ANNALES D'ÉCONOMIE ET DE STATISTIQUE. – N° 51 — 1998

Weak Laws of Large Numbers for Dependent Random Variables

Robert M. DE JONG*

ABSTRACT. – In this paper we will prove several weak laws of large numbers for dependent random variables. The weak dependence concept that is used is the mixingale concept. From the weak laws of large numbers for triangular arrays of mixingale random variables, weak laws for mixing and near epoch dependent random variables follow. Features of the weak laws of large numbers that are proven here is that they impose tradeoff conditions between dependence and trending of the summands.

Lois faibles de grands nombres pour des variables aléatoires dépendantes

RÉSUMÉ. – Dans cet article nous démontrerons plusieurs lois faibles des grands nombres pour des variables aléatoires dépendantes. Le concept de la dépendance faible qui est utilisé est le concept de mixingale. Des lois faibles de grands nombres pour des variables aléatoires de mixingale, pour des variables aléatoires mélangeantes et les variables aléatoires dépendantes d'époque proche suivent. Les caractéristiques des lois faibles des grands nombres qui sont prouvées ici correspondent à imposer des conditions d'arbitrage entre dépendance dans les sommations.

* R.M. de JONG: Department of Economics, Michigan State University, East Lansing, MI 48824, USA. The author thanks two anonymous referees for their perceptive comments and remarks.

1 Introduction

It has long been recognized that for random variables that are time series, it is problematic to make the assumption of independence. Therefore, in order to justify asymptotic inference for time series, laws of large numbers (LLNs) and central limit theorems for dependent random variables are important. Traditionally, so-called *mixing* assumptions are used in statistical literature. Mixing conditions can be used for many applications and are strong enough to provide the foundations for complex proofs. However, mixing assumptions may be restrictive and unnecessary and rule out potentially interesting types of random variables. For proofs of LLNs, usually the *mixingale* assumption suffices. Therefore, weak LLNs for mixingales are important for obtaining weak convergence and asymptotic normality results for estimators in the case of dependent random variables. GALLANT and WHITE [1988] and PÖTSCHER and PRUCHA [1991a, 1991b] establish such results. A mixingale sequence can be viewed as an asymptotic equivalent of a martingale difference sequence. Strong LLNs for mixingale sequences have been established in MCLEISH [1975], HANSEN [1991, 1992], DE JONG [1995] and DAVIDSON and DE JONG [1995]. Weak LLNs for mixingale sequences can be found in ANDREWS [1988], DAVIDSON [1993a] and De JONG [1995]. In this paper, we establish four weak LLNs for mixingales.
While the mixingale assumption is a convenient and minimal assumption for proving LLNs, the mixingale assumption usually is not suitable for serving as a primitive assumption. This is because the mixingale property does not necessarily carry over to functions of mixingales. Also, the mixingale assumption is not strong enough for deriving central limit theorems with easily verifiable conditions. Therefore, in nonlinear estimation procedures such as maximum likelihood estimation, the mixingale assumption cannot easily be used. To overcome these problems, the concept of *near epoch dependent* (NED) random variables has been introduced. The concept of near epoch dependence allows for a relatively large amount of dependence. Consider for example the $MA(\infty)$ process

(1) $$X_t = \sum_{j=0}^{\infty} \rho_j V_{t-j}$$

where V_t is a sequence of random variables, and ρ_j is a sequence of deterministic constants. Clearly, ARMA processes are a special case of such a process. It is well-known (See ANDREWS [1985]) that even for the special case of i.i.d. V_t and an AR(1) process, X_t need not be mixing. However, in view of the fact that it is widely recognized that $MA(\infty)$ processes can provide good approximations to economic processes, it is desirable that results such as weak LLNs are available for such processes. Usually, for proofs of LLNs and for central limit theorems, it is not even necessary that V_t is assumed to be independent; a mixing assumption on V_t will usually be sufficient. Moreover, it is desirable that such results can also be established for (classes of) functions $g(.)$ of processes such as X_t as defined above. Typically, the strategy for proving a weak LLN for statistics such as

$$n^{-1}\sum_{t=1}^{n} g(Y_t) \tag{2}$$

for some function $g(.)$ is as follows. First, we make the NED assumption for Y_t, thereby allowing for Y_t that are possibly MA(∞) processes. Second, it is shown that $g(Y_t)$ is NED. Third, it is proven that $g(Y_t)$ is a mixingale by using the NED property of $g(Y_t)$. Finally, a LLN for mixingales is applied.

The new results in this paper are the following. Firstly, we establish an L_p-type convergence result for triangular mixingale arrays, for $1 \leq p \leq 2$, that contains the results of ANDREWS [1988], DAVIDSON [1993a], and DE JONG [1995] as special cases. Secondly, we establish an L_p-type weak LLN for triangular mixingale arrays that avoids a condition that allows truncation of the summands, such as uniform integrability. Thirdly, an L_2-type weak LLN for mixingales is established by bounding the sum of the covariances. Finally, another L_2-type weak LLN is derived; neither one of these two L_2-type LLNs that are derived dominates the other.

The plan of this paper is as follows. Section 2 gives the various definitions that are needed and provides an overview of weak LLNs for mixingales that have been reported in the literature. The new weak LLNs that are established are stated in section 3. In section 4, our result are applied to obtain weak LLNs for NED random variables. The Mathematical Appendix contains the proofs of the results.

2 Definitions and Weak LLNs

A triangular mixingale array is defined as follows. Let $\mathcal{F}_t^n$ denote an infinite array of σ-fields that is increasing in t for each n, and let X_{nt} and V_{nt} for $t = 1, 2, \ldots, n$ denote a triangular array of random variables defined on the probability space $(\Omega, \mathcal{F}, P)$. Let $\| X \|_p$ denote $(E|X|^p)^{1/p}$ for $p \geq 1$.

DEFINITION 1 : $\{X_{nt}, \mathcal{F}_t^n\}$ is called an L_p-mixingale if for $\psi(m) \geq 0$ and $c_{nt} \geq 0$ and for all $m \geq 0$ and $t \geq 1$,

$$\| X_{nt} - E(X_{nt}|\mathcal{F}_{t+m}^n) \|_p \leq c_{nt}\psi(m+1) \tag{3}$$

and

$$\| E(X_{nt}|\mathcal{F}_{t-m}^n) \|_p \leq c_{nt}\psi(m) \tag{4}$$

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

The mixingale concept is a weak dependence concept that is strong enough for establishing proofs of weak and strong LLNs. However, it is desirable that transformations of weakly dependent random variables are again weakly dependent in some sense. The mixingale concept does not necessarily satisfy this requirement. Therefore, the dependence concept of near epoch dependence has been introduced in the literature. See GALLANT and WHITE [1988] and PÖTSCHER and PRUCHA [1991a] for detailed discussions of this and related dependence concepts. Also note that a similar dependence concept is already used by BILLINGSLEY [1968], who uses the term "functions of mixing processes".

DEFINITION 2 : The triangular array of random variables X_{nt} is called L_p-NED, $p \geq 1$, on V_{nt} if for $\nu(m) \geq 0$ and $d_{nt} \geq 0$ and for all $m \geq 0$ and $t \geq 1$,

$$\| X_{nt} - E(X_{nt}|V_{n,t-m}, \ldots, V_{n,t+m}) \|_p \leq d_{nt}\nu(m) \tag{5}$$

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

The basis process V_{nt} typically needs to satisfy a weak dependence concept such as some mixing condition in order to obtain useful results. For the definition of strong (α-) and uniform (ϕ-) mixing random variables see e.g. GALLANT and WHITE [1988, p. 23] and PÖTSCHER and PRUCHA [1991a, p. 164]. If we consider the process

$$X_t = \sum_{j=0}^{\infty} \rho_j V_{t-j}, \tag{6}$$

where we assume that $\| V_t \|_p < \infty$ for all t, it can be easily shown that it is L_p-NED, for $p > 1$, on V_t. This is because

$$\begin{aligned} &\| X_t - E(X_t|V_{t-m}, \ldots, V_{t+m}) \|_p \\ &\leq \| \sum_{j=m+1}^{\infty} \rho_j V_{t-j} \|_p \\ &\leq \left(\sum_{j=m+1}^{\infty} |\rho_j| \right) (\sup_{s \leq t} \| V_s \|_p) \end{aligned} \tag{7}$$

and therefore we conclude that X_t is L_p-NED on V_t with respect to constants $d_t = \sup_{s \leq t} \| V_s \|_p$ and $\nu(m) = \sum_{j=m+1}^{\infty} |\rho_j|$, provided that $\sum_{j=0}^{\infty} |\rho_j| < \infty$. Therefore, weaker conditions for weak LLNs for NED random variables translate immediately to a wider class of (functions of) MA(∞) random variables for which a weak LLNs holds. The example illustrates the advantages of assuming near epoch dependence over mixing assumptions: under general regularity conditions, ARMA processes can be shown to be NED, while showing mixing conditions can be hard. Also, transformations $g(X_t)$ can be shown to be NED as well; if for example, $g(.)$ is differentiable with uniformly bounded derivative, $g(X_t)$ will be NED with the same $\nu(m)$ sequence that X_t had, while the d_t has been multiplied by a constant (or vice versa). More general results on transformations of

NED processes (see e.g.GALLANT and WHITE [1988] and PÖTSCHER and PRUCHA [1991a]) indicate that transformations of X_t that satisfy smoothness conditions will be NED under regularity conditions, but with sequences of the type d_t^β and $\nu(m)^\alpha$, for constants $\alpha, \beta > 0$ that depend on $g(.)$. Note that NED random variables are mixingales because

$$(8) \qquad \begin{aligned} &\| E(X_{nt}|V_{n,t-m}, V_{n,t-m-1}, \ldots) \|_p \\ &\le d_{nt}\nu([m/2]) + 2\phi([m/2])^{1-1/r} \| X_{nt} \|_r \end{aligned}$$

and

$$(9) \qquad \begin{aligned} &\| E(X_{nt}|V_{n,t-m}, V_{n,t-m-1}, \ldots) \|_p \\ &\le d_{nt}\nu([m/2]) + 6\alpha([m/2])^{1/p-1/r} \| X_{nt} \|_r \end{aligned}$$

where $p \le r$ and $\alpha(m)$ and $\phi(m)$ denote the mixing sequences of V_{nt}. See ANDREWS [1988] for proofs of these results. In ANDREWS [1988], the following theorem can be found as well.

THEOREM 1 : (ANDREWS [1988]) Suppose $\{X_{nt}, \mathcal{F}_t^n\}$ is an L_p-mixingale, for $p \ge 1$, such that
1. $\lim_{m\to\infty} \psi(m) = 0$;
2. X_{nt} is uniformly integrable;
3. $\limsup_{n\to\infty} n^{-1}\sum_{t=1}^n c_{nt} < \infty$.
Then

$$(10) \qquad \lim_{n\to\infty} \| n^{-1}\sum_{t=1}^n X_{nt} \|_1 = 0.$$

PÖTSCHER and PRUCHA [1991] noted that Andrews' weak LLN as stated above can be improved upon a little, but not so as to allow c_{nt} sequences that are trending to infinity with n or t. For L_2-mixingales, a slightly different result is available.

THEOREM 2 : (ANDREWS [1988]) Suppose $\{X_{nt}, \mathcal{F}_t^n\}$ satisfies the conditions for an L_2-mixingale, but instead of the condition that $\psi(m) \to 0$ as $m \to \infty$, we now impose that $m^{-1}\sum_{j=0}^m \psi(j) \to 0$ as $m \to \infty$. In addition, assume that $\sup_{t,n\ge 1} c_{nt} < \infty$. Then

$$(11) \qquad \lim_{n\to\infty} \| n^{-1}\sum_{t=1}^n X_{nt} \|_2 = 0.$$

In a recent paper, DAVIDSON [1993a] establishes a result that is useful for establishing a central limit theorem for sums of asymptotically degenerate NED random variables (that is, for random variables for which the variance approaches zero as the observation number tends to infinity). Note that the latter result can be found in DAVIDSON [1993b]; note that DE JONG

[1997] establishes a central limit theorem that is even more general. In our notation, Davidson's weak LLN takes the following form. In the sequel of this paper, a_n denotes a strictly positive nondecreasing sequence such that $a_n \to \infty$ as $n \to \infty$.

THEOREM 3 : (DAVIDSON [1993a]) Suppose $\{X_{nt}, \mathcal{F}_t^n\}$ is an L_1-mixingale such that
1. $\lim_{m\to\infty} \psi(m) = 0$;
2. $\limsup_{n\to\infty} a_n^{-1} \sum_{t=1}^n c_{nt} < \infty$; 3. $\limsup_{n\to\infty} a_n^{-2} \sum_{t=1}^n c_{nt}^2 = 0$;
4. $|X_{nt} c_{nt}^{-1}|$ is uniformly integrable.
Then

$$\lim_{n\to\infty} \| a_n^{-1} \sum_{t=1}^n X_{nt} \|_1 = 0. \tag{12}$$

In DE JONG [1995], the first weak LLN is stated that is capable of providing a result for c_{nt} sequences that are trending to infinity as n or t tends to infinity. This typically occurs if the random variables X_{nt} are trended with n or t.

THEOREM 4 : (DE JONG [1995]) Suppose $\{X_{nt}, \mathcal{F}_t^n\}$ is an $L_p$mixingale, $1 \le p \le 2$, such that for some sequence $B_n \ge 1$, $B_n = o(n^{1/2})$,
1. $\lim_{K\to\infty} \limsup_{n\to\infty} n^{-1} \sum_{t=1}^n \| X_{nt} I(|X_{nt}| > KB_n) \|_p = 0$;
2. For all $K > 0$, $\lim_{n\to\infty} n^{-1} \sum_{t=1}^n c_{nt} \psi([K n^{1/2} B_n^{-1}]) = 0$.
Then

$$\lim_{n\to\infty} \| n^{-1} \sum_{t=1}^n X_{nt} \|_p = 0. \tag{13}$$

3 Four Weak LLNs

In this section, similarly to DAVIDSON and DE JONG [1995], we will characterize a weak LLN as a theorem citing conditions such that $a_n^{-1} \sum_{t=1}^n X_{nt} \to 0$ in probability or in L_p. Our first result is a result that contains the weak LLNs for triangular L_p-mixingale arrays, $1 \le p \le 2$, of ANDREWS [1988], DAVIDSON [1993a] and DE JONG [1995] as special cases. Note that the LLNs that are discussed in this section all allow trending of the random variables involved. This is contrary to the results that can be found in the literature, apart from Theorem 4. Note that the theorem below essentially is Theorem 4, where B_n is allowed to depend on t also, and instead of averages, $a_n^{-1} \sum_{t=1}^n X_{nt}$ is considered.

THEOREM 5 : Suppose $\{X_{nt}, \mathcal{F}_t^n\}$ is an L_p-mixingale, where $1 \leq p \leq 2$. Assume there exists a triangular array $B_{nt} \geq 1$ such that

1. $\lim_{K\to\infty} \limsup_{n\to\infty} a_n^{-1} \sum_{t=1}^n \| X_{nt} I(|X_{nt}| > B_{nt}K) \|_p = 0$;
2. $\lim_{n\to\infty} a_n^{-2} \sum_{t=1}^n B_{nt}^2 = 0$;
3. For all $K > 0$, $\lim_{n\to\infty} a_n^{-1} \sum_{t=1}^n c_{nt} \psi([K(a_n^{-2} \sum_{t=1}^n B_{nt}^2)^{-1/2}]) = 0$.

Then

$$(14) \qquad \lim_{n\to\infty} \| a_n^{-1} \sum_{t=1}^n X_{nt} \|_p = 0.$$

Proof: See appendix.

Note that this theorem and the theorems that follow all impose a tradeoff between dependence as measured by the $\psi(m)$ sequence and heterogeneity as measured by the c_{nt}. If we assume that X_{nt} is uniformly integrable, choose $B_{nt} = 1$ and $a_n = n$, and if it is assumed that

$$(15) \qquad \limsup_{n\to\infty} n^{-1} \sum_{t=1}^n c_{nt} < \infty,$$

we obtain Andrews' weak LLN as a special case. The conditions of Davidson's weak LLN are obtained if we choose $B_{nt} = c_{nt}$ and $p = 1$ since

$$(16) \lim_{K\to\infty} \limsup_{n\to\infty} a_n^{-1} \sum_{t=1}^n \| X_{nt} I(|X_{nt}| > c_{nt}K) \|_1$$
$$\leq \left[\limsup_{n\to\infty} a_n^{-1} \sum_{t=1}^n c_{nt}\right] [\lim_{K\to\infty} \sup_{n,t\geq 1} \| X_{nt} c_{nt}^{-1} I(|X_{nt}| > c_{nt}K) \|_1] = 0$$

by uniform integrability of $|X_{nt} c_{nt}^{-1}|$. Note that Assumption (5.3) is implied by Assumptions (3.1) and (3.3) of Davidson's weak LLN. Therefore, Davidson's weak LLN is a special case of Theorem 5 too.

If we want to avoid a condition of the type of condition (5.1) that allows the truncation of the summands, the following alternative new L_p-type weak LLN can be used:

THEOREM 6 : Suppose $\{X_{nt}, \mathcal{F}_t^n\}$ is an L_p-mixingale for $1 \leq p \leq 2$ such that

$$(17) \qquad \lim_{n\to\infty} \left(a_n^{-p} \sum_{t=1}^n c_{nt}^p\right)^{1/p} \sum_{m=0}^{[n^{(p-1)/p}]+1} \psi(m) = 0.$$

Then

$$(18) \qquad \lim_{n\to\infty} \| a_n^{-1} \sum_{t=1}^n X_{nt} \|_p = 0.$$

Proof: See Appendix.

By extending the argument used for the proof of Theorem 2, it is possible to derive the following weak LLN for L_2-mixingales:

THEOREM 7 : Suppose $\{X_{nt}, \mathcal{F}_t^n\}$ is an L_2-mixingale such that

$$\lim_{n\to\infty}\left(a_n^{-2}\sum_{t=1}^{n}c_{nt}^2\right)\sum_{m=0}^{n-1}\psi(m)=0. \tag{19}$$

Then

$$\lim_{n\to\infty}\| a_n^{-1}\sum_{t=1}^{n}X_{nt}\|_2=0. \tag{20}$$

Adapting an argument as appearing in DE JONG [1995] and DAVIDSON and DE JONG [1995], another L_2-type weak LLN can be obtained:

THEOREM 8 : Suppose $\{X_{nt}, \mathcal{F}_t^n\}$ is an L_2-mixingale such that for some $\eta > 0$

$$\lim_{n\to\infty}\left(a_n^{-2}\sum_{t=1}^{n}c_{nt}^2\right)\sum_{m=0}^{n}(\log(m+2))^{1+\eta}\psi(m)^2=0. \tag{21}$$

Then

$$\lim_{n\to\infty}\| a_n^{-1}\sum_{t=1}^{n}X_{nt}\|_2=0. \tag{20}$$

Clearly, Theorem 8 improves upon Theorem 7 if $\psi(m) = O((\log(m+2))^{-1-\eta})$ for some $\eta > 0$. One may conjecture that it is possible to eliminate the $(\log(m+2))^{1+\eta}$ term in the above theorem, which would render a result that dominates both Theorem 7 and Theorem 8. However, in spite of his efforts, the author failed to obtain such a result. If we compare Theorem 6 for the case $p = 2$ with Theorem 7, it is easily seen that none of these theorems dominates the other. If $\psi(m) = (m+1)^{-1}$, the central condition of Theorem 6 reduces to the condition

$$\lim_{n\to\infty}a_n^{-2}\sum_{t=1}^{n}c_{nt}^2(\log(n+1))^2=0, \tag{23}$$

while the central condition of Theorem 7 reduces to

$$\lim_{n\to\infty}a_n^{-2}\sum_{t=1}^{n}c_{nt}^2\log(n+1)=0, \tag{24}$$

and therefore Theorem 7 renders a better result than Theorem 6 for this case. Note that Theorem 8 states that for this $\psi(m)$ sequence, the condition

$$\lim_{n\to\infty} a_n^{-2} \sum_{t=1}^{n} c_{nt}^2 = 0 \tag{25}$$

suffices. If $\psi(m) = (\log(m+2))^{-1}$ however, the central condition of Theorem 6 reduces to

$$\lim_{n\to\infty} a_n^{-2} \sum_{t=1}^{n} c_{nt}^2 n(\log(n+1))^{-2} = 0, \tag{26}$$

while the central condition of Theorem 8 reduces to

$$\lim_{n\to\infty} a_n^{-2} \sum_{t=1}^{n} c_{nt}^2 n(\log(n+1))^{-1+\eta} = 0 \tag{27}$$

for some $\eta > 0$, and therefore Theorem 6 gives the best result in this case. Note that for this case, the central condition of Theorem 7 reduces to

$$\lim_{n\to\infty} a_n^{-2} \sum_{t=1}^{n} c_{nt}^2 n(\log(n+1))^{-1} = 0. \tag{28}$$

Moreover, if we assume that $a_n = n^{-1}$ and $c_{nt} = t^{\alpha}$ and $\psi(m) = m^{-\beta}$ for $\alpha, \beta \in (0,1)$, then for the case $p = 2$, from Theorems 6 and 7, the condition $\alpha < \beta/2$ is obtained from the central conditions of these theorems, while the condition $\alpha < \beta$ is obtained from Theorem 8. Clearly, the last result is the best one available for this case. Note that in the cases discussed above, the best results are obtained from Theorems 6 and 8. However, as was pointed out by a referee, for the case

$$\psi(m) = (1/(\log k))I(m = 2^k \text{ for some integer } k)$$

and $p = 2$, Theorem 7 will dominate both Theorem 6 and 8. Finally, note that a fair comparison of Theorem 5 with Theorems 6, 7 and 8 is complicated in view of the integrability-type condition (5.1) that appears in that theorem.

4 Results for NED Random Variables

For applications such as stating conditions for consistency and asymptotic normality, the mixingale results are not directly useful; instead, weak LLNs for NED or mixing random variables are required. The weak LLNs for mixingales can be transformed easily to weak LLN results for NED random variables by using the inequalities of Equations (4) and (5). The conditions that are needed are listed in the following assumptions, that are derived straightforwardly from the conditions for the weak LLNs of the previous section.

ASSUMPTION 1 : There exists a triangular array B_{nt} such that
1. $\lim_{K\to\infty} \limsup_{n\to\infty} a_n^{-1} \sum_{t=1}^n \parallel X_{nt} I(|X_{nt}| > B_{nt}K) \parallel_p = 0$;
2. $\lim_{n\to\infty} a_n^{-2} \sum_{t=1}^n B_{nt}^2 = 0$.
3. For all $K > 0$, $\lim_{n\to\infty} a_n^{-1} \sum_{t=1}^n \parallel X_{nt} \parallel_r \nu([K(a_n{}^{-2} \sum_{t=1}^n B_{nt}^2)^{-1/2}]) = 0$;
4. For all $K > 0$,

$$\lim_{n\to\infty} a_n^{-1} \sum_{t=1}^n \parallel X_{nt} \parallel_r \alpha([K(a_n^{-2} \sum_{t=1}^n B_{nt}^2)^{-1/2}])^{1/p-1/r} = 0,$$

or
$\lim_{n\to\infty} a_n^{-1} \sum_{t=1}^n \parallel X_{nt} \parallel_r \phi([K(a_n^{-2} \sum_{t=1}^n B_{nt}^2)^{-1/2}])^{1-1/r} = 0.$

ASSUMPTION 2 :
1. $\lim_{n\to\infty}(a_n^{-p} \sum_{t=1}^n \parallel X_{nt} \parallel_r^p)^{1/p} \sum_{m=0}^{[n^{(p-1)/p}]+1} \nu(m) = 0$;
2. $\lim_{n\to\infty}(a_n^{-p} \sum_{t=1}^n \parallel X_{nt} \parallel_r^p)^{1/p} \sum_{m=0}^{[n^{(p-1)/p}]+1} \alpha(m)^{1/p-1/r} = 0$,
or
$\lim_{n\to\infty}(a_n^{-p} \sum_{t=1}^n \parallel X_{nt} \parallel_r^p)^{1/p} \sum_{m=0}^{[n^{(p-1)/p}]+1} \phi(m)^{1-1/r} = 0.$

ASSUMPTION 3 :
1. $p = 2$;
2. $\lim_{n\to\infty}(a_n^{-2} \sum_{t=1}^n \parallel X_{nt} \parallel_r^2) \sum_{m=0}^{n-1} \nu(m) = 0$;
3. $\lim_{n\to\infty}(a_n^{-2} \sum_{t=1}^n \parallel X_{nt} \parallel_r^2) \sum_{m=0}^{n-1} \alpha(m)^{1/2-1/r} = 0$,
or
$\lim_{n\to\infty}(a_n^{-2} \sum_{t=1}^n \parallel X_{nt} \parallel_r^2) \sum_{m=0}^{n-1} \phi(m)^{1-1/r} = 0.$

ASSUMPTION 4 :
1. $p = 2$;
2. $\lim_{n\to\infty}(a_n^{-2} \sum_{t=1}^n \parallel X_{nt} \parallel_r^2) \sum_{m=0}^{n} \nu(m)^2 (\log(m+2))^{1+\eta} = 0$;
3. $\lim_{n\to\infty}(a_n^{-2} \sum_{t=1}^n \parallel X_{nt} \parallel_r^2) \sum_{m=0}^{n} \alpha(m)^{1-2/r} (\log(m+2))^{1+\eta} = 0$,
or
$\lim_{n\to\infty}(a_n^{-2} \sum_{t=1}^n \parallel X_{nt} \parallel_r^2) \sum_{m=0}^{n} \phi(m)^{2-2/r} (\log(m+2))^{1+\eta} = 0.$

The following theorem is a simple combination of the inequalities of Equations (4) and (5) and Theorems 5, 6, 7, and 8.

THEOREM 9 : Suppose that X_{nt} is a triangular array of random variables that are L_p-NED on a strong or uniform mixing array V_{nt}, where $1 \leq p \leq 2$ and $p \leq r$. In addition, assume that $d_{nt}/ \parallel X_{nt} \parallel_r$ is bounded uniformly in t and n. Then if Assumption 1, 2, 3 or 4 holds, we have

$$\lim_{n\to\infty} \parallel a_n^{-1} \sum_{t=1}^n X_{nt} \parallel_p = 0. \tag{29}$$

Proof of Theorem 5:

Note that for all B_{nt}, K, and integer-valued $b \geq 1$

$$
\begin{aligned}
(30) \qquad a_n^{-1}\sum_{t=1}^{n} X_{nt} &= a_n^{-1}\sum_{t=1}^{n}(X_{nt} - E(X_{nt}|\mathcal{F}_{t+b-1}^{n})) \\
&\quad + a_n^{-1}\sum_{t=1}^{n}(E(X_{nt}I(|X_{nt}| \leq B_{nt}K)|\mathcal{F}_{t+b-1}^{n}) \\
&\quad - E(X_{nt}I(|X_{nt}| \leq B_{nt}K)|\mathcal{F}_{t-b}^{n})) \\
&\quad + a_n^{-1}\sum_{t=1}^{n}(E(X_{nt}I(|X_{nt}| > B_{nt}K)|\mathcal{F}_{t+b-1}^{n}) \\
&\quad - E(X_{nt}I(|X_{nt}| > B_{nt}K)|\mathcal{F}_{t-b}^{n})) \\
&\quad + a_n^{-1}\sum_{t=1}^{n} E(X_{nt}|\mathcal{F}_{t-b}^{n}) \\
&\equiv T_{n1} + T_{n2} + T_{n3} + T_{n4}.
\end{aligned}
$$

Because of condition (5.1) it is possible to choose K_ε so large that $\sup_{b\geq 1} \| T_{n3} \|_p < \varepsilon$. Choose

$$
(31) \qquad b_n^\varepsilon = \max\left(1, \left[2^{-1}\varepsilon\left(a_n^{-2}\sum_{t=1}^{n} K_\varepsilon^2 B_{nt}^2\right)^{-1/2}\right]\right).
$$

Define

$$
\begin{aligned}
(32) \qquad y_{jn} = a_n^{-1}\sum_{t=1}^{n}&(E(X_{nt}I(|X_{nt}| \leq K_\varepsilon B_{nt})|\mathcal{F}_{t+j}^{n}) \\
&- E(X_{nt}I(|X_{nt}| \leq K_\varepsilon B_{nt})|\mathcal{F}_{t+j-1}^{n})),
\end{aligned}
$$

and note that

$$
(33) \qquad T_{n2} = \sum_{j=-b_n+1}^{b_n-1} y_{jn}.
$$

Now note that

$$
\begin{aligned}
(34) \quad \limsup_{n\to\infty} \| T_{n2} \|_p &\leq \limsup_{n\to\infty} \| T_{n2} \|_2 \leq \limsup_{n\to\infty} \sum_{j=-b_n+1}^{b_n-1} \| y_{jn} \|_2 \\
&\leq \limsup_{n\to\infty} 2b_n^\varepsilon \left(a_n^{-2}\sum_{t=1}^{n} K_\varepsilon^2 B_{nt}^2\right)^{1/2} \leq \varepsilon
\end{aligned}
$$

because $[x] \leq x$, $p \leq 2$ and $\lim_{n\to\infty} a_n^{-2} \sum_{t=1}^n B_{nt}^2 = 0$. Finally note that

$$\text{(35)} \qquad \begin{aligned} &\lim_{n\to\infty} (\| T_{n1} \|_p + \| T_{n4} \|_p) \\ &\leq \lim_{n\to\infty} a_n^{-1} \sum_{t=1}^n c_{nt} [\psi(b_n^\varepsilon) + \psi(b_n^\varepsilon)] = 0 \end{aligned}$$

by Assumption 3 of Theorem 5. □

Proof of Theorem 6:

Define $b_n = [1 + n^{(p-1)/p}]$, and note that

$$\text{(36)} \qquad \begin{aligned} & a_n^{-1} \sum_{t=1}^n X_{nt} \\ &= a_n^{-1} \sum_{t=1}^n E(X_{nt} | \mathcal{F}_{t-b_n}^n) \\ &\quad + a_n^{-1} \sum_{t=1}^n \sum_{j=1}^{b_n-1} (E(X_{nt} | \mathcal{F}_{t+j}^n) - E(X_{nt} | \mathcal{F}_{t+j-1}^n)) \\ &\quad + a_n^{-1} \sum_{t=1}^n \sum_{j=-b_n+1}^{0} (E(X_{nt} | \mathcal{F}_{t+j}^n) - E(X_{nt} | \mathcal{F}_{t+j-1}^n)) \\ &\quad + a_n^{-1} \sum_{t=1}^n (X_{nt} - E(X_{nt} | \mathcal{F}_{t+b_n-1}^n)) \\ &\equiv T_{n1} + T_{n2} + T_{n3} + T_{4n}. \end{aligned}$$

Next, note that

$$\text{(37)} \qquad \begin{aligned} \| T_{n2} \|_p &\leq a_n^{-1} \sum_{j=1}^{b_n-1} \| \sum_{t=1}^n E(X_{nt} | \mathcal{F}_{t+j}^n) - E(X_{nt} | \mathcal{F}_{t+j-1}^n) \|_p \\ &= O\left(a_n^{-1} \sum_{j=0}^{b_n-1} \left(\sum_{t=1}^n E |E(X_{nt} | \mathcal{F}_{t+j}^n) - E(X_{nt} | \mathcal{F}_{t+j-1}^n)|^p \right)^{1/p} \right) \\ &= O\left(\sum_{j=0}^{b_n} \psi(j) \left(a_n^{-p} \sum_{t=1}^n c_{nt}^p \right)^{1/p} \right) = o(1) \end{aligned}$$

by assumption. Note that the first inequality is the norm inequality, the first equality is Equation (16.46) of DAVIDSON [1994], and the second equality uses some rearranging and the mixingale definition, and the conclusion follows from the condition of the theorem. Similarly,

$$\begin{aligned}
(38)\qquad \| T_{n3} \|_p &\leq a_n^{-1} \sum_{j=0}^{b_n-1} \| \sum_{t=1}^{n} E(X_{nt}|\mathcal{F}_{t-j}^n) - E(X_{nt}|\mathcal{F}_{t-j-1}^n) \|_p \\
&= O\left(a_n^{-1} \sum_{j=0}^{b_n-1} \left(\sum_{t=1}^{n} E(|E(X_{nt}|\mathcal{F}_{t-j}^n) - E(X_{nt}|\mathcal{F}_{t-j-1}^n)|)^p \right)^{1/p} \right) \\
&= O\left(\sum_{j=0}^{b_n} \psi(j) \left(a_n^{-p} \sum_{t=1}^{n} c_{nt}^p \right)^{1/p} \right) = o(1)
\end{aligned}$$

by assumption. Moreover,

$$\begin{aligned}
(39)\qquad \| T_{n1} \|_p &= \| a_n^{-1} \sum_{t=1}^{n} E(X_{nt}|\mathcal{F}_{t-b_n}^n) \|_p \\
&= \sum_{j=1}^{b_n} b_n^{-1} n a_n^{-1} \| n^{-1} \sum_{t=1}^{n} E(X_{nt}|\mathcal{F}_{t-b_n}^n) \|_p \\
&\leq \sum_{j=1}^{b_n} b_n^{-1} n a_n^{-1} \left(n^{-1} \sum_{t=1}^{n} E|E(X_{nt}|\mathcal{F}_{t-j}^n)|^p \right)^{1/p} \\
&\leq \sum_{j=1}^{b_n} \psi(j) \left(a_n^{-p} b_n^{-p} n^{p-1} \sum_{t=1}^{n} c_{nt}^p \right)^{1/p} \\
&= O\left(\sum_{j=1}^{b_n} \psi(j) \left(a_n^{-p} \sum_{t=1}^{n} c_{nt}^p \right)^{1/p} \right) = o(1)
\end{aligned}$$

where the first inequality is the conditional expectations property and Jensen's inequality, the second is the mixingale definition and the conclusion follows by assumption. Finally, note that

$$\begin{aligned}
(40)\ \| T_{n4} \|_p &= \| a_n^{-1} \sum_{t=1}^{n} (X_{nt} - E(X_{nt}|\mathcal{F}_{t+b_n-1}^n)) \|_p \\
&= O\Bigg(\sum_{j=0}^{b_n-1} b_n^{-1} n a_n^{-1} \| n^{-1} \sum_{t=1}^{n} |X_{nt} - E(X_{nt}|\mathcal{F}_{t+j}^n)| \\
&\qquad + |E(X_{nt}|\mathcal{F}_{t+j}^n) - E(X_{nt}|\mathcal{F}_{t+b_n-1}^n)| \|_p \Bigg) \\
&= O\left(\sum_{j=0}^{b_n-1} n b_n^{-1} a_n^{-1} \left(n^{-1} \sum_{t=1}^{n} E|X_{nt} - E(X_{nt}|\mathcal{F}_{t+j}^n)|^p \right)^{1/p} \right) \\
&= O\left(\sum_{j=0}^{b_n} \psi(j) \left(a_n^{-p} \sum_{t=1}^{n} c_{nt}^p \right)^{1/p} \right) = o(1)
\end{aligned}$$

by assumption. The second equality is some rearranging and the triangle inequality, and the second equality uses conditional expectations properties and Jensen's inequality, and the conclusion follows by assumption. □

Proof of Theorem 7:

Firstly note that for all $m \geq 0$

$$
\begin{aligned}
(41) \quad EX_{nt}X_{n,t+m} &= E(X_{nt} - E(X_{nt}|\mathcal{F}^n_{t+[m/2]}))X_{n,t+m} \\
&\quad + EE(X_{nt}|\mathcal{F}^n_{t+[m/2]})X_{n,t+m} \\
&\leq \| X_{n,t+m} \|_2 \| X_{nt} - E(X_{nt}|\mathcal{F}^n_{t+[m/2]}) \|_2 \\
&\quad + EE(X_{nt}|\mathcal{F}^n_{t+[m/2]})E(X_{n,t+m}|\mathcal{F}^n_{t+[m/2]}) \\
&\leq (\psi(0)+\psi(1))c_{n,t+m}c_{nt}\psi([m/2]+1) \\
&\quad + \| E(X_{nt}|\mathcal{F}^n_{t+[m/2]}) \|_2 \| E(X_{n,t+m}|\mathcal{F}^n_{t+[m/2]}) \|_2 \\
&\leq (\psi(0)+\psi(1))c_{n,t+m}c_{nt}\psi([m/2]+1) \\
&\quad + (\psi(0)+\psi(1))c_{nt}c_{n,t+m}\psi(m-[m/2]) \\
&= c_{nt}c_{n,t+m}\psi'(m),
\end{aligned}
$$

where the first equality is rearranging of the expression, the first inequality is Cauchy-Schwartz's and rearranging, the second inequality uses the L_2-mixingale definition (from which it follows that $\| X_{nt} \|_2 \leq (\psi(0) + \psi(1))c_{nt})$ and the Cauchy-Schwartz inequality, and the third inequality uses the mixingale definition. In the last equality, we defined

$$(42) \qquad \psi'(m) = (\psi(0)+\psi(1))(\psi(m-[m/2]) + \psi([m/2]+1)).$$

Next, note that

$$
\begin{aligned}
(43)\ E\left(a_n^{-1}\sum_{t=1}^{n} X_{nt}\right)^2 &\leq 2a_n^{-2}\sum_{t=1}^{n}\sum_{m=0}^{n-t} EX_{nt}X_{n,t+m} \\
&\leq 2a_n^{-2}\sum_{t=1}^{n}\sum_{m=0}^{n-t} c_{nt}c_{n,t+m}\psi'(m) \\
&= 2a_n^{-2}\sum_{m=0}^{n-1}\psi'(m)\sum_{t=1}^{n-m} c_{nt}c_{n,t+m} \\
&\leq 2a_n^{-2}\sum_{m=0}^{n-1}\psi'(m)\left(\sum_{t=1}^{n-m} c_{nt}^2\right)^{1/2}\left(\sum_{t=1}^{n-m} c_{n,t+m}^2\right)^{1/2} \\
&= O\left(a_n^{-2}\sum_{m=0}^{n-1}\psi(m)\left(\sum_{t=1}^{n} c_{nt}^2\right)\right) = o(1)
\end{aligned}
$$

by assumption, where the third inequality is Cauchy-Schwartz's. □

Proof of Theorem 8:

Firstly note that

$$(44)\quad a_n^{-1}\sum_{t=1}^{n}X_{nt} = a_n^{-1}\sum_{t=1}^{n}E(X_{nt}|\mathcal{F}_{t-n}^{n}) + a_n^{-1}\sum_{t=1}^{n}\sum_{j=1}^{n-1}(E(X_{nt}|\mathcal{F}_{t+j}^{n}) - E(X_{nt}|\mathcal{F}_{t+j-1}^{n})) + a_n^{-1}\sum_{t=1}^{n}\sum_{j=-n+1}^{0}(E(X_{nt}|\mathcal{F}_{t+j}^{n}) - E(X_{nt}|\mathcal{F}_{t+j-1}^{n})) + a_n^{-1}\sum_{t=1}^{n}(X_{nt} - E(X_{nt}|\mathcal{F}_{t+n-1}^{n})) \equiv T_{n1} + T_{n2} + T_{n3} + T_{n4},$$

and also note that for all integer-valued j

$$(45)\quad E(E(X_{nt}|\mathcal{F}_{t+j}^{n}) - E(X_{nt}|\mathcal{F}_{t+j-1}^{n}))^2 = E(E(X_{nt}|\mathcal{F}_{t+j}^{n}))^2 - E(E(X_{nt}|\mathcal{F}_{t+j-1}^{n}))^2 = E(X_{nt} - E(X_{nt}|\mathcal{F}_{t+j-1}^{n}))^2 - E(X_{nt} - E(X_{nt}|\mathcal{F}_{t+j}^{n}))^2.$$

See also HALL and HEYDE [1980, p. 21]. Let $\xi_j = (|j|+1)^{-1}(\log(|j|+2))^{-1-\eta}$. Then

(46)

$$ET_{n2}^2 = E\left(\sum_{j=1}^{n-1}a_n^{-1}\sum_{t=1}^{n}(E(X_{nt}|\mathcal{F}_{t+j}^{n}) - E(X_{nt}|\mathcal{F}_{t+j-1}^{n}))\right)^2 \le \left(\sum_{j=1}^{n-1}\xi_j\right)\left(\sum_{j=1}^{n-1}\xi_j^{-1}a_n^{-2}\sum_{t=1}^{n}E(E(X_{nt}|\mathcal{F}_{t+j}^{n}) - E(X_{nt}|\mathcal{F}_{t+j-1}^{n}))^2\right)$$
$$= O\bigg(a_n^{-2}\sum_{t=1}^{n}\sum_{j=1}^{n-1}\xi_j^{-1}[E(X_{nt} - E(X_{nt}|\mathcal{F}_{t+j-1}^{n}))^2 - E(X_{nt} - E(X_{nt}|\mathcal{F}_{t+j}^{n}))^2]\bigg)$$
$$= O\bigg(a_n^{-2}\sum_{t=1}^{n}\bigg[a_0\psi(0)^2c_{nt}^2 + \sum_{j=1}^{n-1}(\log(j+2))^{1+\eta}E(X_{nt} - E(X_{nt}|\mathcal{F}_{t+j}^{n}))^2\bigg]\bigg)$$
$$= O\bigg(a_n^{-2}\sum_{t=1}^{n}c_{nt}^2 + a_n^{-2}\sum_{t=1}^{n}c_{nt}^2\sum_{j=1}^{n-1}(\log(j+2))^{1+\eta}\psi(j+1)^2\bigg) = o(1)$$

where the first inequality is Cauchy-Schwartz's of the type

$$\left(\sum_{j=1}^{n-1} b_j\right)^2 \le \left(\sum_{j=1}^{n-1} \xi_j\right)\left(\sum_{j=1}^{n-1} \xi_j^{-1} b_j^2\right),$$

and the conclusion follows by assumption. A similar argument holds for T_{n3}:

$$\begin{aligned}(47)\quad ET_{n3}^2 &= E\left(\sum_{j=0}^{n-1} a_n^{-1} \sum_{t=1}^{n}(E(X_{nt}|\mathcal{F}_{t-j}^n) - E(X_{nt}|\mathcal{F}_{t-j-1}^n))\right)^2 \\ &= O\left(a_n^{-2} \sum_{t=1}^{n}\sum_{j=0}^{n-1} \xi_j^{-1} E(E(X_{nt}|\mathcal{F}_{t-j}^n) - E(X_{nt}|\mathcal{F}_{t-j-1}^n))^2\right) \\ &= O\left(a_n^{-2} \sum_{t=1}^{n}[c_{nt}^2 \psi(0)^2 + \sum_{j=0}^{n-1}(\log(j+2))^{-1-\eta} E(E(X_{nt}|\mathcal{F}_{t-j}^n))^2\right) \\ &= O\left(a_n^{-2} \sum_{t=1}^{n} c_{nt}^2 + a_n^{-2} \sum_{t=1}^{n} c_{nt}^2 \sum_{j=0}^{n-1}(\log(j+2)^{-1-\eta}\psi(j)^2\right) = o(1).\end{aligned}$$

Moreover,

$$\begin{aligned}(48)\qquad ET_{n1}^2 &\le a_n^{-2} \sum_{t=1}^{n}(nE|E(X_{nt}|\mathcal{F}_{t-n}^n)|^2) \\ &\le a_n^{-2} \sum_{t=1}^{n}\sum_{j=0}^{n} E|E(X_{nt}|\mathcal{F}_{t-j}^n)|^2 \\ &\le a_n^{-2} \sum_{t=1}^{n} c_{nt}^2 \sum_{j=0}^{n} \psi(j)^2 = o(1)\end{aligned}$$

by assumption. The first inequality is Jensen's, the second uses conditional expectations properties. The proof is completed by noting that

$$\begin{aligned}(49)\quad ET_{n4}^2 &\le a_n^{-2} \sum_{t=1}^{n}(nE|X_{nt} - E(X_{nt}|\mathcal{F}_{t+n-1}^n)|^2) \\ &\le a_n^{-2} \sum_{t=1}^{n}\sum_{j=0}^{n-1}(E|X_{nt} - E(X_{nt}|\mathcal{F}_{t+j}^n)|^2 \\ &\qquad + E|E(X_{nt}|\mathcal{F}_{t+j}^n) - E(X_{nt}|\mathcal{F}_{t+n-1}^n)|^2) \\ &= O\left(a_n^{-2} \sum_{t=1}^{n} c_{nt}^2 \sum_{j=0}^{n-1} \psi(j+1)^2\right) = o(1)\end{aligned}$$

by assumption, where the steps are similar to the that of Equation (48). □

• References

ANDREWS, D.W.K. (1985). – "A Nearly Independent, but Non-Strong Mixing, Triangular Array", *Journal of Applied Probability*, 21, pp. 930-934.

ANDREWS, D.W.K. (1988). – "Laws of Large Numbers for Dependent Non-Identically Distributed Random Variables", *Econometric Theory*, 4, pp. 458-467.

DAVIDSON, J. (1992). – "A Central Limit Theorem for Globally Nonstationary Near-Epoch Dependent Functions of Mixing Processes", *Econometric Theory*, 8, pp. 313-334.

DAVIDSON, J. (1993a). – "An L_1-Convergence Theorem for Heterogeneous Mixingale arrays with Trending Moments", *Statistics and Probability Letters*, 16, pp. 301-306.

DAVIDSON, J. (1993b). – "The Central Limit Theorem for Globally Nonstationary Near-Epoch Dependent Functions of Mixing Processes: the Asymptotically Degenerate Case", *Econometric Theory*, 9, pp. 402-412.

DAVIDSON, J. (1994). – *Stochastic Limit Theory*. Oxford: Oxford University Press.

DAVIDSON, J., DE JONG, R.M. (1995). – "Strong Laws of Large Numbers for Dependent Heterogeneous Processes: a Synthesis of New Results", *Working paper*.

DE JONG, R.M. (1995). – "Laws of Large Numbers for Dependent Heterogeneous Processes", *Econometric Theory*, 11, pp. 347-358.

DE JONG, R.M. (1997). – "Central Limit Theorems for Dependent Heterogeneous Random Variables", *Econometric Theory*, 13, pp. 353-367.

GALLANT, A.R., WHITE, H. (1988). – *A unified theory of estimation and inference for nonlinear dynamic models*. New York: Basil Blackwell.

HANSEN, B.E. (1991). – "Strong Laws for Dependent Heterogeneous Processes", *Econometric Theory*, 7, pp. 213-221.

HANSEN, B.E. (1992). – "Strong Laws for Dependent Heterogeneous Processes. Erratum", *Econometric Theory*, 8, pp. 421-422.

MCLEISH, D.L. (1975). – "A Maximal Inequality and Dependent Strong Laws", *Annals of Probability*, 3, pp. 829-839.

PÖTSCHER, B.M., PRUCHA, I.R. (1991a). – "Basic Structure of the Asymptotic Theory in Dynamic Nonlinear Econometric Models, part 1: Consistency and Approximation Concepts", *Econometric Reviews*, 10, 125-216.

PÖTSCHER, B.M., PRUCHA, I.R. (1991b). – "Basic Structure of the Asymptotic Theory in Dynamic Nonlinear Econometric Models, part 2: Asymptotic Normality", *Econometric Reviews*, 10, pp. 253-325.