Laws of Large Numbers for Dependent Heterogeneous Processes
Author(s): R. M. de Jong
Source: *Econometric Theory*, Vol. 11, No. 2 (Jun., 1995), pp. 347-358
Published by: Cambridge University Press
Stable URL: http://www.jstor.org/stable/3532577
Accessed: 16-03-2016 12:24 UTC

Econometric Theory, **11**, 1995, 347–358. Printed in the United States of America.

LAWS OF LARGE NUMBERS FOR DEPENDENT HETEROGENEOUS PROCESSES

R.M. DE JONG
Free University Amsterdam

This paper provides weak and strong laws of large numbers for weakly dependent heterogeneous random variables. The weak laws of large numbers presented extend known results to the case of trended random variables. The main feature of our strong law of large numbers for mixingale sequences is the less strict decay rate that is imposed on the mixingale numbers as compared to previous results.

1. INTRODUCTION

Laws of large numbers (LLN's) for mixingale sequences are an essential tool in the proofs of consistency of estimators when dependent processes are considered. A mixingale sequence can be viewed as an asymptotic equivalent of a martingale difference sequence. In a recent paper, Andrews (1988) extended the mixingale concept and established weak L_1-type LLN's for mixingales. Andrews' work extends the results of McLeish (1975), who introduced the mixingale condition. In Section 2 of this paper, Andrews' conditions for an L_1-type LLN for L_p-mixingale sequences will be extended such as to cover the case of trended mixingale sequences. Recently, Hansen (1991, 1992) proved some new strong LLN's using Andrews' L_p-mixingale concept. Those results extend the results of McLeish. In Section 3 of this paper, the conditions of both McLeish and Hansen for a strong LLN to hold will be relaxed substantially for the important case of L_p-mixingales with uniformly bounded pth moments and bounded indices of magnitude. The results obtained in Section 3 are complementary to those obtained by Hansen and McLeish since our results will not be powerful in the case of increasing indices of magnitude or the case that no uniform moment bound exists. In Section 4, we apply our results to near epoch dependent sequences. Section 5 concludes this paper with the proofs of the theorems.

I thank Dr. P. Spreij and two anonymous referees for their valuable comments. The first version of this paper appeared as Free University Department of Economics and Econometrics Research Memorandum 1991-88. This paper was presented at the ESEM conference, August 1992, Brussels, Belgium. This research corresponds partly to Chapter 3 of my Ph.D. thesis.

2. PRELIMINARIES AND A WEAK LLN FOR MIXINGALES

Let (Ω, F, P) denote a probability space. Let $\{X_i : i \geq 1\}$ be a sequence of random variables on (Ω, F, P). Let $\{F_i : i = \ldots, 0, 1, \ldots,\}$ be any nondecreasing sequence of sub σ-fields of F. Often one will take $F_i = \sigma(X_1, \ldots, X_i)$ for $i \geq 1$ and $F_i = \{\emptyset, \Omega\}$ for $i \leq 0$. $E(X_i | F_j)$ denotes the conditional expectation of X_i given F_j. Whenever $E(X_i | F_j)$ is used, we assume $E|X_i|$ to be finite. Let $\|X\|_p$ denote $(E|X|^p)^{1/p}$. Andrews (1988) defined an L_p-mixingale as follows.

DEFINITION 1. *The sequence* $\{X_i, F_i\}$ *is called an* L_p*-mixingale,* $p \geq 1$, *if there exist nonnegative constants* $\{c_i : i \geq 1\}$ *and* $\{\psi(m) : m \geq 0\}$ *such that* $\psi(m) \to 0$ *as* $m \to \infty$ *and for all* $i \geq 0$ *and* $m \geq 0$ *we have*

(A) $\|E(X_i | F_{i-m})\|_p \leq c_i \psi(m)$ *and*
(B) $\|X_i - E(X_i | F_{i+m})\|_p \leq c_i \psi(m+1)$.

The mixingale concept is one of weak dependence. The mixingale property is strong enough for proofs of weak and strong LLN's. However, the mixingale property is not preserved under many transformations, and as a consequence in the literature the mixingale concept is mainly used as a tool for showing that a LLN holds. Note that according to Definition 1 L_p-mixingales are necessarily mean zero random variables. The weak LLN of Andrews (1988) typically requires

$$\limsup_{n \to \infty} n^{-1} \sum_{i=1}^{n} c_i < \infty$$

and no restriction on the $\psi(m)$ sequence, except for the condition that $\psi(m) \to 0$ as $m \to \infty$, for a weak L_1-type LLN to hold. In addition, X_i is assumed to be uniformly integrable. Because the mixingale magnitude indices c_i can in many cases be assumed bounded, e.g., by $\sup_{i \geq 1} E|X_i|$ as in Andrews (1988, Theorem 1b), in many cases those assumptions will be reasonable ones. On the other hand, even for very slowly increasing c_i sequences, Andrews' result is not valid. Increasing c_i sequences can occur in the case of trending moments of the X_i. For example, consider an L_2-mixingale sequence $Z_i = i^{1/12} W_i$, where W_i is a nondegenerate L_2-mixingale sequence with uniformly bounded mixingale magnitude index series c_i. From the mixingale definition, it is easily seen that the Z_i sequence needs to have a mixingale magnitude index series of the form $Ci^{1/12}$, for some constant $C > 0$. Unlike the weak LLN for L_p-mixingales of Andrews (1988), existing strong LLN's for L_p-mixingale sequences (see Section 3) do allow for increasing c_i sequences. In the case of an i.i.d. W_i sequence, a weak law for the Z_i sequence is easily shown by an elementary variance calculation (and noting that W_1 possesses a finite second moment due to the L_2-mixingale condition), without imposing the uniform integrability condition. This motivates finding weak LLN's for possibly trended mixingale sequences with possibly increasing c_i sequences.

The following theorem provides a result that is capable of showing that only a tradeoff condition between the rate of decay of the mixingale numbers and the rate of increase of the average of the L_p-mixingale magnitude indices is required for a LLN to hold.

THEOREM 1. *Suppose the sequence* $\{X_i, F_i\}$ *is an* L_p*-mixingale*, $1 \leq p \leq 2$, *such that for some sequence* $C_n \geq 1$, $C_n = o(n^{1/2})$,

(A) $\lim_{B\to\infty} \limsup_{n\to\infty} n^{-1} \sum_{i=1}^{n} \| X_i I(|X_i| > BC_n) \|_p = 0$ *and*
(B) *for all* $K > 0$, $\lim_{n\to\infty} n^{-1} \sum_{i=1}^{n} c_i \psi([Kn^{1/2} C_n^{-1}]) = 0$.

Then,

$$\left\| n^{-1} \sum_{i=1}^{n} X_i \right\|_p \to 0$$

as $n \to \infty$ $\left(\textit{and therefore } |n^{-1} \sum_{i=1}^{n} X_i| \to 0 \textit{ in prob.}\right)$

Proof. See Section 5.

Note that by the Hölder and Markov inequalities, (A) follows from the condition $\limsup_{n\to\infty} C_n^{1-r/p} n^{-1} \sum_{i=1}^{n} \|X_i\|_r^{r/p} < \infty$ for some $r > p$. For $C_n = 1$, a condition similar to (A) can be found in Gut (1992) and Pötscher and Prucha (1991). Also, for the choice $C_n = 1$, (A) follows from uniform integrability of $|X_i|^p$. The case of a trended X_i sequence will typically correspond to choosing an increasing C_n sequence. A weak L_1-type law for the Z_i sequence described earlier can be established by noting that conditions (A) and (B) of Theorem 1 can be verified by assuming $\psi(m) = o(m^{-1/4})$ and choosing $C_n = n^{1/6}$. Note that Andrews' (1988) weak LLN is easily seen to be a special case of Theorem 1. Also note that the results of this section hold if triangular arrays X_{ni} are considered similarly to those of Andrews (1988).

3. A STRONG LLN FOR MIXINGALE SEQUENCES

Strong LLN's for mixingales were elaborated upon by McLeish (1975) and Hansen (1991, 1992). Their approach is proving almost sure convergence of

$$\sum_{i=1}^{n} X_i$$

under some conditions, and they obtain an almost sure law for sample averages by imposing those conditions on $X_i i^{-1}$ and concluding that

$$\sum_{i=1}^{n} X_i i^{-1}$$

converges almost surely to some random variable. The almost sure law for sample averages now follows by the Kronecker Lemma; i.e., if a_i is a sequence of positive real numbers that is nondecreasing in i and $a_i \to \infty$ if $i \to \infty$,

$$\sum_{i=1}^{n} X_i a_i^{-1} < \infty \Rightarrow a_n^{-1} \sum_{i=1}^{n} X_i \to 0.$$

(see, e.g., Chung, 1974, p. 123). For L_p-mixingales, this approach typically results in conditions on the c_i sequence of the type

$$\sum_{i=1}^{\infty} c_i^p i^{-p} < \infty$$

as imposed by Hansen (1991, 1992) and McLeish (1975). This requirement allows for increasing c_i sequences that are $O(i^{\varepsilon})$, for $\varepsilon < 1 - 1/p$, if $p > 1$. Although increasing c_i sequences can be important, one might wish to trade some of the room that is allowed for the c_i sequence against less strict assumptions on the $\psi(m)$ sequence. Hansen (1991, 1992) assumed

$$\sum_{m=1}^{\infty} \psi(m) < \infty.$$

McLeish (1975) required the $\psi(m)$ sequence to be of *size* $-\frac{1}{2}$. A sequence $\psi(m)$ is said to be of size $-\beta$, $\beta > 0$, if $\psi(m) = O(m^{-\beta-\varepsilon})$ for some $\varepsilon > 0$. McLeish (1975) only considered L_2-mixingales. If we assume $\sup_{i\geq 1} E|X_i|^p < \infty$ and restrict attention to the case of indices of magnitude that are bounded or increasing very slowly, we can improve upon previous results by means of the following theorem and the corollary to it.

THEOREM 2. *Suppose the sequence* $\{X_i, F_i\}$ *is an* L_p*-mixingale,* $p > 1$, *such that* $\sup_{i\geq 1} E|X_i|^p < \infty$. *Suppose we can find strictly positive nondecreasing sequences* B_i *and* m_i, m_i *integer-valued and* $m(1) = 1$, *such that the following conditions hold*:

(A) $\sum_{i=1}^{\infty} i^{-1} B_i^{1-p} < \infty$,
(B) $\sum_{i=1}^{\infty} c_i i^{-1} \psi(m_i) < \infty$, *and*
(C) *for all* $\delta > 0$, $\sum_{n=1}^{\infty} m_n \exp(-n\delta^2 m_n^{-2} B_n^{-2}) < \infty$.

Then,

$$n^{-1} \sum_{i=1}^{n} X_i \to 0$$

almost surely.

The following corollary is now easily established.

COROLLARY 1. *Suppose the sequence* $\{X_i, F_i\}$ *is an* L_p*-mixingale,* $p > 1$, *such that for the* L_p*-indices of magnitude we have* $\sup_{i\geq 1}|c_i| < \infty$. *Suppose* $\psi(m)$ *is of order* $O(\log^{-1-\eta}(m))$ *as* $m \to \infty$, *for some* $\eta > 0$. *Then,*

$$n^{-1}\sum_{i=1}^{n} X_i \to 0$$

almost surely.

Proof. See Section 5.

Corollary 1 shows that, although restrictions on the rate of convergence of $\psi(m)$ are necessary, sample averages of L_p-mixingales can converge almost surely if the magnitude indices c_i are uniformly bounded and if some size of the $\psi(m)$ sequence is in evidence.

4. WEAK AND STRONG LAWS FOR NEAR-EPOCH DEPENDENT PROCESSES

In the theory of consistency and asymptotic normality of (non)parametric estimators for dependent samples, use is made of the near epoch dependence concept. The introduction of the concept of near epoch dependence is motivated by problems that occur when merely mixing sequences are considered (for the definitions of strong (α-) and uniform (ϕ-) mixing sequences, see, e.g., Gallant and White, 1988, p. 23; Pötscher and Prucha, 1991, p. 164). As is well known, functions of a finite number of lagged variables of a mixing process are again mixing, but this is not necessarily the case if a function of the entire history of a mixing process is considered. Even simple AR(1) processes can fail to be either ϕ- or α-mixing. It may, however, be possible to verify the L_p-near epoch dependence condition.

In what follows, $\{V_i : i \geq 1\}$ denotes a sequence of random variables on (Ω, F, P). Andrews (1988) defined the L_p-near epoch dependence concept as follows.

DEFINITION 2. *The sequence $\{X_i\}$ is called L_p-near epoch dependent, $p \geq 1$, on $\{V_i\}$ if for $m \geq 0$*

$$\| X_i - E(X_i \,|\, V_{i-m}, \ldots, V_{i+m}) \|_p \leq d_i \nu(m)$$

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

Typically, we will have to assume that the basis process $\{V_i\}$ is either ϕ- or α-mixing in order to obtain results. Note that mixing processes can be considered L_p-near epoch dependent for all $p \geq 1$ on themselves with $\nu(m) = 0$ for $m \geq 0$ and $d_i = 1$ for $i \geq 1$. Gallant and White (1988) considered L_2-near epoch dependent sequences with uniformly bounded d_i and referred to such sequences as "near epoch dependent." Pötscher and Prucha (1991) introduced the concept of L_p-approximability. This concept unifies Gallant and White's near epoch dependence concept and the stochastic stability and ν-stability concepts of Bierens (1981, 1984). The L_p-approximability concept is very useful in establishing weak LLN's and uniform weak LLN's for functions of L_p-approximable processes. However, for obtaining almost sure results, the L_p-near epoch dependence concept is more appropriate (see Pötscher and Prucha, 1991). Moreover, the L_p-approximability concept is not convenient for considering trended random variables.

Inequalities of Serfling (1968) for the ϕ-mixing case and McLeish (1975) for the α-mixing case can now be used to show that random variables that are L_p-near epoch dependent on some mixing sequence are L_p-mixingales. The following theorem provides a weak law of large numbers for L_p-near epoch dependent sequences that illustrates the possible use of Theorem 1.

THEOREM 3. *Suppose the sequence* $\{X_i\}$ *is* L_p*-near epoch dependent,* $1 \leq p \leq 2$, *on an* α*-mixing sequence* $\{V_i\}$ *and assume that for all* i, $EX_i = 0$. *If for some sequence* $C_n \geq 1$, $C_n = o(n^{1/2})$, *and some* $r > p$

(A) $\limsup_{n\to\infty} C_n^{1-r/p} n^{-1} \sum_{i=1}^n \|X_i\|_r^{r/p} < \infty$ *and*
(B) *for all* $K > 0$,

$$\lim_{n\to\infty} n^{-1} \sum_{i=1}^{n} (d_i + \|X_i\|_r)(\nu([Kn^{1/2}C_n^{-1}/2]) + \alpha([Kn^{1/2}C_n^{-1}/2])^{1/p-1/r}) = 0,$$

we have

$$\left\| n^{-1} \sum_{i=1}^{n} X_i \right\|_p \to 0$$

as $n \to \infty$.

Proof. See Section 5.

An application of Theorem 2 establishes the following strong LLN for L_p-near epoch dependent processes.

THEOREM 4. *Suppose the sequence* $\{X_i\}$ *is* L_p*-near epoch dependent,* $p > 1$, *on* $\{V_i\}$ *with uniformly bounded* d_i *sequence, where* $\{V_i\}$ *is* α*-mixing with coefficients* $\alpha(m)$ *and assume that for all* i, $EX_i = 0$. *Suppose* $\sup_{i\geq 1} E|X_i|^r < \infty$ *for some* $r > p$. *Suppose the mixing numbers* $\alpha(m)$ *and the near epoch dependence numbers* $\nu(m)$ *satisfy*

$$\nu([m/2]) + 6\alpha([m/2])^{1/p-1/r} = O(\log^{-1-\eta}(m))$$

as $m \to \infty$ *for some* $\eta > 0$. *Then,*

$$n^{-1} \sum_{i=1}^{n} X_i \to 0$$

almost surely.

Proof. See Section 5.

A potentially useful application of the strong law presented here is to stochastically Lipschitz-continuous functions of L_p-near epoch dependent processes (for a definition of stochastically Lipschitz-continuous functions, see Gallant and White, 1988; Pötscher and Prucha, 1991). Theorems of Gallant and White (1988) and Pötscher and Prucha (1991) show under regularity conditions that stochastically Lipschitz-continuous functions of processes that are L_p-near epoch dependent of some size on an α-mixing sequence are again L_p-near epoch dependent on the same sequence of some size. The latter size, however, need not be the same as the former one and can depend on quantities involving the existence of certain moments and on the size of the original sequence. Because Theorem 4 applies to processes that are L_p-near epoch dependent on some α-mixing sequence for which both the α-mixing sequence and the ν-sequence decay polynomially, Theorem 4 will apply to such transformations of a near epoch dependent process of some size, regardless of the exact size of the sequences involved. Note, however, that the transformation result of Andrews (1991), which involves Lipschitz-continuous functions of an L_p-near epoch dependent process with a constant Lipschitz function, leaves near-epoch dependence numbers intact. Moreover, note that for obtaining weak LLN's for transformations of near epoch dependent processes no conditions on the $\psi(m)$ sequence are needed except that it has zero limit as $m \to \infty$ (see Pötscher and Prucha, 1991).

5. PROOFS

This section contains the proofs of the theorems and the lemma. We will start with a proof of Azuma's (1967) inequality for martingale difference sequences. This inequality is used in the proof of Theorem 2. The proof is nearly identical to the proof of Hoeffding's inequality given in Pollard (1984, Appendix B).

LEMMA 1 (Azuma). *If $\{X_i, F_i\}$ is a zero mean martingale difference sequence and*

$$|X_i| \le B \quad a.s.,$$

then for all $\varepsilon > 0$

$$P\left[\left|\sum_{i=1}^{n} X_i\right| > \varepsilon\right] \le 2\exp(-\varepsilon^2/(2nB^2)).$$

Proof. Consider $E(\exp(tX_i) \mid F_{i-1})$. By convexity,

$$\exp(tX_i) \le \exp(-tB)(B - X_i)/(2B) + \exp(tB)(X_i + B)/(2B),$$

so

$$E(\exp(tX_i) \mid F_{i-1}) \le \exp(-tB)/2 + \exp(tB)/2$$

because

$E(X_i \mid F_{i-1}) = 0.$

Analogously to Pollard (1984, Appendix B), it can now be shown that

$\log(E(\exp(tX_i) \mid F_{i-1})) \le (\frac{1}{2})t^2B^2.$

We will use the successive conditioning strategy that is employed also in proofs of central limit theorems for martingale differences (e.g., Pollard, 1984, Ch. 8). Then it easily follows that

$$E\exp\left(t\sum_{i=1}^{n} X_i\right) \le \exp(nt^2B^2/2).$$

Because, by the Markov inequality, for all $t \ge 0$,

$$P\left[\sum_{i=1}^{n} X_i \ge \varepsilon\right] \le \exp(-\varepsilon t)E\exp\left(t\sum_{i=1}^{n} X_i\right) \le \exp(-\varepsilon t + nt^2B^2/2)$$

the result now follows by setting $t = \varepsilon/nB^2$ and applying the same result to $\{-X_i\}$. ■

Proof of Theorem 1. The line of proof is essentially that of Andrews (1988), but we allow for $m = m(n)$. Note that, for all $B > 0$ and all integer-valued $m \ge 1$,

$$\begin{aligned} n^{-1}\sum_{i=1}^{n} X_i &= n^{-1}\sum_{i=1}^{n} X_i - E(X_i \mid F_{i+m-1}) \\ &\quad + n^{-1}\sum_{i=1}^{n} E(X_i I(|X_i| \le BC_n) \mid F_{i+m-1}) \\ &\qquad\qquad - E(X_i I(|X_i| \le BC_n) \mid F_{i-m}) \\ &\quad + n^{-1}\sum_{i=1}^{n} E(X_i I(|X_i| > BC_n) \mid F_{i+m-1}) \\ &\qquad\qquad - E(X_i I(|X_i| > BC_n) \mid F_{i-m}) \\ &\quad + n^{-1}\sum_{i=1}^{n} E(X_i \mid F_{i-m}) \\ &= T_1 + T_2 + T_3 + T_4. \end{aligned} \tag{5.1}$$

We will take $m = m(n) = m_n$. By Assumption A, using the conditional Jensen inequality, we can pick B so large that

$$\limsup_{n\to\infty} \|T_3\|_p \le \limsup_{n\to\infty} 2n^{-1}\sum_{i=1}^{n} \|X_i I(|X_i| > BC_n)\|_p < \varepsilon.$$

Define

$$Y_{ji} = E\big(X_i I(|X_i| \le BC_n)\,|\,F_{i+j}\big) - E\big(X_i I(|X_i| \le BC_n)\,|\,F_{i+j-1}\big),$$

and note that for all j and $1 \le i \le n$ $\{Y_{ji}, F_{i+j}\}$ is a bounded martingale difference sequence. We now have, similarly to Andrews (1988),

$$\begin{aligned}\|T_2\|_2 &= \left\| n^{-1} \sum_{i=1}^{n} \sum_{j=-m_n+1}^{m_n-1} Y_{ji} \right\|_2 \\ &\le \sum_{j=-m_n+1}^{m_n-1} \left\| n^{-1} \sum_{i=1}^{n} Y_{ji} \right\|_2 \\ &\le 4 m_n n^{-1/2} BC_n.\end{aligned}$$

So for the choice $m_n = \max(1, [C_n^{-1} n^{1/2} B^{-1} \varepsilon/4])$, we have by Lyapounov's inequality

$$\begin{aligned}\limsup_{n\to\infty} \|T_2\|_p \le \limsup_{n\to\infty} \|T_2\|_2 \\ \le \limsup_{n\to\infty} 4[C_n^{-1} n^{1/2} B^{-1} \varepsilon/4]\, n^{-1/2} BC_n \le \varepsilon\end{aligned}$$

if $p \le 2$ because $[x] \le x$ and $C_n = o(n^{1/2})$. Clearly by the mixingale definition for this choice of m_n

$$\begin{aligned}&\limsup_{n\to\infty} \|T_1 + T_4\|_p \\ &\quad\le \limsup_{n\to\infty} n^{-1} \sum_{i=1}^{n} \big(\|X_i - E(X_i\,|\,F_{i+m_n-1})\|_p + \|E(X_i\,|\,F_{i-m_n})\|_p \big) \\ &\quad\le 2 \limsup_{n\to\infty} n^{-1} \sum_{i=1}^{n} c_i \psi([C_n^{-1} n^{1/2} B^{-1} \varepsilon/4]) = 0\end{aligned}$$

by Assumption B. Because ε was arbitrary, the result follows. ■

Proof of Theorem 2. Once again, for all $B > 0$ and integer-valued $m \ge 1$, consider equation (5.1). For proving a strong LLN we will make both B and m depend on i, and we set $C_n = 1$ for all $n \ge 1$. To obtain the result, we prove the following three steps:

STEP 1. *If*

$$\sum_{i=1}^{\infty} i^{-1} B_i^{1-p} < \infty$$

and

$$\sup_{i\ge 1} E|X_i|^p < \infty,$$

then $T_3 \to 0$ *almost surely.*

Proof. Let $S_n = \sum_{i=1}^n \big(E\big(X_i I(|X_i| > B_i)\,|F_{i+m_i-1}\big) - E\big(X_i I(|X_i| > B_i)\,|F_{i-m_i}\big)\big)i^{-1}$. Then, for any $\delta > 0$, using the Markov inequality and monotone convergence, and the Hölder and Markov inequalities,

$$P\big[\sup_{j\geq 1}|S_{n+j} - S_n| > \delta\big]$$

$$\leq 2\delta^{-1}\sum_{i=n+1}^{\infty} E|X_i|I(|X_i| > B_i)i^{-1}$$

$$\leq 2\sum_{i=n+1}^{\infty} i^{-1}\delta^{-1}E|X_i|^p B_i^{1-p} \to 0 \quad \text{as } n \to \infty$$

if $\sum_{i=1}^{\infty} i^{-1}B_i^{1-p} < \infty$ and $\sup_{i\geq 1}E|X_i|^p < \infty$, which is imposed. So we conclude by the Cauchy criterion that S_n converges almost surely, so by the Kronecker Lemma, $T_3 \to 0$ almost surely. ■

STEP 2. *If $\{X_i, F_i\}$ is an L_p-mixingale sequence such that*

$$\sum_{i=1}^{\infty} c_i i^{-1}\psi(m_i) < \infty,$$

then $T_1 + T_4 \to 0$ almost surely.

Proof. Let $S_n' = \sum_{i=1}^n E(X_i|F_{i-m_i})i^{-1}$. Then, for any $\delta > 0$, using the Markov inequality and monotone convergence, the Lyapounov inequality, and the L_p-mixingale definition,

$$\begin{aligned} P\big[\sup_{j\geq 1}|S_{n+j}' - S_n'| > \delta\big] &\leq P\left[\sum_{i=n+1}^{\infty} |E(X_i|F_{i-m_i})|i^{-1} > \delta\right] \\ &\leq \delta^{-1}\sum_{i=n+1}^{\infty} E|E(X_i|F_{i-m_i})|i^{-1} \\ &\leq \delta^{-1}\sum_{i=n+1}^{\infty} \|E(X_i|F_{i-m_i})\|_p i^{-1} \\ &\leq \delta^{-1}\sum_{i=n+1}^{\infty} c_i\psi(m_i)i^{-1} \to 0 \end{aligned}$$

as $n \to \infty$ if $\sum_{i=1}^{\infty} c_i i^{-1}\psi(m_i) < \infty$, which is imposed. So S_n' converges almost surely to some random variable, by the Cauchy criterion. So $T_4 \to 0$ almost surely by the Kronecker Lemma. The same argument holds for T_1. ■

STEP 3. *Let m_i and B_i be strictly positive nondecreasing sequences such that $m(1) = 1$ and m_i is integer-valued and assume that for all $\delta > 0$*

$$\sum_{n=1}^{\infty} m_n \exp(-n\delta^2 m_n^{-2} B_n^{-2}) < \infty.$$

Then, $T_2 \to 0$ *almost surely.*

Proof. In this proof, we use Azuma's Lemma, together with the Borel-Cantelli Lemma. For $j \geq 1$ define

$$v(j) = \inf\{l \in \{1,2,\dots\} : m(l) \geq j\},$$

and for $j \geq 1$ define $q(j) = v(j)$; for $j \leq 0$ define $q(j) = v(1-j)$. Note that

$$P\left[\left|n^{-1}\sum_{i=1}^{n} E(X_i I(|X_i| \leq B_i) \,|\, F_{i+m_i-1}) - E(X_i I(|X_i| \leq B_i) \,|\, F_{i-m_i})\right| > \delta\right]$$

$$= P\left[\left|n^{-1}\sum_{i=1}^{n}\sum_{j=-m_i+1}^{m_i-1} E(X_i I(|X_i| \leq B_i) \,|\, F_{i+j}) - E(X_i I(|X_i| \leq B_i) \,|\, F_{i+j-1})\right| > \delta\right]$$

$$= P\left[\left|n^{-1}\sum_{j=-m_n+1}^{m_n-1}\sum_{i=q(j)}^{n} E(X_i I(|X_i| \leq B_i) \,|\, F_{i+j}) - E(X_i I(|X_i| \leq B_i) \,|\, F_{i+j-1})\right| > \delta\right]$$

$$\leq \sum_{j=-m_n+1}^{m_n-1} P\left[\left|n^{-1}\sum_{i=q_j}^{n} E(X_i I(|X_i| \leq B_i) \,|\, F_{i+j}) - E(X_i I(|X_i| \leq B_i) \,|\, F_{i+j-1})\right| > \delta/2m_n\right]$$

$$\leq \sum_{j=-m_n+1}^{m_n-1} 2\exp(-\delta^2 n^2 B_n^{-2} m_n^{-2}(n - q(j) + 1)/32)$$

$$\leq 4m_n \exp(-n\delta^2 m_n^{-2} B_n^{-2}/32).$$

This implies that, by virtue of the Borel-Cantelli Lemma, $T_2 \to 0$ almost surely. Corollary 1 now follows by taking $m_i = [i^{\varepsilon}]$, for some $0 < \varepsilon < \frac{1}{2}$, and setting $B_i^{1-p} = (\log(i+1))^{-2}$. Note that the condition $\sup_{i\geq 1} E|X_i|^p < \infty$ follows from the requirement $\sup_{i\geq 1}|c_i| < \infty$ on the L_p-mixingale magnitude indices in view of the inequality

$$\|X_i\|_p \leq (\psi(0) + \psi(1))c_i$$

by the definition of an L_p-mixingale. ∎

Finally, we will prove our results regarding near epoch dependent sequences.

Proof of Theorem 3. Note that (A) of Theorem 1 holds because, by the Hölder and Markov inequalities,

$$\lim_{B\to\infty} \limsup_{n\to\infty} n^{-1} \sum_{i=1}^{n} \| X_i I(|X_i| > BC_n) \|_p$$

$$\leq \lim_{B\to\infty} \limsup_{n\to\infty} B^{1-r/p} C_n^{1-r/p} n^{-1} \sum_{i=1}^{n} \| X_i \|_r^{r/p} = 0$$

in view of Assumption A of Theorem 3. Assumption B of Theorem 1 is verified by noting that X_i is an L_p-mixingale with mixingale coefficients $c_i = 2d_i + \|X_i\|_r$ and mixingale numbers $\psi(m) = \nu([m/2]) + 6\alpha([m/2])^{1/p-1/r}$ (see Andrews, 1988, equation 2) and by Assumption B of Theorem 3. ■

Proof of Theorem 4. Again consider Andrews' (1988) equation 2 as in the proof of Theorem 3. Noting that X_i is a mixingale and that the associated c_i sequence is uniformly bounded by construction and that $\psi(m) = O((\log(m))^{-1-\eta})$ for some $\eta > 0$ by assumption the theorem follows by an application of Corollary 1 to Theorem 2. ■

REFERENCES

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

Andrews, D.W.K. (1991) An empirical process central limit theorem for dependent nonidentically distributed random variables. *Journal of Multivariate Analysis* 38, 187–203.

Azuma, K. (1967) Weighted sums of certain dependent random variables. *Tohoku Mathematical Journal* 19, 357–367.

Bierens, H.J. (1981) *Robust Methods and Asymptotic Theory.* Lecture notes in economics and mathematical systems 192. Berlin: Springer-Verlag.

Bierens, H.J. (1984) Model specification testing of time series regressions. *Journal of Econometrics* 24, 323–353.

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

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

Gut, A. (1992) The weak law of large numbers for arrays. *Statistics and Probability Letters* 14, 49–52.

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. *Annals of Probability* 3, 829–839.

Pollard, D. (1984) *Convergence of Stochastic Processes.* New York: Springer-Verlag.

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

Serfling, R.J. (1968) Contributions to central limit theory for dependent variables. *Annals of Mathematical Statistics* 39, 1158–1175.