Econometric Theory, 4, 1988, 210–230. Printed in the United States of America.

SOME INVARIANCE PRINCIPLES AND CENTRAL LIMIT THEOREMS FOR DEPENDENT HETEROGENEOUS PROCESSES

JEFFREY M. WOOLDRIDGE AND HALBERT WHITE
Massachusetts Institute of Technology
and University of California, San Diego

Building on work of McLeish [14,15], we present a number of invariance principles for doubly indexed arrays of stochastic processes which may exhibit considerable dependence, heterogeneity, and/or trending moments. In particular, we consider possibly time-varying functions of infinite histories of heterogeneous mixing processes and obtain general invariance results, with central limit theorems following as corollaries. These results are formulated so as to apply to economic time series, which may exhibit some or all of the features allowed in our theorems. Results are given for the case of both scalar and vector stochastic processes. Using an approach recently pioneered by Phillips [19–21], and Phillips and Durlauf [23], we apply our results to least squares estimation of unit root models.

1. INTRODUCTION

One of the most useful tools for studying the asymptotic distribution of econometric estimators is the central limit theorem (CLT), which provides conditions ensuring that the standardized sum of a sequence of random variables has the standard normal distribution approximately, in large samples. More recently, the functional central limit theorem (FCLT), or invariance principle, has found application in studying the asymptotic distribution of econometric estimators applied to data generated by unit root processes (Phillips [19–21], Phillips and Durlauf [23], Park and Phillips [16,17]). The FCLT generalizes the CLT to metric spaces other than finite-dimensional Euclidean spaces; a common application is to provide conditions ensuring the convergence of the partial sums of a sequence of random variables to Brownian motion.

In this paper, we present FCLT's and CLT's closely related to earlier results of McLeish [14,15] which apply to the partial sums of a class of dependent heterogeneous stochastic processes. McLeish [14, Theorem 4.2] proves an FCLT for singly indexed sequences of random variables which are near epoch dependent functions of ϕ- or α-mixing processes under conditions

The authors gratefully acknowledge the helpful comments of Clive Granger, A.R. Gallant, Lars Hansen, Alberto Holly, Peter Phillips, and the referees. Any errors are the authors' responsibility. This work was supported by NSF Grant SES-8510637 and by a Sloan Foundation Dissertation Fellowship.

  0266-4666/88 $5.00 + .00

which rule out trending moments. McLeish [15, Corollary 2.11] gives an FCLT for doubly indexed sequences of ϕ- or α-mixing processes under conditions allowing for trending moments. We use McLeish's approach to fill a gap in the results so far available. Specifically, we give an FCLT for doubly indexed sequences of near epoch dependent functions of ϕ- or α-mixing processes under conditions which permit trending moments. Our results are not particularly technically innovative; however, the relatively primitive and general conditions under which our results are obtained may make them readily applicable to a variety of economic contexts.

The plan of the paper is as follows. In Section 2 we present a general FCLT and CLT for a particular class of univariate mixingales (McLeish [13]) with possibly trending moments. Section 3 contains simpler but still fairly general results under conditions ruling out trends. Section 4 contains results for multivariate time series. In each section, we provide examples involving least squares estimation of AR(1) models on data generated by unit root processes which somewhat extend work of Phillips [20] and Phillips and Durlauf [23]. Section 5 contains a summary and concluding remarks.

2. AN INVARIANCE PRINCIPLE FOR DEPENDENT HETEROGENEOUS ARRAYS

Our results obtain as corollaries to a general invariance principle (Theorem 19.2) of Billingsley [5]. (See Billingsley [5,6] for definitions and useful background.) We refer to $C \equiv C[0,1]$, the space of all continuous functions on $[0,1]$ and to $D \equiv D[0,1]$, the space of all right continuous function with left limits on $[0,1]$. For convenience and ease of reference, we give a slight restatement of Billingsley's result.

THEOREM 2.1. *Let $\{W_n: n = 1,2,\dots\}$ be a sequence of random elements of D such that*

(i) *$\{W_n\}$ is tight in the Skorohod topology on D, and the weak limit process of any convergent subsequence is in C with probability one;*
(ii) *$\{W_n^2(a): n = 1,2,\dots\}$ is uniformly integrable for all $a \in [0,1]$;*
(iii) *$E[W_n(a)] \to 0$ and $E[W_n^2(a)] \to a$ as $n \to \infty$ for all $a \in [0,1]$;*
(iv) *$\{W_n\}$ has asymptotically independent increments.*

Then

$$W_n \xrightarrow{L} W,$$

where W is a standard Brownian motion on D. Further,

$$W_n(1) \xrightarrow{L} N(0,1).$$

The conditions of Theorem 2.1 are fairly abstract. The following result (McLeish [15, Lemma 3.6]) provides some useful sufficient conditions for Theorem 2.1, parts (i) and (ii).

LEMMA 2.2. *Let* $\{W_n\}$ *be a sequence of random elements of D such that* $W_n(0) = 0$ *and*

$$U \equiv \{ \max_{a \leq b \leq a+\delta} [W_n(b) - W_n(a)]^2/\delta : n \geq N(a,\delta),\ 0 \leq a \leq 1,\ 0 \leq \delta \leq 1\}$$

is uniformly integrable for some nonrandom function $N(a,\delta) \in \mathbb{N} \equiv \{1,2,\cdots\}$. *(If* $a + \delta \geq 1$ *then the max is taken over* $[a,1]$.)

Then $\{W_n\}$ *is tight in the Skorohod topology on D, and the limit in law of any convergent subsequence of* $\{W_n\}$ *is almost surely in C. Furthermore, for each* $a \in [0,1]$,

$$\{W_n^2(a): n = 1,2,\ldots\}$$

is uniformly integrable.

The particular choice for W_n which is of interest here is given in the following assumption.

Assumption A.1. $\{X_{nt}: n,t = 1,2,\ldots\}$ is a double array of real-valued random variables on the probability space (Ω,F,P), and for $a \in [0,1]$,

$$W_n(a) \equiv \sum_{t=1}^{k_n(a)} X_{nt},$$

where $k_n(\cdot)$ is a nondecreasing right continuous integer-valued function on $[0,1]$, with $k_n(0) = 0$, $n = 1,2,\ldots$.

A leading case occurs when $X_{nt} = \sigma_n^{-1}Z_t$, $\sigma_n^2 \equiv \operatorname{var}\left(\sum_{t=1}^{n} Z_t\right)$ and $k_n(a) = [na]$, where $[na]$ denotes the integer part of na. We place conditions on the dependence and moments of $\{X_{nt}\}$ which ensure that the hypotheses of Lemma 2.2 hold. Specifically, we require that $\{X_{nt}\}$ be a mixingale sequence, a concept introduced by McLeish [13]. For convenience, we provide the following definition.

DEFINITION 2.3. *Let* $\{X_{nt}: n,t = 1,2,\ldots\}$ *and* $\{F_{nt}: n = 1,2,\ldots; t = \ldots -1,0,1,\ldots\}$ *be double arrays such that* F_{nt} *is a sub* σ*-algebra of* F *and* $F_{nt} \subset F_{n,t+1}$, *for all* $n \in \mathbb{N}$, $t \in Z \equiv \{\cdots -1,0,1,\cdots\}$. *Let* $\{\psi_m: m = 0,1,\ldots\}$ *be a sequence of nonnegative constants with* $\psi_m \downarrow 0$ *as* $m \to \infty$ *and* $\{c_{nt} > 0: n,t = 1,2,\ldots\}$ *be a double array of positive constants. Letting* $\|\cdot\|_2$ *denote the* $L_2(P)$ *norm, assume that for all* $n,t \geq 1$, $m \geq 0$,

(i) $\|E[X_{nt}|F_{n,t-m}]\|_2 \leq \psi_m c_{nt}$,
(ii) $\|X_{nt} - E[X_{nt}|F_{n,t+m}]\|_2 \leq \psi_{m+1} c_{nt}$.

Then $\{X_{nt}\}$ *is a mixingale with respect to* $\{F_{nt}\}$.

When $\psi_m = O(m^{-\lambda})$ for all $\lambda > q$, we say that $\{\psi_m\}$ is of size $-q$.

This definition implies $E(X_{nt}) = 0$ for all n,t. Special cases of mixingales are mixing sequences and martingale difference sequences ($\psi_m = 0$ for all

$m \neq 0$). Mixingales allow for substantial dependence and heterogeneity. Our main result, obtained using McLeish's [15] general approach, is an invariance principle for a particular class of mixingales. For now it suffices to impose the following conditions.

Assumption A.2. $\{X_{nt}, F_{nt}\}$ is a mixingale with $\{\psi_m\}$ of size $-\frac{1}{2}$ and constants $\{c_{nt}\}$.

Assumption A.3. $\{X_{nt}^2/c_{nt}^2\colon n,t = 1,2,\ldots\}$ is uniformly integrable.

We define the partial sum

$$S_{nj} \equiv \sum_{t=1}^{j} X_{nt}, \qquad n,j = 1,2,\ldots.$$

The following inequality for the partial sum due to McLeish [15, Lemma 3.2] is fundamental for the analysis of mixingales; among other things, it allows verification of the conditions of Lemma 2.2. For a detailed proof, see Gallant [8, Lemma 2, p. 511].

LEMMA 2.4. *Let $\{X_{nt}, F_{nt}\}$ be a mixingale and put $S_{n0} \equiv 0$, $S_{nj} \equiv \sum_{t=1}^{j} X_{nt}, j \geq 1$. Suppose $\{a_k\colon k \in Z\}$ is a doubly infinite sequence of positive constants with $a_k = a_{-k}$. Then for all $k \geq 0$, $l,n \geq 1$,*

$$E[\max_{j\leq l} (S_{n,k+j} - S_{nk})^2] \leq 4\left(\sum_{t=k+1}^{k+l} c_{nt}^2\right)\left(\sum_{i=-\infty}^{\infty} a_i\right)\left[(\psi_0^2 + \psi_1^2)/a_0 + 2\sum_{t=1}^{\infty} \psi_t^2 |a_t^{-1} - a_{t-1}^{-1}|\right].$$

If $\psi_m > 0$ for all $m \geq 0$, then

$$E[\max_{j\leq l} (S_{n,k+j} - S_{nk})^2] \leq 16 \sum_{t=k+1}^{k+l} c_{nt}^2 \left\{\sum_{m=0}^{\infty} \left(\sum_{i=0}^{m} \psi_i^{-2}\right)^{-1/2}\right\}^2$$

for all $k \geq 0$, $l,n \geq 1$.

Applying Lemma 2.4 gives the following result (McLeish [15, Lemma 3.5]).

Proposition 2.5. Suppose that Assumptions A.1–A.3 hold. Then

$$\left\{\max_{j\leq l} (S_{n,k+j} - S_{nk})^2 \Big/ \sum_{t=k+1}^{k+l} c_{nt}^2,\ k = 0,1,\ldots;\ l,n = 1,2,\ldots\right\}$$

is uniformly integrable.

The following assumption (identical to Assumption 2.2a of McLeish [15]), along with Proposition 2.5, allows us to verify the uniform integrability requirement in Lemma 2.2.

Assumption A.4.

$$\sup_{\substack{0<\delta\leq 1-a \\ 0\leq a<1}} \limsup_{n\to\infty} \delta^{-1} \sum_{t=k_n(a)}^{k_n(a+\delta)} c_{nt}^2 < \infty.$$

Although this assumption appears somewhat complicated, it is typically straightforward to verify. In cases where $X_{nt} = \sigma_n^{-1} Z_t$ and growing moments in Z_t can be ruled out, simpler conditions are available. These are considered in Section 3. The present assumption permits treatment of cases with growing moments.

The following result establishes conditions under which Lemma 2.2 applies to the choice for W_n of Assumption A.1.

Proposition 2.6. Suppose that Assumptions A.1–A.4 hold. Then there exists a nonrandom finite integer function $N(a,\delta)$ such that

$$\{ \max_{a\leq b\leq a+\delta} [W_n(b) - W_n(a)]^2/\delta\colon n \geq N(a,\delta),\ 0 \leq a \leq 1,\ 0 < \delta \leq 1\}$$

is uniformly integrable.

Applying Lemma 2.2 immediately yields

COROLLARY 2.7. *Given Assumptions A.1, A.2, A.3, and A.4,* $\{W_n\}$ *is tight in D, and any limit process is almost surely continuous. Furthermore, for each* $a \in [0,1]$, $\{W_n^2(a)\colon n = 1,2,\dots\}$ *is uniformly integrable.*

Thus, Assumptions A.1–A.4 suffice for Theorem 2.1, parts (i) and (ii). Because $\{X_{nt}\}$ is a mixingale, it follows that $E(W_n(a)) = 0$ for all $a \in [0,1]$. To ensure that part (iii) of Theorem 2.1 holds, it suffices simply to impose

Assumption A.5. For each $a \in [0,1]$, $E(W_n^2(a)) \to a$ as $n \to \infty$.

The flexibility allowed by choice of k_n permits treatment of cases with $X_{nt} = \sigma_n^{-1} Z_t$ and heterogeneous or trending moments in Z_t under Assumption A.5. However, certain non-ergodic processes (Hall and Heyde [12]) are ruled out. Including these would require letting k_n be a random element of D, as in Durrett and Resnick [7]. This is beyond the scope of this paper.

The conclusions of Theorem 2.1 now follow given Assumptions A.1–A.5 when $W_n(a)$ has asymptotically independent increments. Mixingales do not possess enough structure to ensure this property. Although it is possible to obtain an FCLT for mixingales without imposing asymptotically independent increments (McLeish [15], Theorem 2.4), the condition needed in place of this property is not primitive (McLeish's [15] condition 2.6). Therefore, following the earlier work of McLeish [14], we specialize to a subclass of mixingales, near epoch-dependent functions of mixing processes, which do have the asymptotically independent increments property. First, we define the concept of near-epoch dependence.

DEFINITION 2.8. *Let* $\{Y_{nt}: t = 0, \pm 1, \ldots, n = 1,2, \ldots\}$ *be a doubly infinite double array of random vectors defined on* (Ω, F, P). *Let* $G_{n,j}^{k} \equiv \sigma(Y_{nj}, \ldots, Y_{nk})$, *for all* $j \leq k$, $j,k \in Z$, $n \in \mathbb{N}$. *The process* $\{X_{nt}\}$ *is near-epoch dependent with respect to* $\{Y_{nt}\}$ *if and only if* $X_{nt} \in L_2(P)$ *for all* $n,t \in \mathbb{N}$ *and there exist constants* $\{\nu_m \geq 0: m = 0,1, \ldots\}$ *with* $\nu_m \downarrow 0$ *as* $m \to \infty$ *and* $\{d_{nt} > 0: n,t = 1,2, \ldots\}$ *such that*

$$\|X_{nt} - E[X_{nt} | G_{n,t-m}^{t+m}]\|_2 \leq \nu_m d_{nt}.$$

This condition ensures that X_{nt} may be a function of the entire history of Y_{nt}, but depends primarily on the "near epoch" of Y_{nt}.

This notion of near-epoch dependence extends a concept introduced by Billingsley [5] and somewhat generalizes definitions appearing elsewhere. For example, in Bierens [4] (who refers to this property as "ν-stability"), McLeish [13,14], and Gallant and White [9], $d_{nt} \equiv 1$. When $X_{nt} = \sigma_n^{-1} Z_t$ and Z_t has growing moments, it is useful to allow d_{nt} to grow with t and shrink with n.

We define the uniform ($\phi-$) and strong ($\alpha-$) mixing coefficients for $\{Y_{nt}\}$ as

$$\phi_m \equiv \sup_n \sup_t \sup_{\{G \in G_{n,-\infty}^{t}, H \in G_{n,t+m}^{\infty}: P(G) > 0\}} |P(H|G) - P(H)|,$$

$$\alpha_m \equiv \sup_n \sup_t \sup_{\{G \in G_{n,-\infty}^{t}, H \in G_{n,t+m}^{\infty}\}} |P(G \cap H) - P(G)P(H)|.$$

The next result, proved along the lines of Theorem 3.1 of McLeish [13], establishes conditions ensuring that near-epoch dependent functions of mixing processes are mixingales.

Proposition 2.9. Let $\{Y_{nt}: t = 0, \pm 1, \ldots, n = 1,2, \ldots\}$ and $\{X_{nt}: n,t = 1,2, \ldots\}$ be random double arrays on (Ω, F, P). Suppose

(i) for some $r \geq 2$, $\|X_{nt}\|_r < \infty$ and $E(X_{nt}) = 0$, for all $n,t \in \mathbb{N}$,
(ii) $\{X_{nt}\}$ is near-epoch dependent with respect to $\{Y_{nt}\}$ with $\{\nu_m\}$ of size $-\frac{1}{2}$
(iii) $\{Y_{nt}\}$ is mixing with either $\{\phi_m\}$ of size $-r/2(r-1)$ or $\{\alpha_m\}$ of size $-r/(r-2)$, $r > 2$.

Let $F_{nt} \equiv G_{n,-\infty}^{t} = \sigma(\ldots, Y_{nt})$ and $c_{nt} \geq \max(\|X_{nt}\|_r, d_{nt})$. Then $\{X_{nt}, F_{nt}\}$ *is a mixingale with* $\{\psi_m \equiv 2\phi_{[m/2]}^{1-1/r} + \nu_{[m/2]}\}$ or $\{\psi_m \equiv 5\alpha_{[m/2]}^{1/2-1/r} + \nu_{[m/2]}\}$ of size $-\frac{1}{2}$ and constants $\{c_{nt}\}$.

Further, if (i) and (ii) hold with $r > 2$, then $\{X_{nt}^2/c_{nt}^2: n,t = 1,2, \ldots\}$ is uniformly integrable.

This result allows Assumption A.2 to be replaced with the conditions of Proposition 2.9. By imposing conditions on d_{nt} which can be traded off with conditions on ν_m, the asymptotically independent increments condition for $\{W_n\}$ can be established. We replace Assumption A.2 with

Assumption B.2.

(i) For some $r \geq 2$, $\|X_{nt}\|_r < \infty$, and $E(X_{nt}) = 0$, $n, t = 1, 2, \ldots$,
(ii) $\{X_{nt}\}$ is near-epoch dependent with respect to $\{Y_{nt}\}$ with $\{\nu_m\}$ of size $-\gamma, \frac{1}{2} \leq \gamma < 1$, and $\max_{1 \leq t \leq k_n(1)} d_{nt} = O(k_n(1)^{\gamma - 1})$,
(iii) $\{Y_{nt}\}$ is mixing with either $\{\phi_m\}$ of size $-r/2(r-1)$ or $\{\alpha_m\}$ of size $-r/(r-2)$, $r > 2$,
(iv) $c_{nt} \geq \max(\|X_{nt}\|_r, d_{nt})$, $n, t = 1, 2, \ldots$.

LEMMA 2.10. *Given Assumptions A.1 with* $k_n(a) - k_n(b) \to \infty$ *as* $n \to \infty$ *for all* $0 \leq b < a \leq 1$, *B.2, A.3, and A.4,* $\{W_n\}$ *has asymptotically independent increments.*

Note that in this result (and in the following), the constants c_{nt} used in Assumptions A.3 and A.4 are those of Assumption B.2(iv) and not of A.2. The invariance principle can now be given.

THEOREM 2.11. *Given Assumptions A.1 with* $k_n(a) - k_n(b) \to \infty$ *as* $n \to \infty$ *for all* $0 \leq b < a \leq 1$, *B.2, A.3, and A.4,* $\{W_n\}$ *is tight in the Skorohod topology on D, and any weak limit process is almost surely continuous. If Assumption A.5 also holds, then* $W_n \xrightarrow{L} W$, *a standard Brownian motion on D. Further,* $\sum_{t=1}^{k_n(1)} X_{nt} \xrightarrow{L} N(0,1)$.

This result extends Theorem 4.2 of McLeish [14] and Corollary 2.11 of McLeish [15]. McLeish's [14] Theorem 4.2 applies to singly indexed sequences with uniformly bounded moments. McLeish's [15] Corollary 2.11 allows doubly indexed arrays and processes with trending moments, but restricts attention to ϕ- and α-mixing processes. Thus, Theorem 2.11 synthesizes the two earlier results of McLeish in a way that allows straightforward application in economic contexts. McLeish's [15] Theorem 2.4 is potentially applicable in contexts broader than our Theorem 2.11, but the condition used in place of the asymptotically independent increments property (condition 2.6) is far from being primitive. Although McLeish [15] derives Corollary 2.11 for ϕ- and α-mixing processes from his Theorem 2.4, we have not been able to show that the more general class of near-epoch dependent processes considered here satisfies McLeish's condition 2.6.

Observe that Assumption A.3 is automatically satisfied when $r > 2$ by Proposition 2.9. However, by not imposing $r > 2$, we allow for cases in which conditions are not placed on moments beyond the second, as in the Lindeberg–Lévy CLT.

As an example, we obtain the asymptotic distribution for the least squares estimator of an AR(1) model when the underlying process is generated as a random walk, the innovations of which have growing variances and considerable time dependence. The example extends a result of Phillips [20, Theorem 3.1] who assumes strong mixing disturbances with uniformly bounded variances.

Example 2.12. Suppose that the univariate process $\{y_t\}$ is generated according to

$$y_t = y_{t-1} + u_t,$$

$$u_t = t^\rho \epsilon_t, \qquad t = 1,2,\dots$$

with y_0 having a given distribution and $\rho \geq 0$. Suppose further that

i. For some $r \geq 2$, $\sup_t \|\epsilon_t\|_{2r} < \infty$, and $E(\epsilon_t) = 0$, $t = 1,2,\dots,$
ii. (a) $\{\epsilon_t\colon t = 1,2,\dots\}$ is near-epoch dependent with respect to a process $\{\eta_t\colon t = 0, \pm 1,\dots\}$ with ν_m of size $-\frac{1}{2}$ and $d_t = 1$; (b) $\{\epsilon_t^2\colon t = 1,2,\dots\}$ is near-epoch dependent with respect to $\{\eta_t\colon t = 0, \pm 1,\dots\}$ with ν_m of size $-\frac{1}{2}$ and $d_t = 1$,
iii. $\{\eta_t\}$ is mixing with $\{\phi_m\}$ of size $-r/(2r-1)$ or $\{\alpha_m\}$ of size $-r(r-2)$, $r > 2$,
iv. (a) $n^{-(2\rho+1)}\sum_{t=1}^{n} E(u_t^2) \to \sigma_u^2$ as $n \to \infty$; (b) $E\left[\left(\sum_{t=1}^{n} u_t\right)^2\right] = \sigma_o^2 n^{2\rho+1} + o(n^{2\rho+1})$, where $\sigma_o^2 > 0$.

Define $\hat{\beta}_n \equiv \left(\sum_{t=1}^{n} y_{t-1}^2\right)^{-1} \sum_{t=1}^{n} y_{t-1} y_t$ and $Z(a) \equiv W(a^{2\rho+1})$. Then

$$n(\hat{\beta}_n - 1) \xrightarrow{L} \left[\int_0^1 Z^2(a)da\right]^{-1} (\tfrac{1}{2}(Z^2(1) - \sigma_u^2/\sigma_o^2).$$

Further,

$$2\sigma_o^{-2} n^{-(2\rho+1)}\left(\sum_{t=1}^{n} y_{t-1}^2\right)(\hat{\beta}_n - 1) + \sigma_u^2/\sigma_o^2 \xrightarrow{L} \chi_1^2.$$

One novel feature of this result is that the growth of var u_t (which necessitates the choice $k_n(a) = [na^{-1/(2\rho+1)}]$) leads to the presence of $Z^2(a)$ in the stochastic integral, rather than $W^2(a)$ (cf. Phillips [20, Theorem 3.1] and Example 3.3 which follows), so that the limiting distribution of $\hat{\beta}_n$ is sensitive to growth in the disturbance variance. Andrews [3] also derives consistency and limiting distribution results for a class of models with I(1) regressors and strong mixing errors which may have exploding variances. Example 2.12 falls outside his framework because he assumes that the disturbances driving the regressors have uniformly bounded variances. Because the regressor in Example 2.12 is the lag of y_t, Andrews' assumptions do not hold.

The near-epoch dependence of condition (ii) of Example 2.12 is a rather weak dependence condition. As we see in the next section, it allows for infinite moving averages in η_t with geometrically declining weights. For clarity, conditions on ϵ_t and ϵ_t^2 have been imposed separately. A sufficient condition for both is that $\{\epsilon_t\}$ is near-epoch dependent with respect to $\{\eta_t\}$ with ν_m of size $-(r-1)/(r-2)$ (Gallant and White [9, Corollary 4.3(b)]).

Although the distribution of $\hat{\beta}_n$ is of interest in its own right, it does not provide a convenient basis for testing for a unit root (H_0: $\beta_o = 1$) due to the presence of the stochastic integral. A convenient test can be based on the second result (see Phillips and Durlauf [23, Lemma 3.1(d)]), with σ_u^2 and σ_o^2 replaced by consistent estimators, as in Phillips [20]. Both tails of the χ_1^2 distribution are relevant for testing for unit roots. Small or negative values indicate the presence of a stable root, while large values indicate the presence of an explosive root.

We point out that the χ_1^2 statistic can be obtained without resorting to an invariance principle. An appropriate central limit theorem will suffice.

Another application of Theorem 2.11, which we mention in passing, is the provision of conditions under which the processes generated by certain of the models of Granger [10] can be shown to converge to standard Brownian motions.

3. SOME SPECIAL CASES

Simpler conditions sufficient for those of Theorem 2.11 may often suffice for econometric practice. Our next results apply in situations in which $X_{nt} = \sigma_n^{-1} Z_{nt}$, and Z_{nt} is not allowed to have exploding moments. The relevant conditions are the following.

Assumption C.1. $\{Z_{nt}: n,t = 1,2,\dots\}$ is a double array of real-valued random variables on (Ω, F, P), $\sigma_n^2 \equiv \text{var}\left(\sum_{t=1}^{n} Z_{nt}\right)$ is such that $\{\sigma_n^{-2}\}$ is $O(n^{-1})$, and for $a \in [0,1]$

$$W_n(a) \equiv \sum_{t=1}^{[na]} X_{nt},$$

where $X_{nt} = \sigma_n^{-1} Z_{nt}$.

Assumption C.2.

(i) For some $r > 2$, $\|Z_{nt}\|_r < \Delta < \infty$, $E(Z_{nt}) = 0$, $n,t = 1,2,\dots,$
(ii) $\{Z_{nt}\}$ is near-epoch dependent with respect to $\{Y_{nt}\}$ with ν_m of size $-\frac{1}{2}$ and $d_{nt} \equiv 1$,
(iii) $\{Y_{nt}\}$ is mixing with either $\{\phi_m\}$ of size $-r/2(r-1)$ or $\{\alpha_m\}$ of size $-r/(r-2)$,
(iv) $c_{nt} \equiv \sigma_n^{-1} \max(1, \|Z_{nt}\|_r)$.

The condition that $\{\sigma_n^{-2}\}$ be $O(n^{-1})$ in Assumption C.1 ensures the nondegeneracy of the limiting distribution of $\{W_n\}$. The Liapounov-like condition of Assumption C.2(i) rules out cases with growing moments and permits the choice $k_n(a) = [na]$ explicit in Assumption C.1.

COROLLARY 3.1. *Given Assumptions C.1 and C.2, $\{W_n\}$ is tight in D and any weak limit process is almost surely continuous. If Assumption A.5 also holds, then $W_n \xrightarrow{L} W$. Moreover, $\sigma_n^{-1}\sum_{t=1}^{n} Z_{nt} \xrightarrow{L} N(0,1)$.*

With $k_n(a) = [na]$, Assumption A.5 imposes an asymptotic covariance stationarity condition. The present condition for the CLT is nevertheless weaker than that of Serfling [24], McLeish [14, Theorem 4.2] or White and Domowitz [26]. For a CLT which imposes no such requirement and also allows for growing moments, see Wooldridge [27].

When (and only when) the n index on Z_{nt} and Y_{nt} is unnecessary (as when one considers correctly specified models of seasonally unadjusted processes without Pitman drift), a simple sufficient condition for Assumption A.5 is that $\sigma_n^2/n \to \sigma_o^2 > 0$. We have

COROLLARY 3.2. *Given Assumptions C.1 and C.2, if $Z_{nt} = Z_{1t}$, $n,t = 1,2,\ldots$, and $Y_{nt} = Y_{1t}$, $t = 0, \pm 1, \ldots$, $n = 1,2,\ldots$, then $\{W_n\}$ is tight in D and any weak limit process is almost surely continuous. If $\sigma_n^2/n \to \sigma_o^2 > 0$ as $n \to \infty$, then $W_n \xrightarrow{L} W$. Further, $\sigma_n^{-1}\sum_{t=1}^{n} Z_t \xrightarrow{L} N(0,1)$.*

This result is a special case of Theorem 4.2 of McLeish [14]; it simplifies McLeish's conditions. It considerably generalizes the CLT (Theorem 2.4) of White and Domowitz [26].

As an example, we extend a result of Phillips [20, Theorem 3.1] to obtain the limiting distribution of the least squares estimator for an AR(1) model when the observations are generated by a random walk, the disturbances of which are an infinite moving average of independent but not identically distributed random variables. Phillips [20] assumes that the disturbances are strong mixing, which potentially rules out application to the infinite moving average case.

Example 3.3. Suppose that the univariate process $\{y_t\}$ is generated according to

$$y_t = y_{t-1} + u_t,$$

$$u_t = \sum_{\tau=0}^{\infty} a_\tau \epsilon_{t-\tau}, \qquad t = 1,2,\ldots,$$

with y_0 having a given distribution. Suppose further that

i. $\sup_t \|\epsilon_t\|_4 < \infty$, and $E(\epsilon_t) = 0$, $t = 0, \pm 1, \ldots$,
ii. $\{a_\tau\}$ is of size $-3/2$,
iii. $\{\epsilon_t\colon t = 0, \pm 1, \ldots\}$ is an independent sequence of random variables,
iv. (a) $n^{-1}\sum_{t=1}^{n} E(u_t^2) \to \sigma_u^2 > 0$ as $n \to \infty$; (b) $n^{-1}E\left[\left(\sum_{t=1}^{n} u_t\right)^2\right] \to \sigma_o^2 > 0$ as $n \to \infty$.

Define $\hat{\beta}_n \equiv \left(\sum_{t=1}^{n} y_{t-1}^2\right)^{-1} \sum_{t=1}^{n} y_{t-1} y_t$. Then

$$n(\hat{\beta}_n - 1) \xrightarrow{L} \left[\int_0^1 W^2(a)da\right]^{-1} (1/2)(W^2(1) - \sigma_u^2/\sigma_o^2).$$

Further,

$$2\sigma_o^{-2} n^{-1} \left(\sum_{t=1}^{n} y_{t-1}^2\right) (\hat{\beta}_n - 1) + \sigma_u^2/\sigma_o^2 \xrightarrow{L} \chi_1^2.$$

The novelty of this result is that it handles disturbances which are infinite moving averages of random variables which need not be identically distributed. Further, the moving average coefficients are required to decay at a geometric rate, rather than at the exponential rate which would be implied by an ARMA(p,q) process with finite p and q. Thus, the present condition allows for processes with a fairly long memory. For example, the fractionally differenced ARIMA (p,d,q) processes (Granger and Joyeaux [11]) are allowed, provided $d < -\frac{1}{2}$. The present example also covers cases which cannot be handled by assuming that $\{u_t\}$ is itself a mixing process; this is useful, as the mixing assumption can rule out certain AR(1) processes (Andrews [1,2]).

Again, a convenient test for the unit root can be based on the χ_1^2 statistic, with σ_u^2 and σ_o^2 replaced by consistent estimators. Consistent estimators for the present context are available from results of Gallant and White [9, Chapter 6]. This statistic will have the same form as that of Example 2.12, illustrating the invariance of this test statistic to the possibility of trending moments in u_t.

4. A MULTIVARIATE INVARIANCE PRINCIPLE

In many applications it is useful to consider the limiting distribution of random elements whose paths are in a multidimensional Euclidean space. Here we consider processes $\{W_n: \Omega \times [0,1] \to \mathbb{R}^l\}$; i.e., for each $\omega \in \Omega$, $W_n(\omega)$: $[0,1] \to \mathbb{R}^l$. Thus, we may view $W_n(\omega)$ as an element of the space $D^l \equiv D \times \cdots \times D$. Because for any X: $[0,1] \to \mathbb{R}^l, X$ is r.c.l.l. (right continuous with left limits) if and only if X_i: $[0,1] \to \mathbb{R}$ is r.c.l.l. for each $i = 1, \ldots, l$ (here $X(a) \equiv (X_1(a), \ldots, X_l(a))$, $a \in [0,1]$), it follows that the space D^l is indeed the appropriate one for considering these random elements. We equip D^l with the metric

$$d(x,y) = \sum_{i=1}^{l} \rho(x_i, y_i),$$

where ρ is the Skorohod metric on D.

The following result, derived from Lemmas 1, 2, and 3 in Phillips and Durlauf [23], establishes that an analog of the Cramér–Wold device holds for random elements of D^l, so that attention can be restricted to linear combinations of random elements of D^l.

Proposition 4.1. Let $\{W_n: n \in \mathbb{N}\}$ be a sequence of random elements of D^l, and let W be a random element of D^l (not necessarily Brownian motion). Then $W_n \xrightarrow{L} W$ if and only if

$$\sum_{i=1}^{l} \lambda_i W_{ni} \xrightarrow{L} \sum_{i=1}^{l} \lambda_i W_i$$

for each linear combination $\lambda' = (\lambda_1, \ldots, \lambda_l)$ with $\lambda'\lambda = 1$.

Combining this result and Theorem 2.11 gives a multivariate analog of Corollary 3.1. We use the following assumptions.

Assumption D.1. $\{Z_{nt}: n,t = 1,2,\ldots\}$ is a double array of real-valued random $l \times 1$ vectors on (Ω, F, P), $V_n \equiv \text{var}\left(\sum_{t=1}^{n} Z_{nt}\right)$ is such that $\{\text{diag}\,[\xi_{n1}^{-1}, \ldots, \xi_{nl}^{-1}] \equiv \Xi_n^{-1}\}$ is $O(n^{-1})$, where Ξ_n is the $l \times l$ diagonal matrix with the eigenvalues $(\xi_{n1}, \ldots, \xi_{nl})$ of V_n along the diagonal, and for $a \in [0,1]$,

$$W_n(a) = \sum_{t=1}^{[na]} X_{nt},$$

where $X_{nt} \equiv V_n^{-1/2} Z_{nt}$, $V_n^{-1/2} \equiv \Xi_n^{-1/2} C_n$ with C_n the orthogonal matrix of eigenvectors of V_n.

Assumption D.2.

(i) For some $r > 2$, $\|Z_{nti}\|_r < \Delta < \infty$, $E(Z_{nti}) = 0$, $n,t = 1,2,\ldots$, $i = 1,\ldots,l$,
(ii) $\{Z_{nti}\}$ is near-epoch dependent with respect to $\{Y_{nt}\}$ with ν_{mi} of size $-\frac{1}{2}$ and $d_{nti} \equiv 1$, $i = 1,\ldots,l$,
(iii) $\{Y_{nt}\}$ is mixing with either $\{\phi_m\}$ of size $-r/2(r-1)$ or $\{\alpha_m\}$ of size $-r/(r-2)$,
(iv) $c_{ntj} = l \cdot \xi_{nj}^{-1/2} \max(1, \|Z_{ntj}\|_r)$, $n,t = 1,2,\ldots$, $j = 1,\ldots,l$.

We also impose a multivariate analog of Assumption A.5.

Assumption D.3. For each $a \in [0,1]$, $E[W_n(a)W_n(a)'] \to aI_l$ as $n \to \infty$.

The multivariate invariance result is the following.

COROLLARY 4.2. *Given Assumptions D.1 and D.2, then $\{W_n\}$ is tight in D^l, and any weak limit process is almost surely continuous. If Assumption D.3 also holds, then $W_n \xrightarrow{L} W$, multivariate Brownian motion. Moreover $V_n^{-1/2} \sum_{t=1}^{n} Z_{nt} \xrightarrow{L} N(0, I_l)$.*

The conditions of this theorem, particularly Assumption D.2, rule out processes with growing moments. We restrict attention to this case because the current framework is not broad enough to allow multiple time series with elements having moments which trend at different rates. In such contexts, a theory of convergence to Gaussian processes more general than Brownian motion is needed. Such a theory can be developed using Proposition 4.1 and the multivariate CLT in Wooldridge [27, Corollary 5.3], which does allow for moments growing at different rates. This is beyond the scope of this paper.

As an example, we obtain the limiting distribution of the least squares estimator for a VAR(1) model when the observations are generated by a multivariate random walk with disturbances obtained from an infinite vector moving average process. This result extends Phillips and Durlauf [23, Theorem 3.2] in essentially the same way that Example 3.3 extends Phillips [20, Theorem 3.1].

Example 4.3. Suppose that the l-variate ($l \in \mathbb{N}$) process $\{y_t\}$ is generated according to

$$y_t = y_{t-1} + u_t,$$

$$u_t = \sum_{\tau=0}^{\infty} A_\tau \epsilon_{t-\tau}, \qquad t = 1,2,\ldots,$$

with the vector y_0 having a given distribution. Suppose further that

i. $\sup_t \|\epsilon_{ti}\|_4 < \infty$, and $E(\epsilon_{ti}) = 0$, $i = 1,\ldots,l$, $t = 0, \pm 1,\ldots,$
ii. $\{A_{\tau ij}\}$ is of size $-\frac{3}{2}$, $i,j = 1,\ldots,l$,
iii. $\{\epsilon_t\colon t = 0, \pm 1,\ldots\}$ is an independent sequence of random vectors,
iv. (a) $\{u_t\}$ is weakly stationary; (b) $\Sigma_u \equiv E(u_t u_t')$ is positive definite; (c) $\Sigma_o \equiv \lim_{n\to\infty} n^{-1}E\left[\left(\sum_{t=1}^{n} u_t\right)\left(\sum_{t=1}^{n} u_t'\right)\right]$ is positive definite.

Define $\hat{B}_n \equiv Y'Y_{-1}(Y_{-1}'Y_{-1})^{-1}$, where Y is the $n \times l$ matrix with rows y_t', and Y_{-1} is the $n \times l$ matrix with rows y_{t-1}'. Then, defining $\Sigma_1 \equiv \sum_{k=2}^{\infty} E(u_1 u_k')$,

$$n(\hat{B}_n - I_l) \xrightarrow{L} \left[\Sigma_o^{1/2}\int_0^1 W(a)dW(a)'\Sigma_o^{1/2} + \Sigma_1'\right] \times \left[\Sigma_o^{1/2}\int_0^1 W(a)W(a)'da\Sigma_o^{1/2}\right]^{-1}.$$

Further,

$$2n^{-1}\operatorname{tr}(\hat{B}_n - I_l)(Y_{-1}'Y_{-1})\Sigma_o^{-1} + \operatorname{tr}\Sigma_u\Sigma_o^{-1} \xrightarrow{L} \chi_l^2.$$

We have imposed covariance stationarity on $\{u_t\}$ so that the result of Phillips [18, Theorem 2.6] for convergence to a matrix stochastic integral can be easily adapted to the present situation without altering the mixing or near-epoch dependence rates. (See also Phillips [22].) Phillips [18, p. 9] notes that heterogeneity in the second moments of $\{u_t\}$ could be allowed at the cost of faster mixing rates. We do not attempt this generalization here.

5. SUMMARY AND CONCLUDING REMARKS

This paper presents several FCLT's and CLT's applicable to stochastic processes exhibiting considerable dependence and heterogeneity as well as possibly trending moments. Our results synthesize some results of McLeish [14,15], and provide relatively primitive conditions which make the present results somewhat more amenable to econometric application. We have given several examples in which we obtain the limiting distribution of the least squares estimator for an AR(1) model applied to data generated by a unit root process, extending results of Phillips [20] and Phillips and Durlauf [23].

Because of the presence of stochastic integrals in the asymptotic distribution for the least squares estimators, these distributions are not convenient for hypothesis testing. However, convenient statistics for hypothesis testing can be obtained from transformations of the least squares estimator. Interestingly, the distributions of these statistics do not require application of an FCLT, but only a CLT. If one is only interested in obtaining tractable test statistics, it may therefore be productive to focus attention on obtaining statistics which require only a CLT for the analysis of asymptotic distributions. However, if one wishes to understand the behavior of the least squares estimator or the familiar associated test statistics in unit root contexts, then the FCLT becomes an indispensable tool.

REFERENCES

1. Andrews, D.W.K. Non-strong mixing autoregressive processes. *Journal of Applied Probability* 21 (1983): 930-934.
2. Andrews, D.W.K. A near independent, but non-strong mixing triangular array. *Journal of Applied Probability* 22 (1984): 729-731.
3. Andrews, D.W.K. Least squares regression with integrated or dynamic regressors under weak error assumptions. *Econometric Theory* 3 (1987): forthcoming.
4. Bierens, H.J. Uniform consistency of kernel estimators of a regression function under generalized conditions. *Journal of the American Statistical Association* 77 (1983): 699-707.
5. Billingsley, P. *Convergence of probability measures.* New York: Wiley, 1968.
6. Billingsley, P. *Probability and measure.* New York: Wiley, 1979.
7. Durrett, R. & S. Resnick. Functional limit theorems for dependent variables. *Annals of Probability* 6 (1978): 829-846.
8. Gallant, A.R. *Nonlinear statistical models.* New York: Wiley, 1987.
9. Gallant, A.R. & H. White. *A unified theory of estimation and inference for nonlinear dynamic models.* Oxford: Basil Blackwell, 1988.
10. Granger, C.W.J. Models that generate trends. UCSD Department of Economics Discussion Paper, 1985.

11. Granger, C.W.J. & R. Joyeaux. An introduction to long-memory time series models and fractional differencing. *Journal of Time Series Analysis* 1 (1980): 15–29.
12. Hall, P. & C.C. Heyde. *Martingale limit theory and its application.* New York: Academic Press, 1980.
13. McLeish, D.L. A maximal inequality and dependent strong laws. *Annals of Probability* 3 (1975): 826–836.
14. McLeish, D.L. Invariance principles for dependent variables. *Zeitschrift für Wahrscheinlichskeitstheorie und Verwandete Gebiete* 32 (1975): 165–178.
15. McLeish, D.L. On the invariance principle for nonstationary mixingales. *Annals of Probability* 5 (1977): 616–621.
16. Park, J.Y. & P.C.B. Phillips. Statistical inference in regressions with integrated processes: part 1. Cowles Foundation Discussion Paper 811, 1986.
17. Park, J.Y. & P.C.B. Phillips. Statistical inference in regressions with integrated processes: part 2. Cowles Foundation Discussion Paper 819, 1987.
18. Phillips, P.C.B. Weak convergence to the matrix stochastic integral $\int BdB'$. Cowles Foundation Discussion Paper 796, 1986.
19. Phillips, P.C.B. Understanding spurious regressions in econometrics. *Journal of Econometrics* 33 (1986): 311–340.
20. Phillips, P.C.B. Time series regression with a unit root. *Econometrica* 55 (1987): 277–302.
21. Phillips, P.C.B. Asymptotic expansions in nonstationary vector autoregressions. *Econometric Theory* 3 (1987): forthcoming.
22. Phillips, P.C.B. Weak convergence of sample covariance matrices to stochastic integrals via martingale approximations. Cowles Foundation Discussion Paper 846, 1987.
23. Phillips, P.C.B. & S.N. Durlauf. Multiple time series regression with integrated processes. *Review of Economic Studies* 53 (1986): 473–496.
24. Serfling, R.J. Contributions to central limit theory for dependent variables. *Annals of Mathematical Statistics* 39 (1968): 1158–1175.
25. White, H. *Asymptotic theory for econometricians.* New York: Academic Press, 1984.
26. White, H. & I. Domowitz. Nonlinear regression with dependent observations. *Econometrica* 52 (1984): 143–162.
27. Wooldridge, J.M. Asymptotic properties of econometric estimators. Ph.D. Dissertation, University of California, San Diego, 1986.
28. Wooldridge, J.M. & H. White. Some invariance and central limit theorems for dependent heterogeneous processes. UCSD Department of Economics Discussion Paper 86-18, 1986.

APPENDIX

For proofs of Theorem 2.1, Lemma 2.2, Proposition 2.5, Proposition 2.6, Corollary 2.7, Proposition 2.9, and Lemma 2.10, see Wooldridge and White [28] (available on request). Gallant [8, pp. 511–514] contains a detailed proof of Lemma 2.4.

Proof of Theorem 2.11. We verify the conditions of Theorem 2.1. By Proposition 2.9, Assumption B.2 implies Assumptions A.1 and A.2. Because Assumptions A.3 and A.4 are also assumed, it follows from Corollary 2.7 that $\{W_n\}$ is tight in D, and any limit process is almost surely continuous. Corollary 2.7 also shows that for each $a \in [0,1]$, $\{W_n^2(a): n = 1,2,\dots\}$ is uniformly integrable, so part (ii) of Theorem 2.1

is satisfied as well. Because $E[X_{nt}] = 0$ for $n, t = 1,2,\ldots$, it follows that $E[W_n(a)] = 0$, all $a \in [0,1]$. $E[W_n^2(a)] \to a$ as $n \to \infty$ is given as Assumption A.5. Thus, parts (i), (ii), and (iii) of Theorem 2.1 are satisfied. It only remains to show that $\{W_n\}$ has asymptotically independent increments. This follows from Lemma 2.10. Because the conditions of Theorem 2.1 hold, we have $W_n \xrightarrow{L} W$. That $\sum_{t=1}^{n} X_{nt} \xrightarrow{L} N(0,1)$ follows immediately. ■

Proof of Example 2.12. Let $u_t \equiv t^\rho \epsilon_t$, $k_n(a) \equiv [na^{1/(2\rho+1)}]$, and

$$W_n(a) \equiv \sigma_o^{-1} n^{-(\rho+1/2)} \sum_{t=1}^{k_n(a)} u_t, \qquad a \in (0,1].$$

We first verify the conditions of Theorem 2.11 for $Y_{nt} \equiv \eta_t$, $X_{nt} \equiv \sigma_o^{-1} n^{-(\rho+1/2)} t^\rho \epsilon_t$, $d_{nt} \equiv \sigma_o^{-1} n^{-(\rho+1/2)} t^\rho$, $c_{nt} \equiv d_{nt} \max(1, \|\epsilon_t\|_r)$, and W_n as defined above. Assumption A.1 holds because $k_n(a) \equiv [na^{1/(2\rho+1)}]$ is nondecreasing, right continuous and $k_n(0) = 0$. Assumption B.2(i) holds given Condition 2.12(i). For Assumption B.2(ii), note that $k_n(1) = n$; therefore,

$$\max_{1 \le t \le k_n(1)} d_{nt} = \sigma_o^{-1} n^{-1/2},$$

so we take $\gamma = \frac{1}{2}$. Assumption B.2(iii) holds by Condition 2.12(iii), and Assumption B.2(iv) follows by our choices of d_{nt} and c_{nt}.

For the uniform integrability requirement of Assumption A.3, note that $X_{nt}^2/c_{nt}^2 = [\max(1, \|\epsilon_t\|_r)]^{-1} \epsilon_t^2$. Given 2.12(i), $\{\epsilon_t^2\}$ is uniformly integrable, and therefore Assumption A.3 holds. Next, consider Assumption A.4. Now for $\Delta < \infty$ sufficiently large

$$\begin{aligned}\sum_{k_n(a)}^{k_n(a+\delta)} c_{nt}^2 &\le \Delta \sigma_o^{-2} n^{-(2\rho+1)} \sum_{k_n(a)}^{k_n(a+\delta)} t^{2\rho}, \\ &= \Delta \sigma_o^{-2} n^{-(2\rho+1)} \{[n(a+\delta)^{1/(2\rho+1)}]^{(2\rho+1)} - [na^{1/(2\rho+1)}]^{(2\rho+1)}\} + o(1), \\ &= \Delta \sigma_o^{-2} \delta + o(1).\end{aligned}$$

Thus, for all $0 < \delta < 1 - a$, $0 \le a < 1$,

$$\limsup_{n\to\infty} \delta^{-1} \sum_{k_n(a)}^{k_n(a+\delta)} c_{nt}^2 \le \Delta \sigma_o^{-2}$$

so that Assumption A.4 is satisfied. Lastly, consider Assumption A.5. Because

$$W_n^2(a) = \sigma_o^{-2} n^{-(2\rho+1)} \left(\sum_{t=1}^{[na^{1/(2\rho+1)}]} t^\rho \epsilon_t \right)^2,$$

it follows by Condition 2.12(iv.b) that

$$\begin{aligned}E[W_n^2(a)] &= \sigma_o^{-2} n^{-(2\rho+1)} (\sigma_o^2 [na^{1/(2\rho+1)}]^{(2\rho+1)} + o([na^{1/(2\rho+1)}]^{(2\rho+1)})), \\ &= ([na^{1/(2\rho+1)}]/n)^{(2\rho+1)} + o(1) \to a \text{ as } n \to \infty.\end{aligned}$$

This verifies the conditions of Theorem 2.11; therefore $W_n \xrightarrow{L} W$, a standard Brownian motion on $D[0,1]$. Define $Z_n(a) \equiv W_n(a^{(2\rho+1)}) = \sum_{t=1}^{[na]} X_{nt}$. By the continuous mapping theorem, $Z_n(a) \xrightarrow{L} Z(a)$, where Z is the Gaussian process defined by $Z(a) \equiv W(a^{(2\rho+1)})$, $a \in [0,1]$.

As in Phillips [20], we write the OLS (ordinary least squares) estimator as $\hat{\beta}_n - 1 = \left(\sum_{t=1}^{n} y_{t-1}^2\right)^{-1}\left(\sum_{t=1}^{n} y_{t-1}u_t\right)$. Let $S_t \equiv \sum_{j=1}^{t} u_j$. Then

$$\sum_{t=1}^{n} y_{t-1}u_t = (1/2)S_n^2 - (1/2)\sum_{t=1}^{n} u_t^2 + y_0\sum_{t=1}^{n} u_t.$$

By 2.12(i), 2.12(ii.b), and 2.12(iii), $\{n^{-(2\rho+1)}u_t^2\}$ is near-epoch dependent with respect to $\{\eta_t\}$ with ν_m of size $-\frac{1}{2}$ and $d_{nt} = n^{-(2\rho+1)}t$; therefore $\{n^{-(2\rho+1)}u_t^2\}$ is a mixingale of size $-\frac{1}{2}$ with $c_{nt} = n^{-(2\rho+1)}t\max(1, \|\epsilon_t^2\|_r)$. By the weak law of large numbers for mixingales,

$$n^{-(2\rho+1)}\sum_{t=1}^{n}(u_t^2 - Eu_t^2) \xrightarrow{P} 0,$$

i.e., $n^{-(2\rho+1)}\sum_{t=1}^{n} u_t^2 \xrightarrow{P} \sigma_u^2$. Similarly, $n^{-(2\rho+1)}\sum_{t=1}^{n} u_t \xrightarrow{P} 0$. Because $n^{-(2\rho+1)}S_n^2 = \sigma_o^2 Z_n^2(1)$, $n^{-(2\rho+1)}S_n^2 \xrightarrow{L} \sigma_o^2 Z^2(1)$. Collecting these results gives

$$n^{-(2\rho+1)}\sum_{t=1}^{n} y_{t-1}u_t \xrightarrow{L} (1/2)(\sigma_o^2 Z^2(1) - \sigma_u^2).$$

Following Phillips [20],

$$\sum_{t=1}^{n} y_t^2 = \sum_{t=1}^{n}(S_t^2 + 2y_o S_t + y_0^2).$$

Now

$$n^{-(2\rho+2)}\sum_{t=1}^{n} S_t^2 = \sigma_o^2 n^{-1}\sum_{t=1}^{n} Z_n^2(t/n),$$
$$= \sigma_o^2\int_0^1 Z_n^2(a)da \xrightarrow{L} \sigma_o^2\int_0^1 Z^2(a)da.$$

Also

$$n^{-(2\rho+1)}\sum_{t=1}^{n} S_t = \sigma_o n^{-1}\sum_{t=1}^{n} Z_n(t/n) = \sigma_o\int_0^1 Z_n(a)da \xrightarrow{L} \sigma_o\int_0^1 Z(a)da.$$

Therefore,

$$n^{-(2\rho+2)}\sum_{t=1}^{n} S_t \xrightarrow{P} 0,$$

so that

$$n^{-(2\rho+2)}\sum_{t=1}^{n} y_t^2 \xrightarrow{L} \sigma_o^2\int_0^1 Z^2(a)da.$$

Combining the results for the numerator and denominator yields

$$n(\hat{\beta}_n - 1) \xrightarrow{L} \left\{\sigma_o^2 \int_0^1 Z^2(a)da\right\}^{-1} (1/2)(\sigma_o^2 Z^2(1) - \sigma_u^2)$$

and completes the first part of the proof.

To show that

$$2\sigma_o^{-2} n^{-(2\rho+1)}\left(\sum_{t=1}^{n} y_{t-1}^2\right)(\hat{\beta}_n - 1) + \sigma_u^2/\sigma_o^2 \xrightarrow{L} \chi_1^2,$$

substitute the expression for $\hat{\beta}_n$ to obtain

$$2\sigma_o^{-2} n^{-(2\rho+1)}\left[\sum_{t=1}^{n} y_{t-1} y_t - \sum_{t=1}^{n} y_{t-1}^2\right] + \sigma_u^2/\sigma_o^2$$

$$= 2\sigma_o^{-2} n^{-(2\rho+1)}\left[\sum_{t=1}^{n} y_{t-1} u_t\right] + \sigma_u^2/\sigma_o^2.$$

We see here that $n^{-(2\rho+1)}\sum_{t=1}^{n} y_{t-1}u_t \xrightarrow{L} (\sigma_o^2\chi_1^2 - \sigma_u^2)/2$, so the second result follows. Note that the invariance principle plays no role here. ■

Proof of Corollary 3.1. We verify the conditions of Theorem 2.11 for $\{X_{nt}\}$. Assumption A.1 is implied by Assumption C.1 with $k_n(a) = [na]$. This choice for k_n is obviously right continuous and increasing on [0,1]. Further, $[na] - [nb] \geq [n(a-b)] \to \infty$ as $n \to \infty$ for all $0 \leq b < a \leq 1$. That Assumption B.2 holds is straightforward to show. Given Assumption C.2(i), we have $\|X_{nt}\|_r = \sigma_n^{-1}\|Z_{nt}\|_r \leq \sigma_n^{-1}\Delta$, while Assumption C.1 implies $\sigma_n^{-1} \leq \Delta n^{-1/2}$, so that $\|X_{nt}\|_r \leq n^{-1/2}\Delta^2 < \infty$, satisfying Assumption B.2(i). (A referee points out that σ_n^2 could be zero for a finite number of n, violating Assumption B.2, but that this has no effect on the results.) Because $\{Z_{nt}\}$ is near-epoch dependent with respect to $\{Y_{nt}\}$ of size $-\frac{1}{2}$ with $d_{nt} = 1$, it follows immediately that $\{X_{nt}\}$ is near-epoch dependent with respect to $\{Y_{nt}\}$ of size $\gamma = -\frac{1}{2}$ with $d'_{nt} = \sigma_n^{-1}$. Given Assumption C.1, we have that σ_n^{-2} is $O(n^{-1})$, which implies that d'_{nt} is $O(n^{-1/2})$ as required, so that Assumption B.2(ii) holds. Assumption C.2(iii) is identical to Assumption B.2.(iii) and Assumption C.2(iv) immediately implies Assumption B.2(iv). Assumption A.3 holds given Assumption B.2 (with $r > 2$) by Proposition 2.9.

To verify Assumption A.4, let $0 \leq a \leq 1$, $o < \delta \leq 1 - a$ be given, so that $a + \delta \leq 1$. Now

$$\delta^{-1} \sum_{t=[na]}^{[n(a+\delta)]} c_{nt}^2 \leq \delta^{-1}\sigma_n^{-2}\Delta([n(a+\delta)] - [na] + 1) \leq \delta^{-1}\sigma_n^{-2}\Delta 3n\delta$$

for all n sufficiently large. Now Assumption C.1 ensures $n\sigma_n^{-2} \leq \Delta$ for n sufficiently large, which implies that

$$\limsup_{n\to\infty} \delta^{-1} \sum_{t=[na]}^{[n(a+\delta)]} c_{nt}^2 \leq 3\Delta^2 < \infty$$

for all $0 \le a \le 1$, $0 < \delta \le 1 - a$. Hence, Assumption A.4 holds. This establishes the first result.

Assumption A.5 is now directly imposed, which establishes the second result. ■

Proof of Corollary 3.2. We apply Corollary 3.1. Obviously, it suffices to show that $\sigma_n^2/n \to \sigma_o^2 > 0$ implies Assumption A.5 for $k_n(a) = [na]$, i.e., that for each $a \in [0,1]$,

$$\operatorname{var}\left(\sum_{t=1}^{[na]} X_{nt}\right) \to a, \qquad \text{as } n \to \infty.$$

Because $\sigma_n^2/n \to \sigma_o^2$, it follows that $\sigma_{[na]}^2/[na] \to \sigma_o^2$ as $n \to \infty$. But

$$\begin{aligned}\operatorname{var}\left(\sum_{t=1}^{[na]} X_{nt}\right) &= \sigma_n^{-2} \operatorname{var}\left(\sum_{t=1}^{[na]} Z_t\right),\\ &= \sigma_n^{-2}\sigma_{[na]}^2,\\ &= \frac{[na]}{n}\left(\frac{\sigma_{[na]}^2}{[na]}\right)\left(\frac{\sigma_n^2}{n}\right)^{-1} \to a, \qquad \text{as } n \to \infty,\end{aligned}$$

because $[na]/n \to a$ as $n \to \infty$. Thus, Assumption A.5 holds, and the result now follows from Corollary 3.1. ■

Proof of Example 3.3. First we show that u_t and ϵ_t satisfy the conditions of Corollary 3.2. Take $Y_{nt} \equiv \epsilon_t$, $Z_{nt} \equiv u_t$. Because $\sigma_n^2/n = n^{-1}E\left(\sum_{t=1}^{n} u_t\right)^2 \to \sigma_o^2 > 0$, $\sigma_n^{-2} = O(n^{-1})$. Thus, Assumption C.1 holds. For Assumption C.2, fix $2 < r \le 4$. Then Assumption C.2(i) holds by condition 3.3(i). Next, note that

$$\begin{aligned}u_t - E_{t-m}^{t+m}(u_t) &= u_t - E_{t-m}^{t}(u_t),\\ &= \sum_{\tau=m+1}^{\infty} a_\tau \epsilon_{t-\tau}.\end{aligned}$$

By Minkowski's inequality for infinite sums (see White [25, Exercise 3.53])

$$\|u_t - E_{t-m}^{t} u_t\|_2 \le \sum_{\tau=m+1}^{\infty} |a_\tau|\,\|\epsilon_{t-\tau}\|_2 \le \Delta \sum_{\tau=m}^{\infty} |a_\tau|.$$

By Condition 3.3(ii), $|a_\tau| < \tau^{-\gamma}$ for all $\gamma > \frac{3}{2}$, so that $\sum_{\tau=m}^{\infty} |a_\tau| < \int_m^\infty \tau^{-\gamma} d\tau = (\gamma - 1)^{-1} m^{1-\gamma}$; it follows that $\{u_t\}$ is near-epoch dependent with respect to ϵ_t with ν_m of size $-\frac{1}{2}$ and $d_t = 1$. Because $\{\epsilon_t\}$ is an independent sequence, it is ϕ-mixing of any size; consequently Assumption C.2(iii) holds. Finally, $\sigma_n^2/n \to \sigma_o^2$ by Condition 3.3(iv.b). It follows from Corollary 3.2 that $W_n \xrightarrow{L} W$. In view of Example 2.12, it suffices to show that $n^{-1}\sum_{t=1}^{n} (u_t^2 - Eu_t^2) \xrightarrow{P} 0$. To this end, we show that u_t^2 is a mixingale of size $-\frac{1}{2}$. This follows if we show u_t^2 is near-epoch dependent with respect to ϵ_t with ν_m of size $-\frac{1}{2}$. Now

$$\begin{aligned}\|u_t^2 - E_{t-m}^t u_t^2\|_2 &\leq \|u_t^2 - (E_{t-m}^t u_t)^2\|_2,\\ &\leq \|\, |u_t + E_{t-m}^t u_t|\; |u_t - E_{t-m}^t u_t|\, \|_2,\\ &\leq \|u_t + E_{t-m}^t u_t\|_4 \|u_t - E_{t-m}^t u_t\|_4,\\ &\leq 2\Delta \sum_{\tau=m}^{\infty} |a_\tau|\, \|\epsilon_{t-\tau}\|_4,\\ &\leq \Delta' \sum_{\tau=m}^{\infty} |a_\tau|.\end{aligned}$$

By Condition 3.3(ii), $\{u_t^2\}$ satisfies the appropriate near-epoch dependence condition. The results now follow as in Example 2.12 or as in Phillips [20, Theorem 3.1]. ∎

Proof of Proposition 4.1. By Lemma 1 in Phillips and Durlauf [23], the finite dimensional sets $\{\Pi_{a_1,\ldots,a_k}^{-1}H\colon k \in \mathbb{N},\, a_1,\ldots,a_k \in [0,1],\, H \in B^{kl}\}$ form a determining class of D^l. By a straightforward extension of their Lemma 2 (p. A4),

$$\mu_n[\Pi_{a_1,\ldots,a_k}^{-1}H] \to \mu[\Pi_{a_1,\ldots,a_k}^{-1}H],$$

where μ_n is the distribution of W_n, and μ is the distribution of W; this convergence holds for all $k \in \mathbb{N}$, $a_1,\ldots,a_k \in [0,1]$ and Borel sets $H \in B^{kl}$. By the discussion in Billingsley [5, p. 35], it now suffices to show $\{\mu_n\}$ is tight. By Lemma 3 of Phillips and Durlauf [23], it is enough to show each of the marginal distributions of $\{\mu_n\}$ is tight. But letting $\lambda = e_i$, $i = 1,\ldots,l$ it follows by assumption that $W_{ni} \xrightarrow{L} W_i$, so that $\{\mu_{ni}\}$ is sequentially compact. By Prohorov's Theorem (Billingsley [5], Theorem 6.2), $\{\mu_{ni}\}$ is tight. This completes the sufficiency part.

For each $\lambda \in \mathbb{R}^l$, the linear functional defined on D^l by

$$L_\lambda(f) \equiv \sum_{i=1}^{l} \lambda_i f_i, \qquad f \in D^l$$

is continuous. Because $W_n \xrightarrow{L} W$ is assumed, it follows by the continuity theorem that $L_\lambda(W_n) \xrightarrow{L} L_\lambda(W)$, i.e., $\lambda' W_n \xrightarrow{L} \lambda' W$. ∎

Proof of Corollary 4.2. We show that for each $\lambda \in \mathbb{R}^l$ with $\lambda'\lambda = 1$, $\lambda' W_n$ converges in distribution to a univariate Brownian motion on $\mathbb{R}$. To do this, we verify the conditions of Theorem 2.11 for the random variables $\lambda' X_{nt}$, where $\lambda'\lambda = 1$.

As in the proof of Corollary 3.1, the choice $k_n(a) = [na]$ imposed in Assumption D.1 satisfies Assumption A.1 and the required conditions on $k_n(a)$.

Regarding Assumption B.2, it follows from the Minkowski inequality that

$$\begin{aligned}\|\lambda' X_{nt}\|_r &\leq \sum_{i=1}^{l} |\lambda_i| \sum_{j=1}^{l} |C_{nij}|\, \xi_{nj}^{-1/2} \|Z_{ntj}\|_r,\\ &\leq l\Delta \sum_{j=1}^{l} \xi_{nj}^{-1/2} < \infty\end{aligned}$$

given Assumptions D.1 and D.2(ii), so that Assumption B.2(i) holds.

Minkowski's inequality also implies that

$$\|\lambda' X_{nt} - E[\lambda' X_{nt} | G^{t+m}_{n,t-m}]\|_2 \le \sum_{i=1}^{l} |\lambda_i| \; \|X_{nti} - E[X_{nti} | G^{t+m}_{n,t-m}]\|_2,$$

$$\le \sum_{i=1}^{l} |\lambda_i| \sum_{j=1}^{l} |C_{nij}| \xi_{nj}^{-1/2} \|Z_{ntj} - E[Z_{ntj} | G^{t+m}_{n,t-m}]\|_2,$$

where ξ_{nj} is the jth eigenvalue of V_n and $C_n = [C_{nij}]$ is an orthonormal matrix containing the eigenvectors of V_n. Now $\|Z_{ntj} - E[Z_{ntj} | G^{t+m}_{n,t-m}]\|_2 \le \nu_m$, $j = 1, \ldots, l$. Further, $|C_{nij}| \le 1$, while $\lambda'\lambda = 1$ implies $\sum_{i=1}^{l} |\lambda_i| \le l$. Thus, we may take $d'_{nt} = l \sum_{j=1}^{l} \xi_{nj}^{-1/2}$. Because $\xi_{nj}^{-1} = O(n^{-1})$ by Assumption D.1, it follows that d'_{nt} is $O(n^{-1/2})$, thus satisfying Assumption B.2(ii).

Now Assumption D.2(iii) directly imposes Assumption B.2(iii), while the choice $c'_{nt} = \sum_{j=1}^{l} c_{ntj} = l \sum_{j=1}^{l} \xi_{nj}^{-1/2} \max(1, \|Z_{ntj}\|_r)$ satisfies Assumption B.2(iv).

Assumption A.3 holds for $\lambda' X_{nt}$ with this choice of c'_{nt} by Proposition 2.9 given Assumption B.2 with $r > 2$. To verify Assumption A.4, let $0 \le a \le 1$, $0 < \delta \le 1 - a$ be given so that $a + \delta \le 1$. Now

$$\delta^{-1} \sum_{t=[na]}^{[n(a+\delta)]} c'^2_{nt} \le \delta^{-1} \sum_{t=[na]}^{[n(a+\delta)]} l^2 \left[\sum_{j=1}^{l} \xi_{nj}^{-1/2} \max(1, \|Z_{ntj}\|_r) \right]^2,$$

$$= \delta^{-1} \sum_{t=[na]}^{[n(a+\delta)]} l^2 \sum_{i=1}^{l} \sum_{j=1}^{l} \xi_{ni}^{-1/2} \xi_{nj}^{-1/2} \max(1, \|Z_{nti}\|_r) \max(1, \|Z_{ntj}\|_r),$$

$$\le \delta^{-1} l^4 \Delta n^{-1} ([n(a+\delta)] - [na] + 1),$$

because $\xi_{nj}^{-1/2} = O(n^{-1/2})$, and for $\Delta < \infty$ sufficiently large we have $\sum_{i=1}^{l} \sum_{j=1}^{l} \xi_{ni}^{-1/2} \xi_{nj}^{-1/2} \max(1, \|Z_{nti}\|_r) \max(1, \|Z_{ntj}\|_r) \le l^2 \Delta n^{-1}$. As in the proof of Corollary 3.1, we have $[n(a+\delta)] - [na] + 1 \le 3n\delta$ for all n sufficiently large. Therefore,

$$\delta^{-1} \sum_{t=[na]}^{[n(a+\delta)]} c'^2_{nt} \le \delta^{-1} l^4 \Delta n^{-1} 3n\delta = 3l^4 \Delta < \infty$$

for n sufficiently large, verifying Assumption A.4. Thus, the first result is established.

To complete the proof, we verify that Assumption A.5 holds, i.e., that $E[\lambda' W_n(a)]^2 \to a$, i.e., $\lambda' E[W_n(a) W_n(a)'] \lambda \to a$; but this follows from Assumption D.3 because $\lambda'\lambda = 1$. This completes the proof. ∎

Proof of Example 4.3. Showing that u_t, ϵ_t satisfy the conditions of Corollary 4.2 follows along the lines of Example 3.3. Therefore, $W_n \xrightarrow{L} W$. Establishing that $n^{-1} \sum_{t=1}^{n} (u_t u'_t - E u_t u'_t) \xrightarrow{P} 0$ and $n^{-1} \sum_{t=1}^{n} u_t \xrightarrow{P} 0$ also follows as in Example 3.3. The analysis for the first result is now identical to Phillips and Durlauf [23, Theorem 3.2]. For the second result, note that

$$2n^{-1} \operatorname{tr}(\hat{B}_n - I)(Y'_{-1} Y_{-1}) \Sigma_o^{-1} + \operatorname{tr} \Sigma_u \Sigma_o^{-1}$$

$$= \operatorname{tr} \Sigma_o^{-1/2} \left(n^{-1} \sum_{t=1}^{n} u'_t y_{t-1} + n^{-1} \sum_{t=1}^{n} y'_{t-1} u_t + \Sigma_u \right) \Sigma_o^{-1/2} \xrightarrow{L} \operatorname{tr} W(1) W(1)' = \chi^2_l$$

as in Phillips and Durlauf [23, Lemma 3.1(d)]. ∎