Ann. Math. Statist.* **36** 299–303.



DEPARTMENT OF STATISTICS
HARVARD UNIVERSITY
CAMBRIDGE, MASSACHUSETTS 02138