ELSEVIER

Statistics & Probability Letters 27 (1996) 1–9

STATISTICS & PROBABILITY LETTERS

A strong law of large numbers for triangular mixingale arrays

Robert M. de Jong[1]

Department of Econometrics, Tilburg University, P.O. Box 90153, 5000 LE Tilburg, Netherlands

Received September 1994; revised January 1995

Abstract

In this paper a strong law of large numbers for triangular mixingale arrays is proven. The mixingale condition is one of asymptotically weak dependence. A strong law of large numbers for triangular mixingale arrays has not been established previously in the literature. The result is applied to kernel regression function estimation.

Keywords: Strong law of large numbers; Mixingale sequence; Law of large numbers; Triangular array; Dependent random variables; Kernel regression

1. Introduction

In McLeish (1975) the first strong law of large numbers for so-called mixingale sequences was established. The mixingale concept is one of asymptotically weak dependence. The mixingale concept is elaborated upon in e.g. Gallant and White (1988), Pötscher and Prucha (1991) and Davidson (1994), among others. In asymptotic theory for dependent processes (such as mixing processes or near epoch dependent processes), many strong and weak law results are proven in the following way: firstly, it is shown that a particular process at hand is a mixingale sequence; secondly, a strong or weak law result for mixingale sequences is applied. Strong laws for mixingale sequences are important tools for showing strong consistency of estimators. For a treatment of such consistency results we refer to Gallant and White (1988) and Pötscher and Prucha (1991).

Andrews (1988), Davidson (1993) and De Jong (1995) provide weak laws of large numbers for mixingale sequences. Typically, these weak law results are easily converted to a result that holds for triangular mixingale arrays. Hansen (1991, 1992) and De Jong (1995) recently proved strong laws for mixingale sequences that are complementary to the strong law of McLeish (1975), while De Jong (1994) provides an improved version of the strong law of McLeish (1975). A corollary of De Jong (1995) shows the interesting result that a logarithmic decay rate for mixingale numbers can in some cases be sufficient for establishing a strong law of large numbers. Strong law results for triangular arrays of random variables have been considered by Teicher (1985). The technique of proof of strong laws of large numbers in many cases is such that these results are not easily generalized towards triangular array results. Strong law results for triangular arrays can for example

[1] E-mail: rdejong@kub.nl.

SSDI 0167-7152(95)00036-4

be important for showing local power results and for showing strong convergence of kernel estimators. In this paper we will set out to prove a strong law of large numbers for triangular mixingale arrays.

The plan of this paper is as follows. Section 2 will provide various definitions and it will summarize the strong law results for mixingale sequences that are available. In Section 3 we will state our strong law of large numbers for triangular mixingale arrays. Section 4 contains an application of our central result to kernel regression function estimators. The proof of this result will be given in the Appendix.

2. Mixingale sequences and mixingale strong laws

Let (Ω,F,P) denote a probability space. Let $\{X_{nt} : n,t \geqslant 1\}$ denote a triangular array of random variables on (Ω,F,P). Let F_t^n denote sequence of sub sigma-fields of F such that $\{F_t^n\}$ for each $n \geqslant 1$ is nondecreasing in t. Whenever an expression like $E(X_{nt}|F_j^n)$ is used it is assumed that $\|X_{nt}\|_1$ is finite, where $\|X\|_q$, for $q \geqslant 1$, is defined as $(E|X|^q)^{1/q}$. Andrews' definition of a triangular L_q-mixingale array is as follows:

Definition 1. The triangular array $\{X_{nt},F_t^n\}$ is a triangular L_q-mixingale array, $q \geqslant 1$, if there exist nonnegative constants $\{c_{nt} : t \geqslant 1\}$ and $\{\psi(m),\ m \geqslant 0\}$ such that $\psi(m) \to 0$ as $m \to \infty$ and for all $t \geqslant 1$ and $m \geqslant 0$ we have

$$\|E(X_{nt}|F_{t-m}^n)\|_q \leqslant c_{nt}\psi(m) \tag{1}$$

and

$$\|X_{nt} - E(X_{nt}|F_{t+m}^n)\|_q \leqslant c_{nt}\psi(m+1). \tag{2}$$

Note that this definition defines triangular mixingale arrays as triangular arrays of mean zero random variables. Often one will want to choose $F_t^n = \sigma(X_{nt}, X_{n,t-1}, \ldots)$; however, we will not impose this. In what follows we will use the definition of the *size* of a sequence. A sequence $\psi(m)$ is said to be of size $-\beta$, $\beta > 0$, if $\psi(m) = \mathrm{O}(m^{-\beta-\varepsilon})$ for some $\varepsilon > 0$. For mixingale sequences a number of strong law results is available. We mention those results below. The oldest strong law result for averages of mixingale sequences is one by McLeish (1975). We will not state this result here since it is contained in the result by De Jong (1994) (Theorem 3 below). Hansen (1991, 1992) developed the following strong law result.

Theorem 1 (Hansen (1991, 1992)). *Let $\{X_t,F_t\}$ be an L_q-mixingale sequence for some $q > 1$ such that $\sum_{t=1}^{\infty} c_t^q t^{-q} < \infty$. Assume $\sum_{m=1}^{\infty} \psi(m) < \infty$. Then*

$$n^{-1}\sum_{t=1}^{n} X_t \to 0 \tag{3}$$

almost surely.

Corollary 1 of De Jong (1995) provides the following result.

Theorem 2 (De Jong (1995)). *Let the sequence $\{X_t,F_t\}$ be an L_q-mixingale sequence for some $q > 1$. Suppose $\sup_{t \geqslant 1} |c_t| < \infty$ and $\psi(m) = \mathrm{O}((\log m)^{-1-\varepsilon})$ as $m \to \infty$ for some $\varepsilon > 0$. Then*

$$n^{-1}\sum_{t=1}^{n} X_t \to 0 \tag{4}$$

almost surely.

Although the original theorem in De Jong (1995) is slightly more general than Theorem 2, the conditions on the c_t sequence that it allows for are easily seen to be relatively strict. The strong law result of De Jong (1994) that contains McLeish's (1975) result is the following.

Theorem 3. *Let the sequence $\{X_t, F_t\}$ be an L_2-mixingale sequence such that, for some $\varepsilon > 0$,*

$$\sum_{t=1}^{\infty} c_t^2 t^{-2} + \sum_{t=1}^{\infty} c_t^2 t^{-2} \sum_{j=0}^{t} \psi(j+1)^2 (\log(j+1))^{1+\varepsilon}$$
$$= \sum_{t=1}^{\infty} c_t^2 t^{-2} + \sum_{j=0}^{\infty} \psi(j+1)^2 (\log(j+1))^{1+\varepsilon} \sum_{t=\max(1,j)}^{\infty} c_t^2 t^{-2} < \infty. \tag{5}$$

Then

$$n^{-1} \sum_{t=1}^{n} X_t \to 0 \tag{6}$$

almost surely.

3. A strong law for mixingale arrays

The lines of proof used in Hansen (1991, 1992) and De Jong (1994, 1995) are not suitable for rendering a proof of a strong law result. In those papers, use is made of Cauchy criterion in combination with the Kronecker lemma. Since this argument cannot be used in the same way in the triangular array case, we had to apply the Borel–Cantelli lemma in order to obtain a strong law for triangular mixingale arrays. Therefore, we set forth conditions that guarantee that for all $\delta > 0$

$$\sum_{n=1}^{\infty} P\left[\left| n^{-1} \sum_{t=1}^{n} X_{nt} \right| > \delta \right] < \infty \tag{7}$$

for a triangular L_q-mixingale array X_{nt}. The result is as follows:

Theorem 4. *Suppose that the triangular array $\{X_{nt}, F_t^n\}$ is a triangular L_q-mixingale array for some $q \geqslant 1$. If for a deterministic triangular array B_{nt} and a positive integer-valued sequence m_n we have*

1. *One of the following conditions holds:*
   (a) $\sum_{n=1}^{\infty} n^{-1} \sum_{t=1}^{n} c_{nt}^q B_{nt}^{1-q} < \infty$; *or*
   (b) $|X_{nt}| \leqslant B_{nt}$ *almost surely for all* $t \leqslant n$; *or*
   (c) $\sum_{n=1}^{\infty} n^{-1} \sum_{t=1}^{n} E|X_{nt}|^r B_{nt}^{1-r} < \infty$ *for some* $r \geqslant 1$.
2. $\sum_{n=1}^{\infty} \left(n^{-1} \sum_{t=1}^{n} c_{nt} \psi(m_n) \right)^q < \infty$.
3. *For all* $\delta > 0$, $\sum_{n=1}^{\infty} m_n \exp\left(-\delta^2 n^2 m_n^{-2} \left(\sum_{t=1}^{n} B_{nt}^2 \right)^{-1} \right) < \infty$.

Then

$$n^{-1} \sum_{t=1}^{n} X_{nt} \to 0 \tag{8}$$

almost surely.

As an example consider a triangular L_2-mixingale with $c_{nt} = 1$ for all t, $n \geqslant 1$ and $\sup_{t,n\geqslant 1} EX_{nt}^4 < \infty$. Take $r = 4$ for the verification of Assumption 1c of Theorem 4; we have to choose $B_{nt} = n^{1/3+\eta}$, for arbitrary $\eta > 0$, in order to verify this assumption. Assumption 3 of Theorem 4 can then be verified by choosing $m_n = [n^{1/6-2\eta}]$. A strong law of large numbers can then be obtained from verification of Assumption 2 of Theorem 4 by imposing that $\psi(m)$ is of size -3. Note that this result is much more restrictive than the results for mixingale sequences as listed in Section 2.

4. Application to kernel regression estimators

In this section, we will apply Theorem 4 to kernel regression function estimators in order to obtain a pointwise strong convergence result for kernel estimators for the case of dependent random variables. The most common dependence concepts for random variables are the strong and uniform mixing conditions. For a definition of these dependence concepts see e.g. Davidson (1994, p. 209). The kernel regression function estimator $\hat{g}_n(x)$ is defined as

$$\hat{g}_n(x) = \frac{\sum_{t=1}^{n} Y_t k\left(\frac{x-X_t}{\gamma_n}\right)}{\sum_{t=1}^{n} k\left(\frac{x-X_t}{\gamma_n}\right)}. \tag{9}$$

The kernel estimator is a form of nonparametric estimation; for various types of convergence (e.g. pointwise), it can be shown that $\hat{g}_n(x) \to E(Y_t|X_t = x)$. We will assume that $X_t \in \mathbb{R}^p$. Note that standard results on kernel regression function estimation show that

$$\gamma_n^{-p} EY_t k\left(\frac{x-X_t}{\gamma_n}\right) \to E(Y_t|X_t = x). \tag{10}$$

See e.g. Bierens (1994). Therefore, we will consider the stochastic part of the problem, i.e. we will set forth conditions such that

$$n^{-1}\sum_{t=1}^{n} \gamma_n^{-p} Y_t k\left(\frac{x-X_t}{\gamma_n}\right) - E\gamma_n^{-p} Y_t k\left(\frac{x-X_t}{\gamma_n}\right) \to 0 \tag{11}$$

almost surely. Note that such a result also can be applied to the denominator in the formula of $\hat{g}_n(x)$ by setting $Y_t = 1$. Convergence and asymptotic normality results for kernel regression function estimators have been considered by Robinson (1983) and Bierens (1983, 1994). Robinson (1983) shows strong consistency and asymptotic normality using the strong mixing concept. Bierens (1983, 1994) are the only references that this author is aware of that provides convergence results for the kernel regression function estimator under dependence conditions that are weaker than mixing. Bierens (1983, 1994) shows weak consistency and asymptotic normality using the concept of v-stable processes. A related concept of weak dependence that is used more frequently in the literature is the concept of near epoch dependence. This dependence concept is introduced in order to obtain convergence results for general ARMA processes. In general, $AR(\infty)$ processes are not strong mixing or uniform mixing, but in general these processes do satisfy the near epoch dependence condition. For discussions of near epoch random variables see Pötscher and Prucha (1991) and Davidson (1994). Following Andrews (1988), we will define L_q-near epoch dependent random variables in the following way.

Definition 2. The triangular array of random variables X_{nt}, $t = 1, \ldots$ is called L_q-near epoch dependent, $q \geqslant 1$, on V_{nt} if for $m \geqslant 0$

$$\|X_{nt} - E(X_{nt} | V_{n,t-m}, \ldots, V_{n,t+m})\|_q \leqslant d_{nt} v(m) \tag{12}$$

and $v(m) \to 0$ as $m \to \infty$.

Typically we have to assume that the V_t process from Definition 2 satisfies some mixing condition in order to obtain useful results. In order to apply Theorem 4 to the nominator and the denominator of the kernel regression function estimator we first need to show that $Y_t k((x - X_t)/\gamma_n)$ is an L_1-mixingale:

Lemma 1. *If (X_t, Y_t) is L_2-near epoch dependent on a strong mixing sequence V_t and $k(.)$ is a uniformly bounded function such that*

$$|k(x) - k(y)| \leqslant L|x - y| \tag{13}$$

then $Y_t k((x - X_t)/\gamma_n)$ is an L_1-mixingale with

$$c_{nt} = C\left(d_t + d_t \,\|Y_t\|_2\, \gamma_n^{-1} + \|Y_t\|_r\right) \tag{14}$$

and

$$\psi(m) = v([m/2]) + 6\alpha([m/2])^{1-1/r}. \tag{15}$$

The following theorem shows strong convergence of numerator and the denominator of the kernel regression function estimator to its expected value.

Theorem 5. *Assume that the uniformly bounded function $k(.)$ satisfies*

$$|k(x) - k(y)| \leqslant L|x - y|. \tag{16}$$

If (X_t, Y_t) is near epoch dependent on a strong mixing sequence V_t and for some $r > 2$

$$\sup_{n \geqslant 1} n^{-1} \sum_{t=1}^{n} E|Y_t|^r < \infty, \tag{17}$$

$$\sum_{n=1}^{\infty} \gamma_n^{-p} n^{-1} \sum_{t=1}^{n} \left(d_t + \gamma_n^{-1} \,\|Y_t\|_2\, d_t + 1\right)\left(v([m_n/2]) + \alpha([m_n/2])^{1-1/r}\right) < \infty \tag{18}$$

where

$$m_n = \left[n^{1/2-\beta/2} n^{-(1+\varepsilon)/(r-1)} \gamma_n^{pr/(r-1)}\right] \tag{19}$$

for some $\beta, \varepsilon > 0$ then

$$n^{-1} \sum_{t=1}^{n} Y_t \gamma_n^{-p} k\left(\frac{x - X_t}{\gamma_n}\right) - E Y_t \gamma_n^{-p} k\left(\frac{x - X_t}{\gamma_n}\right) \to 0 \tag{20}$$

almost surely.

Proof. See Appendix. □

Appendix. Proof of the theorems and lemma

Proof of Theorem 4. Note that for all deterministic triangular arrays $B_{nt} > 0$ and integer-valued sequences $m_n \geqslant 1$

$$\begin{aligned}
n^{-1}\sum_{t=1}^{n} X_{nt} &= n^{-1}\sum_{t=1}^{n}\left(X_{nt} - E\left(X_{nt}|F^n_{t+m_n-1}\right)\right) \\
&+n^{-1}\sum_{t=1}^{n}\left(E(X_{nt}I\left(|X_{nt}|\leqslant B_{nt}\right)|F^n_{t+m_n-1}) - E\left(X_{nt}I(|X_{nt}|\leqslant B_{nt})|F^n_{t-m_n}\right)\right) \\
&+n^{-1}\sum_{t=1}^{n}\left(E\left(X_{nt}I(|X_{nt}| > B_{nt})|F^n_{t+m_n-1}\right) - E\left(X_{nt}I(|X_{nt}| > B_{nt})|F^n_{t-m_n}\right)\right) \\
&+n^{-1}\sum_{t=1}^{n}E\left(X_{nt}|F^n_{t-m_n}\right) \\
&= T_1 + T_2 + T_3 + T_4.
\end{aligned} \tag{A.1}$$

Note that the result follows from the Borel–Cantelli lemma if we show that for all $\delta > 0$ and for $j = 1,\ldots,4$,

$$\sum_{n=1}^{\infty} P[|T_j| > \delta] < \infty.$$

Next, note that if Assumption 1 of Theorem 4 holds then $T_3 \to 0$ almost surely since, by the Markov inequality,

$$\begin{aligned}
&\sum_{n=1}^{\infty} P\left[|n^{-1}\sum_{t=1}^{n} E(X_{nt}I(|X_{nt}| > B_{nt})|F^n_{t+m_n-1}) - E(X_{nt}I(|X_{nt}| > B_{nt})|F^n_{t-m_n})| > \delta\right] \\
&\quad\leqslant \sum_{n=1}^{\infty} n^{-1}\sum_{t=1}^{n} 2\delta^{-1}E|X_{nt}|I(|X_{nt}| > B_{nt}) \\
&\quad\leqslant 2\delta^{-1}\sum_{n=1}^{\infty} n^{-1}\sum_{t=1}^{n} E|X_{nt}|^r B_{nt}^{1-r} < \infty
\end{aligned} \tag{A.2}$$

or

$$\leqslant 2\delta^{-1}\sum_{n=1}^{\infty} n^{-1}\sum_{t=1}^{n}(\psi(0)+\psi(1))^q c_{nt}^q B_{nt}^{1-q} < \infty \tag{A.3}$$

by the mixingale assumption. The case of Assumption 1c is trivial. Assumption 2 ensures that $T_4 \to 0$ and $T_1 \to 0$ almost surely since, by the Markov inequality

$$\begin{aligned}
&\sum_{n=1}^{\infty} P\left[\left|n^{-1}\sum_{t=1}^{n} E(X_{nt}|F^n_{t-m_n})\right| > \delta\right] \\
&\quad\leqslant \sum_{n=1}^{\infty} \delta^{-p}\left(\left\|n^{-1}\sum_{t=1}^{n} E(X_{nt}|F^n_{t-m_n})\right\|_q\right)^q
\end{aligned}$$

$$
\begin{aligned}
&\leqslant \sum_{n=1}^{\infty} \delta^{-q} \left(n^{-1} \sum_{t=1}^{n} \left\| E(X_{nt} | F_{t-m_n}^n) \right\|_q \right)^q \\
&\leqslant \sum_{n=1}^{\infty} \delta^{-q} \left(n^{-1} \sum_{t=1}^{n} c_{nt} \psi(m_n) \right)^q < \infty
\end{aligned}
\tag{A.4}
$$

by assumption. The same argument holds for T_1. Next, note that Azuma's (1967) inequality as quoted by Davidson (1994, Eq. (15.28)) states that for martingale difference sequences z_t such that $|z_t| \leqslant B_{nt}$ for all $1 \leqslant t \leqslant n$ we have

$$
P\left[\left| \sum_{t=1}^{n} z_t \right| > \delta \right] \leqslant 2 \exp\left(-\delta^2 \left(\sum_{t=1}^{n} B_{nt}^2 \right)^{-1} 2^{-1} \right). \tag{A.5}
$$

Assumption 3 then ensures that $T_2 \to 0$ almost surely since

$$
\begin{aligned}
&\sum_{n=1}^{\infty} P\left[\left| n^{-1} \sum_{t=1}^{n} E(X_{nt} I(|X_{nt}| \leqslant B_{nt}) | F_{t+m_n-1}^n) - E(X_{nt} I(|X_{nt}| \leqslant B_{nt}) | F_{t-m_n}^n) \right| > \delta \right] \\
&\leqslant \sum_{n=1}^{\infty} \sum_{j=-m_n+1}^{m_n-1} P\left[\left| n^{-1} \sum_{t=1}^{n} (E(X_{nt} I(|X_{nt}| \leqslant B_{nt}) | F_{t+j}^n) - E(X_{nt} I(|X_{nt}| \leqslant B_{nt}) | F_{t+j-1}^n)) \right| > \delta 2^{-1} m_n^{-1} \right] \\
&\leqslant \sum_{n=1}^{\infty} 2(2m_n) \exp\left(-(1/2)(\delta/(2m_n))^2 n^2 \left(\sum_{t=1}^{n} (2B_{nt})^2 \right)^{-1} \right) \\
&\leqslant \sum_{n=1}^{\infty} 4m_n \exp\left(-\delta^2 n^2 m_n^{-2} 32^{-1} \left(\sum_{t=1}^{n} B_{nt}^2 \right)^{-1} \right) < \infty
\end{aligned}
\tag{A.6}
$$

by assumption, where the second inequality follows by Azuma's inequality and by noting that $|E(X_{nt} I(|X_{nt}| \leqslant B_{nt}) | F_{t+j}^n) - E(X_{nt} I(|X_{nt}| \leqslant B_{nt}) | F_{t+j-1}^n)| \leqslant 2B_{nt}$.

Proof of Lemma 1. Define $H_{t-m}^{t+m} = \sigma(V_{t-m}, \ldots, V_{t+m})$. Firstly note that $Y_t k((x - X_t)/\gamma_n)$ is L_1-near epoch dependent on V_{nt} since

$$
\begin{aligned}
&\left\| Y_t k\left(\frac{x - X_t}{\gamma_n} \right) - E\left(Y_t k\left(\frac{x - X_t}{\gamma_n} \right) \middle| H_{t-m}^{t+m} \right) \right\|_1 \\
&\leqslant \left\| Y_t k\left(\frac{x - X_t}{\gamma_n} \right) - E(Y_t | H_{t-m}^{t+m}) k\left(\frac{x - X_t}{\gamma_n} \right) \right\|_1 \\
&+ \left\| E(Y_t | H_{t-m}^{t+m}) k\left(\frac{x - X_t}{\gamma_n} \right) - E(Y_t | H_{t-m}^{t+m}) E\left(k\left(\frac{x - X_t}{\gamma_n} \right) \middle| H_{t-m}^{t+m} \right) \right\|_1 \\
&+ \left\| E(Y_t | H_{t-m}^{t+m}) E\left(k\left(\frac{x - X_t}{\gamma_n} \right) \middle| H_{t-m}^{t+m} \right) - E\left(Y_t k\left(\frac{x - X_t}{\gamma_n} \right) \middle| H_{t-m}^{t+m} \right) \right\|_1 \\
&\leqslant C' \| Y_t - E(Y_t | H_{t-m}^{t+m}) \|_1 + \| Y_t \|_2 \left\| k\left(\frac{x - X_t}{\gamma_n} \right) - E\left(k\left(\frac{x - X_t}{\gamma_n} \right) \middle| H_{t-m}^{t+m} \right) \right\|_2 \\
&+ \left\| k\left(\frac{x - X_t}{\gamma_n} \right) E(Y_t | H_{t-m}^{t+m}) - Y_t k\left(\frac{x - X_t}{\gamma_n} \right) \right\|_1
\end{aligned}
$$

$$\leqslant C''d_t v(m) + \|Y_t\|_2 \left\| k\left(\frac{x-X_t}{\gamma_n}\right) - k\left(\frac{x-E(X_t|H_{t-m}^{t+m})}{\gamma_n}\right)\right\|_2$$
$$\leqslant C''d_t v(m) + \|Y_t\|_2 L\gamma_n^{-1} \|X_t - E(X_t|H_{t-m}^{t+m}))\|_2$$
$$\leqslant (C'' + L\|Y_t\|_2 \gamma_n^{-1}) d_t v(m). \tag{A.7}$$

The second inequality follows from the uniform boundedness assumption on the kernel function, the Cauchy–Schwartz inequality and the inequality $\|E(Y|G)\|_1 \leqslant \|Y\|_1$, the third uses the fact that the Borel measurable function $g(X)$ minimizing $\|Y - g(X)\|_2$ is $E(Y|X)$, and the fourth uses the Lipschitz-continuity assumption on the kernel function. Using a result from Andrews (1988), it is now easily seen that $Y_t k((x-X_t)/\gamma_n)$ is an L_1-mixingale with

$$c_{nt} = C(d_t + d_t \|Y_t\|_2 \gamma_n^{-1} + \|Y_t\|_r) \tag{A.8}$$
$$\psi(m) = v([m/2]) + 6\alpha([m/2])^{1-1/r} \tag{A.9}$$

for some $r > 1$.

Proof of Theorem 2. Theorem 2 is proven by choosing $q = 1$,

$$X_{nt} = \gamma_n^{-p} Y_t k\left(\frac{x-X_t}{\gamma_n}\right),$$
$$B_{nt} = B_n = n^{(1+\varepsilon)/(r-1)}\gamma_n^{-rp/(r-1)}$$
$$m_n = [n^{1/2-\beta/2} n^{-(1+\varepsilon)/(r-1)} \gamma_n^{pr/(r-1)}]$$

where $\beta, \varepsilon > 0$. Note that in order for m_n to be not converging to zero we need $r > 2$. Then condition 1c of Theorem 4 is satisfied since

$$\sum_{n=1}^{\infty} n^{-1} \sum_{t=1}^{n} E|X_{nt}|^r B_n^{1-r} \leqslant C \sum_{n=1}^{\infty} \gamma_n^{-pr} n^{-1-\varepsilon} \gamma_n^{rp} n^{-1} \sum_{t=1}^{n} E|Y_t|^r \leqslant C \sum_{n=1}^{\infty} n^{-1-\varepsilon} < \infty. \tag{A.10}$$

Condition 3 of Theorem 4 is satisfied if

$$m_n^2 B_n^2 \leqslant n^{1-\beta}. \tag{A.11}$$

This is the case since

$$m_n^2 B_n^2 \leqslant n^{1-\beta} n^{-2(1+\varepsilon)/(r-1)} \gamma_n^{2pr/(r-1)} n^{2(1+\varepsilon)/(r-1)} \gamma_n^{-2rp/(r-1)} = n^{1-\beta}. \tag{A.12}$$

Condition 2 of Theorem 4 is the key condition from Theorem 5 with the B_n and m_n sequence as specified and the c_{nt} and $\psi(m)$ sequences from Lemma 1.

References

Andrews, D.W.K. (1988), Laws of large numbers for dependent non-identically distributed random variables, *Econometric Theory* **4**, 458–467.

Azuma, K. (1967), Weighted sums of certain dependent random variables, *Tokohu Math. J.* **19**, 357–367.

Bierens, H.J. (1983), Uniform consistency of kernel regression estimators of a regression function under generalized conditions, *J. Amer. Statist. Assoc.* **78**, 699–707.

Bierens, H.J. (1994), *Topics in Advanced Econometrics* (Cambridge University Press, Cambridge).

Chung, K.L. (1974), *A Course in Probability Theory* (Academic Press, New York).

Davidson, J. (1993), An L_1-convergence theorem for heterogeneous mixingale arrays with trending moments, *Statist. Probab. Lett.* **16**, 301–306.

Davidson, J. (1994), *Stochastic Limit Theory* (Cambridge University Press, Cambridge).

De Jong, R.M. (1994), A strong law for L_2-mixingale sequences, Working paper.

De Jong, R.M. (1995), Laws of large numbers for dependent heterogeneous processes, Working paper.

Gallant, A.R. and H. White (1988), *A Unified Theory of Estimation and Inference for Nonlinear Dynamic Models* (Basil Blackwell, New York).

Hall, P. and C.C. Heyde (1980), *Martingale Limit Theory and Its Applications* (Academic Press, New York).

Hansen, B.E. (1991), Strong laws for dependent heterogeneous processes, *Econometric Theory* **7**, 213–221.

Hansen, B.E. (1992), Strong laws for dependent heterogeneous processes. Erratum. *Econometric Theory* **8**, 421–422.

McLeish, D.L. (1975), A maximal inequality and dependent strong laws, *Ann. Probab.* **3**, 829–839.

Pötscher , B.M. and I.R. Prucha (1991), Basic structure of the asymptotic theory in dynamic nonlinear econometric models, part 1: consistency and approximation concepts, *Econometric Rev.* **10**, 125–216.

Robinson, P.M. (1983), Nonparametric estimators for time series, *J. Time Ser. Anal.* **4**, 185–207.

Teicher, H. (1985), Almost certain convergence in double arrays, *Zeitschrift für Wahrscheinlichkeitstheorie und verwandte Gebiete* **69**, 331–345.