Cowles Foundation Paper 717

Econometric Theory, **4**, 1988, 458–467 Printed in the United States of America.

LAWS OF LARGE NUMBERS FOR DEPENDENT NON-IDENTICALLY DISTRIBUTED RANDOM VARIABLES

DONALD W.K. ANDREWS
Cowles Foundation, Yale University

This paper provides L^1 and weak laws of large numbers for uniformly integrable L^1-mixingales. The L^1-mixingale condition is a condition of asymptotic weak temporal dependence that is weaker than most conditions considered in the literature. Processes covered by the laws of large numbers include martingale difference, $\phi(\cdot)$, $\rho(\cdot)$, and $\alpha(\cdot)$ mixing, autoregressive moving average, infinite-order moving average, near epoch dependent, L^1-near epoch dependent, and mixingale sequences and triangular arrays. The random variables need not possess more than one finite moment and the L^1-mixingale numbers need not decay to zero at any particular rate. The proof of the results is remarkably simple and completely self-contained.

1. INTRODUCTION

The importance of laws of large numbers (LLNs) in econometric theory requires little explanation. LLNs and uniform LLNs are used extensively in establishing consistency and asymptotic normality of parametric and nonparametric estimators in all types of econometric models. In addition, they are used in establishing the asymptotic distributions of test statistics under the null and local alternative hypotheses.

This paper presents new L^1 and weak LLNs for dependent non-identically distributed sequences and triangular arrays of random variables (rv's). The results for triangular arrays are needed when sequences of local alternatives are considered. An analogue of McLeish's [10,12] mixingale condition, called an L^1-mixingale, is introduced for sequences and arrays of rv's. This condition is weaker than the mixingale condition.

Examples of L^1-mixingales include: martingale difference sequences, mean-zero M-dependent sequences, mean-zero stationary Gaussian processes whose correlations converge to zero as the time span increases to infinity, mean-zero $\phi(\cdot)$, $\rho(\cdot)$, and $\alpha(\cdot)$ mixing sequences that are L^p bounded (i.e.,

I would like to thank the referees, Hal White, and In Choi for helpful comments, the California Institute of Technology for their hospitality while this research was undertaken, and the Alfred P Sloan Foundation and the National Science Foundation for their financial support through a Research Fellowship and Grant SES-8618617, respectively

$\sup_{i\geq 1} E|X_i|^p < \infty$) for some $p > 1$, mean-zero near-epoch-dependent sequences, mean-zero L^p-near-epoch-dependent sequences for $p \geq 1$, mixingales, and infinite-order moving average processes whose coefficients are absolutely summable and whose innovations are L^p bounded for some $p > 1$ including autoregressive moving average processes.

It is shown that uniformly integrable L^1-mixingales satisfy an L^1 LLN and in consequence a weak LLN given a flexible condition on the relative magnitudes of the rv's. In contrast to McLeish's [10] conditions for a strong LLN for mixingales, our conditions do not require the mixingale numbers $\{\psi_m\}$ to decay to zero at a particular rate, they do not require the rv's to be square integrable, and they apply to both sequences and triangular arrays of rv's rather than to just sequences. Hence, for our results there is no tradeoff between the temporal dependence of a sequence as measured by the decay rate of $\{\psi_m\}$ and the number of finite moments needed for an LLN, as arises in McLeish's strong LLN. (These comments are not to be interpreted as a criticism of McLeish's results, since McLeish was concerned with a mixingale convergence theorem and strong LLN whereas we are concerned with L^1 and weak LLNs.)

The weaker moment conditions of our results are particularly convenient for use with estimators of covariance matrices of other estimators because they only require as many finite moments as typically are needed for the asymptotic normality of the "other" estimators (e.g., they can be applied in Gallant and White's [6] Theorems 6.3, 6.4, and 6.8). The present results also can be used to weaken the assumptions needed to establish the consistency of estimators in nonlinear econometric models, e.g., see Andrews and Fair [2].

The remainder of this paper is organized as follows: Section 2 defines L^1-mixingales and presents the LLNs. Section 3 discusses a number of examples of L^1-mixingales. Section 4 proves the LLNs given in Section 2.

2. LAWS OF LARGE NUMBERS FOR L^1-MIXINGALES

For clarity of presentation, we first consider sequences of rv's. Triangular arrays are discussed below. Let $(\Omega,\boldsymbol{F},P)$ denote a probability space. Let $\{X_i: i \geq 1\}$ be a sequence of rv's on $(\Omega,\boldsymbol{F},P)$. Let $\{\boldsymbol{F}_i: i = \ldots,0,1,\ldots\}$ be any nondecreasing sequence of sub σ-fields of $\boldsymbol{F}$. Often one will take $\boldsymbol{F}_i = \sigma(X_1,\ldots,X_i)$ for $i \geq 1$ and $\boldsymbol{F}_i = \{\phi,\Omega\}$ *for* $i \leq 0$. Let $E(X_i|\boldsymbol{F}_j)$ denote the conditional expectation of X_i given $\boldsymbol{F}_j$. Whenever $E(X_i|\boldsymbol{F}_j)$ is used, it is to be understood that X_i is assumed to be integrable. Let $\|\cdot\|_p$ denote the $L^p(P)$ norm, i.e., $\|X_i\|_p = (E|X_i|^p)^{1/p}$.

DEFINITION 1. *The sequence* $\{X_i,\boldsymbol{F}_i\}$ *is an* L^1*-mixingale 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) $\|E(X_i|\boldsymbol{F}_{i-m})\|_1 \le c_i\psi_m$,

(b) $\|X_i - E(X_i|\boldsymbol{F}_{i+m})\|_1 \le c_i\psi_{m+1}$.

Comments: (1) A mixingale sequence, as defined by McLeish [10], is one that satisfies the L^1-mixingale condition with $\|\cdot\|_1$ replaced by the $L^2(P)$ norm $\|\cdot\|_2$ in (a) and (b) above.

(2) Condition (b) usually holds trivially since X_i is almost always $\boldsymbol{F}_i$ measurable (which implies $E(X_i|\boldsymbol{F}_{i+m}) = X_i$ a.s.). One example where X_i is not $\boldsymbol{F}_i$ measurable (but condition (b) still holds) is a doubly infinite moving average process where $\boldsymbol{F}_i$ is the σ-field generated by the innovations with indices $s \le i$, see Example 5 of Section 3.

(3) L^1-mixingales are necessarily sequences of mean zero rv's, since $|EX_i| = |EE(X_i|\boldsymbol{F}_{i-m})| \le E|E(X_i|\boldsymbol{F}_{i-m})| \le c_i\psi_m \to 0$ as $m \to \infty$. Thus, to apply results for L^1-mixingales to a sequence of rv's $\{Z_i\}$, one must consider the sequence $\{Z_i - EZ_i\}$.

(4) The constants $\{\psi_m\}$ are referred to as the L^1-mixingale numbers. These numbers index the temporal dependence of the sequence $\{X_i\}$. Clearly, if $\{X_i\}$ is independent and $\boldsymbol{F}_i \supset \sigma(X_1,\ldots,X_i)$, then one can choose $\psi_m = 0$ for all $m > 0$. The constants $\{c_i\}$ are chosen to index the "magnitude" of the rv's $\{X_i\}$. For example, one can take $c_i = \|X_i\|_{p_i}$ for some $p_i \ge 1$ for $i = 1,2,\ldots$. The most common choice is simply $p_i = 1$, i.e., $c_i = \|X_i\|_1$ for $i = 1,2,\ldots$.

THEOREM 1. (a) *Suppose the sequence* $\{X_i,\boldsymbol{F}_i\}$ *is a uniformly integrable L^1-mixingale. If* $\overline{\lim}_{n\to\infty} \frac{1}{n}\sum_{i=1}^{n} c_i < \infty$, *then* $E|\bar{X}_n| = E\left|\frac{1}{n}\sum_{i=1}^{n} X_i\right| \to 0$ *as* $n \to \infty$ *and in consequence* $\bar{X}_n \xrightarrow{p} 0$ *as* $n \to \infty$. (b) *If the sequence* $\{X_i,\boldsymbol{F}_i\}$ *is a uniformly integrable L^1-mixingale with constants* $\{c_i\}$ *given by* $\{\|X_i\|_1\}$, *then* $E|\bar{X}_n| \to 0$ *as* $n \to \infty$ *and in consequence* $\bar{X}_n \xrightarrow{p} 0$ *as* $n \to \infty$.

Comments: (1) No rate of decay to zero is imposed on the L^1-mixingale numbers $\{\psi_m\}$. This contrasts with many LLNs in which the constants that index the temporal dependence, such as $\phi(\cdot)$, $\rho(\cdot)$, or $\alpha(\cdot)$ mixing numbers, near-epoch-dependent numbers, or mixingale numbers must converge to zero at a particular rate, e.g., see McLeish [10].

(2) The conclusion of Theorem 1 can be strengthened to convergence in L^p for $1 < p \le 2$ if $\{X_i,\boldsymbol{F}_i\}$ is an L^p-mixingale (i.e., an L^1-mixingale with the $L^1(P)$ norm replaced by the $L^p(P)$ norm) and the uniform integrability of $\{X_i\}$ is replaced by uniform integrability of $\{|X_i|^p\}$. The proof is identical to that given for Theorem 1 in Section 4 with the $L^1(P)$ norm in equations (5) and (7) replaced by the $L^p(P)$ norm. The above conditions can be weakened if $p = 2$. In particular, $\bar{X}_n$ satisfies an L^2 LLN if $\{X_i\}$ are L^2

bounded, $\{X_i, F_i\}$ is an L^2-mixingale (i.e., a mixingale) with $c_i = \sup_{k \geq 1} \|X_k\|_2$ for all i, and the L^2-mixingale numbers $\{\psi_m\}$ satisfy $\frac{1}{n} \sum_{m=1}^{n} \psi_m \to 0$ as $n \to \infty$. See Section 4 for a proof.

(3) One can obtain a uniform weak LLN from Theorem 1 by applying results given in Andrews [1].

We now consider triangular arrays of rv's, $\{X_{ni}: i = 1, \ldots, k_n; n = 1,2,\ldots\}$, where $k_n \uparrow \infty$ as $n \to \infty$. Let $\{F_{ni}: i = \ldots, 0, 1, \ldots; n = 1,2,\ldots\}$ be an array of sub σ-fields of F such that $\{F_{ni}\}$ is nondecreasing in i for each n. In many cases, one can take $F_{ni} = F$ for all $i > k_n$ and $F_{ni} = \{\phi, \Omega\}$ for all $i \leq 0$. The prime example where this is not true is when X_{ni} depends on a singly or doubly infinite sequence of innovations, as in infinite-order moving average or autoregressive processes.

DEFINITION 2. *The triangular array* $\{X_{ni}, F_{ni}\}$ *is an* L^1*-mixingale if there exist nonnegative constants* $\{c_{ni}: i = 1, \ldots, k_n, n = 1,2,\ldots\}$ *and* $\{\psi_m: m = 0,1,\ldots\}$ *such that* $\psi_m \to 0$ *as* $m \to \infty$ *and for all* $i = 1, \ldots, k_n$, $n \geq 1$, *and* $m \geq 0$, *we have*

(a) $\|E(X_{ni} | F_{ni-m})\|_1 \leq c_{ni} \psi_m,$

(b) $\|X_{ni} - E(X_{ni} | F_{ni+m})\|_1 \leq c_{ni} \psi_{m+1}.$

Comment: This definition generalizes McLeish's [12] definition of a mixingale triangular array.

Next, we present a weak LLN for triangular arrays. Its proof is completely analogous to that of Theorem 1, and in consequence, is omitted. (Details are available from the author upon request.)

THEOREM 2. (a) *Suppose the triangular array* $\{X_{ni}, F_{ni}\}$ *is a uniformly integrable* L^1*-mixingale. If* $\overline{\lim}_{n \to \infty} \frac{1}{k_n} \sum_{i=1}^{k_n} c_{ni} < \infty$, *then* $E|\bar{X}_n| = E\left|\frac{1}{k_n} \sum_{i=1}^{k_n} X_{ni}\right| \to 0$ *as* $n \to \infty$ *and in consequence* $\bar{X}_n \xrightarrow{p} 0$ *as* $n \to \infty$. (b) *If the triangular array* $\{X_{ni}, F_{ni}\}$ *is a uniformly integrable* L^1*-mixingale with constants* $\{c_{ni}\}$ *given by* $\{\|X_{ni}\|_1\}$, *then* $E|\bar{X}_n| \to 0$ *as* $n \to \infty$ *and in consequence* $\bar{X}_n \xrightarrow{p} 0$ *as* $n \to \infty$.

Comment: Strong LLNs are not available for triangular arrays of rv's except in very special cases, e.g., see Teicher [13] and references therein. The weak LLN for triangular arrays that can be derived from McLeish's [10, Theorem 1.6] maximal inequality for mixingales requires the mixingale numbers to decline to zero at a particular rate and the rv's to be square integrable. The conditions of Theorem 2 are weaker on both counts.

3. EXAMPLES OF L^1-MIXINGALES

1. A martingale difference array $\{X_{ni}, F_{ni}: 1 \le i \le k_n, n \ge 1\}$ is an L^1-mixingale with $\psi_m = 0$ for $m \ge 1$ and $c_{ni} = \|X_{ni}\|_1$ if we set $F_{ni} = \{\phi, \Omega\}$ for $i \le 0$ and $F_{ni} = F$ for $i > k_n$. Hence, if $\{X_{ni}\}$ is uniformly integrable, then $\bar{X}_n$ satisfies L^1 and weak LLNs as $n \to \infty$. For martingale difference sequences these results are not new, e.g., see Chow [4].

2. A triangular array of mean-zero M-dependent rv's $\{X_{ni}: 1 \le i \le k_n, n \ge 1\}$ is an L^1-mixingale with $\psi_m = 0$ for $m > M$, and $c_{ni} = \|X_{ni}\|_1$ if one takes $F_{ni} = \sigma(X_{n1}, \ldots, X_{ni})$ for $1 \le i \le k_n$, $F_{ni} = \{\phi, \Omega\}$ for $i \le 0$, and $F_{ni} = F$ for $i > k_n$. Hence, if $\{X_{ni}\}$ is uniformly integrable, then $\bar{X}_n$ satisfies L^1 and weak LLNs as $n \to \infty$.

3. Suppose $\{X_i: i \ge 1\}$ is a mean-zero stationary Gaussian process. Define $F_i = \sigma(X_1, \ldots, X_i)$ for $i \ge 1$ and $F_i = \{\phi, \Omega\}$ for $i \le 0$. If the spectrum of $\{X_i: i \ge 1\}$ is continuous and nonzero, then $\{X_i, F_i\}$ is an L^1-mixingale with $c_i = \|X_i\|_1$ and exponentially declining L^1-mixingale numbers. Hence, $\bar{X}_n$ satisfies L^1 and weak LLNs. In fact, by Comment 2 of Theorem 1, it also satisfies an L^2 LLN.

It follows by Theorems 1, 2, and 4 and the discussion following Theorem 4 of Kolmogorov and Rozonov [9] that under the above conditions $\{X_i\}$ is $\alpha(\cdot)$ mixing with exponentially declining $\alpha(\cdot)$ mixing numbers. The assertions above now follow by Example 4 below and Theorem 1(b) above.

4. Let $\{X_{ni}: 1 \le i \le k_n, n = 1,2,\ldots\}$ be a triangular array of mean-zero $\phi(\cdot)$, $\rho(\cdot)$, or $\alpha(\cdot)$ mixing rv's that are L^p bounded for some $p > 1$ (see McLeish [10] and Herrndorf [8] for definitions of these mixing conditions). Define $\{F_{ni}\}$ as in Example 2. Then, $\{X_{ni}, F_{ni}\}$ is a uniformly integrable L^1-mixingale with $\{c_{ni}\} = \{\|X_{ni}\|_p\}$ and $\bar{X}_n$ satisfies L^1 and weak LLNs as $n \to \infty$. These results follow by the mixing inequalities of Lemma 2.1 of McLeish [10] for $\phi(\cdot)$ and $\alpha(\cdot)$ mixing arrays, the inequality $\rho(m) \ge 4\alpha(m)$ for $\rho(\cdot)$ mixing arrays, and Theorem 2.

5. Suppose $X_i = \sum_{j=-\infty}^{\infty} a_{ij}\epsilon_{i-j}$ for $i \ge 1$, where $\{\epsilon_j, F_j: -\infty < j < \infty\}$ is a sequence of martingale difference innovation rv's and corresponding σ-fields and $\{a_{ij}: -\infty < j < \infty, i \ge 1\}$ is a sequence of constants. If $\{\epsilon_j\}$ are L^p bounded for some $p > 1$ and $\sum_{j=-\infty}^{\infty} \sup_{i \ge 1} |a_{ij}| < \infty$, then $\{X_i, F_i\}$ is a uniformly integrable L^1-mixingale with $c_i = \sup_{-\infty < k < \infty} \|\epsilon_k\|_1$ for $i \ge 1$ and $\bar{X}_n$ satisfies L^1 and weak LLNs as $n \to \infty$. Obviously, one-sided infinite-order moving average processes are obtained by taking $a_{ij} = 0$ for all $j < 0$. Exam-

ples of such sequences include autoregressive, moving average, and autoregressive moving average processes.

To establish the assertions above, write

$$\|E(X_t|F_{t-m})\|_1 = \left\| \sum_{j=m}^{\infty} a_{tj}\epsilon_{t-j} \right\|_1 \le \sum_{j=m}^{\infty} \sup_{t\ge 1} |a_{tj}| \cdot \sup_{-\infty<k<\infty} \|\epsilon_k\|_1,$$

$$\|X_t - E(X_t|F_{t+m})\|_1 = \left\| \sum_{j=-\infty}^{-m-1} a_{tj}\epsilon_{t-j} \right\|_1 \le \sum_{j=m+1}^{\infty} \sup_{t\ge 1} |a_{t,-j}| \cdot \sup_{-\infty<k<\infty} \|\epsilon_k\|_1.$$

Let $c_t = \sup_{-\infty<k<\infty} \|\epsilon_k\|_1$ for $i \ge 1$ and $\psi_m = \sum_{j=m}^{\infty} (\sup_{t\ge 1} |a_{tj}| + \sup_{t\ge 1} |a_{t,-j}|)$ for $m \ge 0$. Then, $\{X_t, F_t\}$ is an L^1-mixingale. The uniform integrability of $\{X_t\}$ is established as follows:

$$\begin{aligned}
&\lim_{b\to\infty} \sup_{t\ge 1} E\left| \sum_{j=-\infty}^{\infty} a_{tj}\epsilon_{t-j} \right| 1(|X_t| \ge b) \\
&\quad \le \lim_{b\to\infty} \sum_{j=-\infty}^{\infty} \sup_{v\ge 1} |a_{vj}| \sup_{t\ge 1} E|\epsilon_{t-j}| 1(|X_t| \ge b), \\
&\quad \le \sum_{j=-\infty}^{\infty} \sup_{v\ge 1} |a_{vj}| \cdot \sup_{-\infty<k<\infty} \|\epsilon_k\|_p \cdot \lim_{b\to\infty} \sup_{t\ge 1} P(|X_t| \ge b)^{p/(p-1)}, \\
&\quad \le \sum_{j=-\infty}^{\infty} \sup_{v\ge 1} |a_{vj}| \cdot \sup_{-\infty<k<\infty} \|\epsilon_k\|_p \cdot \lim_{b\to\infty} (\sup_{t\ge 1} E|X_t|/b)^{p/(p-1)}, \\
&\quad = 0, \qquad\qquad (1)
\end{aligned}$$

where the equality uses the fact that $\sup_{t\ge 1} E|X_t| < \infty$ (which follows from the same argument as in equation (1) with "$\lim_{b\to\infty}$ " and "$1(|X_t| > b)$" removed).

6. Here we introduce a generalization of near-epoch-dependent (NED) sequences and triangular arrays of rv's that allow the rv's to have less than two moments finite. The results of Section 2 are shown to establish L^1 and weak LLNs for such rv's. The NED condition was introduced by Billingsley [3] (under the rubric "functions of mixing processes"). McLeish [10–12] provides an LLN, a central limit, and an invariance principle for NED sequences that extend Billingsley's results considerably. More recent developments are given in Wooldridge and White [14], Gallant [5], and Gallant and White [6]. The latter two references give numerous examples of NED sequences and arrays that arise in econometrics. The examples include nonlinear dynamic regression and simultaneous equations models under suitable conditions on the functional forms and underlying rv's in the models.

Suppose $\{Y_{ni}: -\infty < i < \infty, n \geq 1\}$ is an $\alpha(\cdot)$ mixing array of rv's, i.e.,

$$\alpha(m) = \sup_{-\infty<i<\infty, n\geq 1} \sup_{A\in F^n_{-\infty,i}, B\in F^n_{i+m,\infty}} |P(A \cap B) - P(A)P(B)| \to 0$$
$$\text{as } m \to \infty,$$

where $F^n_{i,j} = \sigma(Y_{ni}, \ldots, Y_{nj})$, $F^n_{-\infty,i} = \sigma(\ldots, Y_{ni-1}, Y_{ni})$, and $F^n_{j,\infty} = \sigma(Y_{nj}, Y_{nj+1}, \ldots)$ for $-\infty < i \leq j < \infty, n \geq 1$. A triangular array $\{X_{ni}: 1 \leq i \leq k_n, n \geq 1\}$ is L^p-*near epoch dependent* if there exists an $\alpha(\cdot)$ mixing array $\{Y_{ni}\}$ with corresponding σ-fields $\{F^n_{i,j}: -\infty \leq i \leq j \leq \infty, n \geq 1\}$ and constants $\{d_{ni}: 1 \leq i \leq k_n, n \geq 1\}$ and $\{\nu_m: m \geq 0\}$ such that $\nu_m \downarrow 0$ as $m \to \infty$ and

$$\|X_{ni} - E(X_{ni}|F^n_{i-m,i+m})\|_p \leq d_{ni}\nu_m.$$

By definition, an NED triangular array is an L^2-NED triangular array. By Holder's inequality, every NED array also is an L^p-NED array for $p \in [1,2]$.

Below we show that a mean-zero L^1-NED sequence or array that is L^r bounded for some $r > 1$ is a uniformly integrable L^1-mixingale with $c_{ni} = 2d_{ni} + \|X_{ni}\|_r$ and $\psi_m = \nu_{[m/2]} + 6\alpha([m/2])^{1-1/r}$ for $m \geq 0$, where $[\cdot]$ denotes the integer part of $\cdot$. Hence, $\bar{X}_n$ satisfies L^1 and weak LLNs if $\overline{\lim}_{n\to\infty} \frac{1}{k_n}\sum_{i=1}^{k_n} d_{ni} < \infty$. Furthermore, for $1 \leq p \leq 2$, a mean-zero L^p-NED sequence or array that is L^r bounded for some $r > p$ is a uniformly integrable L^p-mixingale. In this case, $\bar{X}_n$ satisfies L^p and weak LLNs if $\lim_{n\to\infty} \frac{1}{k_n}\sum_{i=1}^{k_n} d_{ni} < \infty$.

Suppose $\{X_{ni}\}$ is a mean-zero L^p-NED triangular array that is L^r-bounded for some $r > p$. To show that $\{X_{ni}\}$ is an L^p-mixingale, we follow the argument of McLeish [10, p. 837]. Let $E_{i-2m}(\cdot)$ denote $E(\cdot|F^n_{-\infty,i-2m})$, then

$$\begin{aligned}
\|E_{i-2m}X_{ni}\|_p &= \|E_{i-2m}(X_{ni} - E(X_{ni}|F^n_{i-m,i+m}) + E(X_{ni}|F^n_{i-m,i+m}))\|_p,\\
&\leq \|X_{ni} - E(X_{ni}|F^n_{i-m,i+m})\|_p + \|E_{i-2m}E(X_{ni}|F^n_{i-m,i+m})\|_p,\\
&\leq d_{ni}\nu_m + 6\alpha(m)^{1/p-1/r}\|E(X_{ni}|F^n_{i-m,i+m})\|_r,\\
&\leq d_{ni}\nu_m + 6\alpha(m)^{1/p-1/r}\|X_{ni}\|_r,\\
&\leq c_{ni}\psi_{2m} \qquad \textbf{(2)}
\end{aligned}$$

for $c_{ni} = 2d_{ni} + \|X_{ni}\|_r$ and $\psi_m = \nu_{[m/2]} + 6\alpha([m/2])^{1/p-1/r}$, where the first inequality above holds by the triangle inequality and the conditional Jensen's inequality, the second inequality holds by the L^p-NED assumption, the mean-zero assumption, and McLeish's [10, Lemma 2.1] $\alpha(\cdot)$ mixing inequality, and the third inequality holds by the conditional Jensen's inequality. By the $\alpha(\cdot)$ mixing property of $\{Y_{ni}\}$ and $\{F^n_{i,j}\}$, $\psi_m \downarrow 0$ as $m \to \infty$.

It remains to show that $\{X_{ni}\}$ satisfies the L^p-mixingale condition (b) (i.e. condition (b) of p. 460 with $\|\cdot\|_1$ replaced by $\|\cdot\|_p$). We have

$$\begin{aligned}\|X_{ni} - E_{i+m}X_{ni}\|_p &\le \|X_{ni} - E(X_{ni}|F^n_{i-m,i+m})\|_p + \|E(X_{ni}|F^n_{i-m,i+m}) \\ &\quad - E_{i+m}X_{ni}\|_p, \\ &= \|X_{ni} - E(X_{ni}|F^n_{i-m,i+m})\|_p + \|E_{i+m}(E(X_{ni}|F^n_{i-m,i+m}) \\ &\quad - X_{ni})\|_p, \\ &\le 2\|X_{ni} - E(X_{ni}|F^n_{i-m,i+m})\|_p, \\ &\le 2d_{ni}\nu_m \le c_{ni}\psi_m, \end{aligned} \tag{3}$$

where the second inequality holds by the conditional Jensen's inequality. Thus, condition (b) holds and $\{X_{ni},F_{ni}\}$ is an L^p-mixingale where $F_{ni} = F^n_{-\infty,i} \forall i$ and $\forall n \ge 1$.

7. By Holder's inequality, mixingales necessarily are L^1-mixingales. Hence, a uniformly integrable mixingale array with mixingale numbers $\{c_{ni}\}$ that satisfy $\lim_{n\to\infty} \frac{1}{k_n}\sum_{i=1}^{k_n} c_{ni} < \infty$ satisfies L^1 and weak LLNs without any assumptions on the decay rate of the mixingale numbers.

4. PROOFS

The proof of Theorem 1 uses an L^p LLN for martingale difference sequences due to Chow [4]. For completeness, we provide a simple proof of this result based on an argument of Hall and Heyde [7, Theorem 2.22].

LEMMA. *Let $\{Y_i,F_i: i \ge 1\}$ be a martingale difference sequence. If $\{|Y_i|^p: i \ge 1\}$ is uniformly integrable for some $1 \le p \le 2$, then*

$$\|\bar{Y}_n\|_p = \left(E\left|\frac{1}{n}\sum_{i=1}^n Y_i\right|^p\right)^{1/p} \to 0 \qquad \text{as } n \to \infty.$$

Proof of Lemma: Given any $\epsilon > 0$, let $B < \infty$ be such that $\sup_{i\ge1} \|Y_i 1(|Y_i| > B)\|_p < \epsilon/4$. Let $W_i = Y_i 1(|Y_i| \le B)$ and $Z_i = Y_i 1(|Y_i| > B)$. Then,

$$\begin{aligned}\|\bar{Y}_n\|_p &\le \left\|\frac{1}{n}\sum_{i=1}^n (W_i - E(W_i|F_{i-1}))\right\|_p + \left\|\frac{1}{n}\sum_{i=1}^n (Z_i - E(Z_i|F_{i-1}))\right\|_p, \\ &\le \left\|\frac{1}{n}\sum_{i=1}^n (W_i - E(W_i|F_{i-1}))\right\|_2 + \frac{2}{n}\sum_{i=1}^n \|Z_i\|_p, \\ &= \left(\frac{1}{n^2}\sum_{i=1}^n E(W_i - E(W_i|F_{i-1}))^2\right)^{1/2} + \frac{2}{n}\sum_{i=1}^n \|Z_i\|_p, \\ &\le \left(\frac{1}{n^2}\sum_{i=1}^n EW_i^2\right)^{1/2} + \epsilon/2, \\ &\le B/\sqrt{n} + \epsilon/2, \end{aligned} \tag{4}$$

where the second inequality uses the conditional Jensen's inequality and the equality uses the fact that $\{W_i - E(W_i|\boldsymbol{F}_{i-1}),\boldsymbol{F}_i: i \geq 1\}$ is a martingale difference sequence. For n sufficiently large, the right-hand side of (4) is less than ϵ. ■

Proof of Theorem 1: First, we prove part (a). Let $Y_{mi} = E(X_i|\boldsymbol{F}_{i+m}) - E(X_i|\boldsymbol{F}_{i+m-1})$ for $i = 1,2,\ldots; m = \ldots,0,1,\ldots$. Then, $\{Y_{mi},\boldsymbol{F}_{i+m}: i \geq 1\}$ is a uniformly integrable martingale difference sequence for each m and the Lemma yields

$$E|\bar{Y}_{mn}| = E\left|\frac{1}{n}\sum_{i=1}^{n} Y_{mi}\right| \to 0 \text{ as } n \to \infty \text{ for } m = \ldots,0,1,\ldots. \tag{5}$$

Next, we write

$$\begin{aligned}\bar{X}_n &= \frac{1}{n}\sum_{i=1}^{n}(X_i - E(X_i|\boldsymbol{F}_{i+M})) + \frac{1}{n}\sum_{i=1}^{n}(E(X_i|\boldsymbol{F}_{i+M}) - E(X_i|\boldsymbol{F}_{i+M-1}))\\ &\quad + \cdots + \frac{1}{n}\sum_{i=1}^{n}(E(X_i|\boldsymbol{F}_{i-M+1}) - E(X_i|\boldsymbol{F}_{i-M})) + \frac{1}{n}\sum_{i=1}^{n}E(X_i|\boldsymbol{F}_{i-M}),\\ &= \frac{1}{n}\sum_{i=1}^{n}(X_i - E(X_i|\boldsymbol{F}_{i+M})) + \sum_{m=-M+1}^{M}\bar{Y}_{mn} + \frac{1}{n}\sum_{i=1}^{n}E(X_i|\boldsymbol{F}_{i-M}),\end{aligned} \tag{6}$$

and so,

$$\begin{aligned}E|\bar{X}_n| &\leq \frac{1}{n}\sum_{i=1}^{n}\|X_i - E(X_i|\boldsymbol{F}_{i+M})\|_1 + \sum_{m=-M+1}^{M}E|\bar{Y}_{mn}|\\ &\quad + \frac{1}{n}\sum_{i=1}^{n}\|E(X_i|\boldsymbol{F}_{i-M})\|_1,\\ &\leq \frac{1}{n}\sum_{i=1}^{n}c_i\psi_{M+1} + \sum_{m=-M+1}^{M}E|\bar{Y}_{mn}| + \frac{1}{n}\sum_{i=1}^{n}c_i\psi_M,\end{aligned} \tag{7}$$

using the assumption that $\{X_i,\boldsymbol{F}_i\}$ is an L^1-mixingale. Given any $\epsilon > 0$, there exists a constant $M < \infty$ such that the first and third summands above are each less than $\epsilon/3$ uniformly over n $\left(\text{since } \overline{\lim}_{n\to\infty}\frac{1}{n}\sum_{i=1}^{n}c_i < \infty\right)$. Given $M < \infty$, there exists a constant $N < \infty$ such that for all $n \geq N$ the second term is less than $\epsilon/3$ by equation (5). Hence, part (a) holds.

Part (b) follows from part (a) because uniformly integrable sequences are L^1 bounded, and hence $\lim_{n\to\infty}\frac{1}{n}\sum_{i=1}^{n}\|X_i\|_1 \leq \sup_{k\geq 1}\|X_k\|_1 < \infty$. ■

Next we prove the assertion of Comment 2 following Theorem 1 which gives conditions for an L^2 LLN. For any integer $s \in [0, i-j]$, we have

$$\begin{aligned}|EX_iX_j| &\le |EX_i(X_j - E(X_j|F_{j+s}))| + |EE(X_iE(X_j|F_{j+s})|F_{j+s})|,\\ &\le \|X_i\|_2 \cdot \|X_j - E(X_j|F_{j+s})\|_2 + |E[E(X_j|F_{j+s})E(X_i|F_{j+s})]|,\\ &\le \sup_{k\ge 1} \|X_k\|_2^2 \cdot \psi_{s+1} + \|E(X_j|F_{j+s})\|_2 \cdot \|E(X_i|F_{j+s})\|_2,\\ &\le \sup_{k\ge 1} \|X_k\|_2^2 \cdot \psi_{s+1} + \sup_{k\ge 1} \|X_k\|_2^2 \cdot \psi_{i-j-s}. \qquad (8)\end{aligned}$$

Using this inequality with $s = [(i-j-1)/2]$ (where $[x]$ denotes the integer part of x), we get

$$\begin{aligned}E|\bar{X}_n|^2 &\le \frac{2}{n^2}\sum_{i=1}^{n}\sum_{j=1}^{i}|EX_iX_j|,\\ &\le \sup_{k\ge 1}\|X_k\|_2^2\frac{4}{n^2}\sum_{i=1}^{n}\sum_{j=1}^{i}\psi_{[(i-j-1)/2]+1},\\ &\le \sup_{k\ge 1}\|X_k\|_2^2\frac{8}{n}\sum_{u=1}^{[n/2]}\psi_u \to 0 \qquad \text{as } n\to\infty. \qquad (9)\end{aligned}$$

■

REFERENCES

1. Andrews, D.W.K. Consistency in nonlinear econometric models: a generic uniform law of large numbers. *Econometrica* 55 (1987): 1465–1472.
2. Andrews, D.W.K. & R C. Fair. Inference in econometric models with structural change. *Review of Economic Studies* 60 (1988): forthcoming.
3. Billingsley, P. *Convergence of probability measures.* New York: Wiley, 1968.
4. Chow, Y.S. On the L^p-convergence for $n^{-1/p}S_n$, $0 < p < 2$. *Annals of Mathematical Statistics* 42 (1971): 393–394.
5. Gallant, A.R. *Nonlinear statistical models.* New York: Wiley, 1987.
6. Gallant, A R. & H White. *A unified theory of estimation and inference for nonlinear dynamic models.* New York: Basil Blackwell, 1988.
7. Hall, P. & C.C. Heyde. *Martingale limit theory and its applications.* New York: Academic Press, 1980.
8. Herrndorf, N. A functional central-limit theorem for ρ-mixing sequences. *Journal of Multivariate Analysis* 15 (1984): 141–146.
9. Kolmogorov, A.N. & Yu. A Rozonov On strong mixing conditions for stationary Gaussian processes. *Theory of Probability and Its Applications* 5 (1960): 204–207.
10. McLeish, D.L. A maximal inequality and dependent strong laws. *Annals of Probability* 3 (1975): 829–839.
11. McLeish, D.L. Invariance principles for dependent variables. *Z. Wahrscheinlichkeitstheorie verw. Gebiete* 32 (1975): 165–178.
12. McLeish, D.L. On the invariance principle for nonstationary mixingales. *Annals of Probability* 5 (1977). 616–621
13. Teicher, H Almost certain convergence in double arrays. *Z. Wahrscheinlichkeitstheorie verw. Gebiete* 69 (1985): 331–345.
14. Wooldridge, J. & H. White. Some invariance principles and central-limit theorems for dependent heterogeneous processes. *Econometric Theory* 4 (1988): forthcoming.