Econometric Theory, **16**, 2000, 621–642. Printed in the United States of America.

THE FUNCTIONAL CENTRAL LIMIT THEOREM AND WEAK CONVERGENCE TO STOCHASTIC INTEGRALS I

Weakly Dependent Processes

ROBERT M. DE JONG
Michigan State University

JAMES DAVIDSON
Cardiff University

This paper gives new conditions for the functional central limit theorem, and weak convergence of stochastic integrals, for near-epoch-dependent functions of mixing processes. These results have fundamental applications in the theory of unit root testing and cointegrating regressions. The conditions given improve on existing results in the literature in terms of the amount of dependence and heterogeneity permitted, and in particular, these appear to be the first such theorems in which virtually the same assumptions are sufficient for both modes of convergence.

1. INTRODUCTION

Asymptotic theory for integrated processes is an area of research where results from functional limit theory are crucial. These results are the main underpinning of the econometric analysis of models with integrated and cointegrated variables. Phillips (1986, 1987), Phillips and Durlauf (1986), Park and Phillips (1988, 1989), and Johansen (1988, 1991) are the well-known seminal contributions to what is now a very extensive literature.

In this theory, the sample statistics whose distributions are sought are typically functions of sample moments in which the data may be (a) stationary or (b) integrated or (c) a mixture of both. The asymptotic analysis of each of these cases requires a different technique. Case (a) is the standard one leading to Gaussian limit distributions. In case (b), weak convergence to functionals of Brownian motion or related Gaussian processes must usually be proved, and the technique of analysis is to combine a multivariate functional central limit theorem (FCLT) with the continuous mapping theorem. It is important, espe-

We thank Bruce Hansen, Peter Phillips, and two anonymous referees for their comments on earlier versions of this paper. Any errors are ours alone. Address correspondence to James Davidson, Cardiff Business School, Cardiff University, Colum Drive, Cardiff CF1 3EU, UK.

 0266-4666/00 $9.50

cially for applications to economic data, that a wide latitude should be permitted in the amount and type of dependence and heterogeneity in the random variables under consideration. In case (c), the limits in question are stochastic integrals (Itô integrals), and except in the univariate case, to show weak convergence calls for a different technique of proof.

Results of type (b) are applied in all the previously cited studies, and results of type (c) are also crucial in all but the first two. It is noteworthy in view of the now routine use of tests based on these asymptotics (with critical values obtained by simulation) that in the latter case, the available proofs of weak convergence impose relatively stringent conditions on the amount and form of permitted dependence. For example, Strasser (1986), Chan and Wei (1988), and Jeganathan (1991) impose a martingale difference assumption, ruling out serial correlation of the increments, at least of the integrator process. Phillips (1988b) considers linear processes with independent and identically distributed (i.i.d.) innovations, and Hansen (1992) allows strong mixing, but all the cited conditions are stronger than are known to be sufficient for the FCLT for the same multivariate process. Moreover, the results given by Phillips (1988a) and Davidson (1994) contain errors.[1]

In this paper, we give new conditions for the multivariate FCLT and stochastic integral convergence. The former result dominates the existing ones in the econometrics literature using comparable assumptions, such as Wooldridge and White (1988) and Davidson (1994, Theorem 29.18). The conditions are only a little stronger than the best comparable ones for the ordinary central limit theorem (CLT). Moreover, our results for stochastic integral convergence impose virtually the same conditions as the FCLT and so represent a substantial improvement over previous results. Section 2 sets out the main assumptions; Section 3 discusses the FCLT and Section 4 the corresponding stochastic integral convergence result. Section 5 concludes the paper. The proofs are gathered in Appendixes A–C.

2. DEFINITIONS AND ASSUMPTIONS

A key issue in this theory is the method of characterizing weak dependence of the underlying time series. We follow authors such as Gallant and White (1988) and Pötscher and Prucha (1991) in working with the concept of *near-epoch dependence on a mixing process*. This framework has considerable generality. Whereas additional dependence can be allowed in specific cases such as linear processes (see Davidson 2000; Phillips and Solo 1992), our assumption is more likely to be robust in cases of partially specified models, in which aspects of the short-run data generation process are unknown. Such situations are endemic in econometric research. In addition to including infinite-order moving averages under suitable summability conditions, near-epoch dependence can be shown to be satisfied in various nonlinear dynamic processes. See Davidson (2000) for examples. Mixing processes are also allowed.

Our definition of near-epoch dependence is as follows. Let X_{nt} denote a triangular array of random variables defined on the probability space $(\Omega, \mathcal{F}, P)$ and let $\|X\|_p$ denote $(E|X|^p)^{1/p}$ for $p \geq 1$.

DEFINITION 1. *X_{nt} is called L_2-NED on random variables V_{nt} if for $m \geq 0$,*

$$\|X_{nt} - E(X_{nt} | \mathcal{F}_{n,t-m}^{t+m})\|_2 \leq d_{nt}\nu(m), \tag{2.1}$$

where $\mathcal{F}_{ns}^t = \sigma(V_{ns}, \ldots, V_{nt}) \subset \mathcal{F}$ for $t \geq s$, d_{nt} is an array of positive constants, and $\nu(m) \to 0$ as $m \to \infty$.

We refer to the d_{nt} as the "near-epoch-dependence (NED) magnitude indices" and to the $\nu(m)$ as the "NED numbers." A sequence such as $\nu(m)$ is said to be of *size* $-\lambda$ if $\nu(m) = O(m^{-\lambda-\varepsilon})$ for some $\varepsilon > 0$, and we also say that X_{nt} is NED of size $-\lambda$ on the process V_{nt}. In the application, V_{nt} can be a mixing process. Of the different mixing concepts that have been defined, the econometrics literature usually adopts either uniform (ϕ-) or strong (α-) mixing, and we consider both of these cases, with similar "size" terminology for the α- and ϕ-mixing numbers. For definitions and details see the preceding references and also Davidson (1994).

De Jong (1997) appears to provide the most general CLT for NED functions of mixing processes currently available. Letting $\{X_{nt}\}$ denote a triangular stochastic array, it is shown in that paper that $\sum_{t=1}^{K_n} X_{nt} \xrightarrow{d} N(0,1)$, where K_n is an integer-valued increasing sequence, if the following assumption holds.[2]

Assumption 1.

(a) X_{nt} has mean zero and $\|\sum_{t=1}^{K_n} X_{nt}\|_2 = 1$.
(b) There exists a positive constant array c_{nt} such that $\{X_{nt}/c_{nt}\}$ is L_r-bounded for $r > 2$ uniformly in t and n.
(c) X_{nt} is L_2-NED of size $-\frac{1}{2}$ on V_{nt}, where V_{nt} is an α-mixing array of size $-r/(r-2)$, or X_{nt} is L_2-NED of size $-\frac{1}{2}$ on V_{nt}, where V_{nt} is a ϕ-mixing array of size $-r/(2(r-1))$, and d_{nt}/c_{nt} is bounded uniformly in t and n.
(d) For some sequence b_n such that $b_n = o(K_n)$ and $b_n^{-1} = o(1)$, letting $r_n = [K_n b_n^{-1}]$, $M_{ni} = \max_{(i-1)b_n+1 \leq t \leq ib_n} c_{nt}$ and $M_{n,r_n+1} = \max_{r_n b_n+1 \leq t \leq K_n} c_{nt}$,

$$\max_{1 \leq i \leq r_n+1} M_{ni} = o(b_n^{-1/2}), \tag{2.2}$$

and

$$\sum_{i=1}^{r_n} M_{ni}^2 = O(b_n^{-1}). \tag{2.3}$$

In the case of ϕ-mixing, $r = 2$ is allowed also if the assumption of uniform integrability of X_{nt}^2/c_{nt}^2 is added to Assumption 1(b).

3. A FUNCTIONAL CENTRAL LIMIT THEOREM

Let

$$X_n(\xi)=\sum_{t=1}^{K_n(\xi)} X_{nt} \quad \text{for } \xi \in [0,1], \tag{3.1}$$

where $\{K_n(\xi), n \geq 1\}$ is a sequence of integer-valued, right-continuous, increasing functions of ξ, with $K_n(0) = 0$ for all $n \geq 1$; $K_n(\xi)$ is nondecreasing in n for all $\xi \in [0,1]$; and $K_n(\xi) - K_n(\xi') \to \infty$ as $n \to \infty$ if $\xi > \xi'$. The reference case obviously is $X_{nt} = n^{-1/2}X_t$, for some sequence of random variables X_t, with $K_n(\xi) = [n\xi]$. This framework is basically the same as that of Wooldridge and White (1988) and Davidson (1994, Ch. 29).

THEOREM 3.1. *Let Assumption* 1 *hold for* X_{nt} *and assume that*

$$\eta(\xi) = \lim_{n\to\infty} EX_n(\xi)^2 \tag{3.2}$$

exists for all $\xi \in [0,1]$ *and that*[3]

$$\lim_{\delta\to 0} \sup_{\xi\in[0,1-\delta]} \limsup_{n\to\infty} \sum_{t=K_n(\xi)+1}^{K_n(\xi+\delta)} c_{nt}^2 = 0. \tag{3.3}$$

Then $X_n(\xi) \xrightarrow{d} X(\xi)$, *where* $X(\xi)$ *is a Gaussian process having almost surely* (*a.s.*) *continuous sample paths and independent increments.*

The line of argument we adopt to prove Theorem 3.1 contrasts with that of Wooldridge and White (1988) and Davidson (1994). They obtain the FCLT by a direct proof that generates the central limit theorem as a corollary, whereas we start with the finite dimensional distributions. Under Assumption 1,

$$(X_n(\xi_1),\ldots,X_n(\xi_k)) \xrightarrow{d} (X(\xi_1),\ldots,X(\xi_k)) \tag{3.4}$$

for any finite collection of coordinates $\xi_1,\ldots,\xi_k \in [0,1]$, where the limit distributions are a.s. continuous and Gaussian. This follows from Theorem 2 of De Jong (1997) and the Cramér–Wold theorem (Davidson, 1994, Theorem 25.5). According to Theorems 15.4 and 15.5 of Billingsley (1968), $X_n \xrightarrow{d} X$, where X is continuous with probability one, if (3.4) holds and X_n is stochastically equicontinuous. This is the property that for all $\varepsilon > 0$,

$$\lim_{\delta\to 0} \limsup_{n\to\infty} P\Big(\sup_{\xi\in[0,1]} \sup_{\{\xi':|\xi-\xi'|<\delta\}} |X_n(\xi) - X_n(\xi')| > \varepsilon\Big) = 0. \tag{3.5}$$

Therefore, to complete the proof it suffices to show that (3.5) holds for the partial sum process and that the increments are independent in the limit. These arguments are set out in Appendix B.

If $\eta(\xi) = \xi$ then X is Brownian motion. More generally, X belongs to an extension of the class of transformed Brownian motion processes B_η defined in Davidson (1994, Ch. 29.4). The term $\eta(\xi)$ must be nondecreasing everywhere on $[0,1]$, but under the present generalization it need not be strictly increasing everywhere, and increments of the process may equal zero a.s.

The dependence and heterogeneity conditions of Theorem 3.1 relax those employed by Wooldridge and White (1988) and Davidson (1994), whose conditions are similar. These latter theorems do not permit a size $-\frac{1}{2}$ of the NED coefficients but employ a trade-off condition. The rate at which $\max_{1\le t\le n} c_{nt}$ is required to approach zero is dictated by this trade-off condition, and also the condition

$$\sup_{\xi\in[0,1),\,\delta\in(0,1-\xi)} \limsup_{n\to\infty} \sum_{t=K_n(\xi)+1}^{K_n(\xi+\delta)} c_{nt}^2/\delta < \infty \tag{3.6}$$

is imposed, which is obviously stronger than (3.3). For example, consider $K_n(\xi) = [n\xi]$ and $c_{nt} = t^{-\beta} n^{\beta-1/2}$ for $\beta \in (0,\frac{1}{2})$, appropriate to the case $X_{nt} = t^{-\beta} n^{\beta-1/2} u_t$, where u_t is i.i.d. with finite variance. Then,

$$\limsup_{n\to\infty} \sum_{t=K_n(\xi)+1}^{K_n(\xi+\delta)} c_{nt}^2/\delta \ge C((\xi+\delta)^{1-2\beta} - \xi^{1-2\beta})/\delta \tag{3.7}$$

for $C > 0$, and clearly the condition from equation (3.6) does not hold, whereas the condition from equation (3.3) does hold.

Plausible examples in which condition (3.3) fails to hold are not easy to construct, but consider the case $c_{nt}^2 = t^{1/4} n^{-1/3} I(t \le n^{4/15})$. Note that $\max_{1\le t\le n} c_{nt} = o(1)$ and that

$$\limsup_{n\to\infty} \sum_{t=1}^{n} c_{nt}^2 = \limsup_{n\to\infty} n^{-1/3} \sum_{t=1}^{n^{4/15}} t^{1/4} = \frac{4}{5}. \tag{3.8}$$

One can clearly find a sequence b_n such that Assumption 1(d) holds here. However,

$$\begin{aligned}
\lim_{\delta\to 0} \sup_{\xi\in[0,1-\delta]} \limsup_{n\to\infty} \sum_{t=[\xi n]+1}^{[(\xi+\delta)n]} c_{nt}^2 &\ge \lim_{\delta\to 0} \limsup_{n\to\infty} \sum_{t=1}^{[\delta n]} t^{1/4} n^{-1/3} I(t \le n^{4/15}) \\
&= \lim_{\delta\to 0} \min\left\{ \limsup_{n\to\infty} \sum_{t=1}^{[\delta n]} t^{1/4} n^{-1/3}, \frac{4}{5} \right\} \\
&= \frac{4}{5}.
\end{aligned} \tag{3.9}$$

The limit of a process whose increments have variances evolving like c_{nt}^2 in this example has a discontinuity at the origin. In other words, $X(0) = 0$ a.s., but for $0 < \xi \le 1$, $X(\xi) = Y$ a.s., where Y is distributed as $N(0,\frac{4}{5})$.

Extending Theorem 3.1 to the multivariate case is straightforward, and we have the following corollary, which follows directly from Theorem 3.1 and Theorem 29.16 of Davidson (1994).

THEOREM 3.2. *Let X_{nt} be an m-vector-valued array and assume that for every m-vector λ of unit length there exists an array c_{nt} such that the conditions of Theorem* 3.1 *hold for $\lambda' X_{nt}$, all with respect to the same functions $K_n(\xi)$. Then $X_n \xrightarrow{d} X$, where X is a m-dimensional Gaussian process having a.s. continuous sample paths and independent increments.*

Implicit in the assumptions of Theorem 3.2 is the existence of a matrix of covariance functions, say, $\eta(\xi)$ $(m \times m)$, having the property that $\lambda'\eta(\xi)\lambda$ is a positive nondecreasing function on [0,1] for all λ of unit length. For example, such a case is given by $\eta(\xi) = \xi^{\beta}\Omega$ for a positive definite matrix Ω and $\beta > 0$. Having the variances trend at different rates is also clearly possible, although it is difficult to state a simple condition on η covering all the possible cases. Apart from this requirement, there should be no difficulty in meeting the conditions of Theorem 3.2 provided Theorem 3.1 holds for each element of the vector. Thus, Corollary 1 of De Jong (1997) shows that any constant array of the form $c_{nt} = t^{\beta} n^{-1/2-\gamma}$ for $\beta \leq \gamma$, with no restriction on signs, will satisfy Assumption 1(d). Clearly, in this case, Assumptions 1(b) and 1(c) are satisfied for all choices of λ by the maximum of the m array constants specified in the elementwise convergence. Condition (3.3) will hold likewise in this case, according to the earlier discussion.

4. WEAK CONVERGENCE TO STOCHASTIC INTEGRALS

Given vector-valued arrays $U_{ns}(p \times 1)$ and $W_{nt}(q \times 1)$, we next consider the convergence in distribution of sums of the type

$$G_n = \sum_{t=1}^{n-1} \sum_{s=1}^{t} U_{ns} W'_{n,t+1} \; (p \times q). \tag{4.1}$$

The case $W_{nt} = U_{nt}$, or more generally of the vectors having elements in common, is permitted in our approach. Letting $U_n(\xi) = \sum_{t=1}^{[n\xi]} U_{nt}$ and $W_n(\xi) = \sum_{t=1}^{[n\xi]} W_{nt}$, $(U_n, W_n) \xrightarrow{d} (U, W)$ under the conditions of Theorem 3.2. Note that with n large enough, $U_n(\xi)$ and $W_n(\xi)$ are arbitrarily well approximated by $\mathcal{F}_{n,-\infty}^{[n\xi]}$-measurable random variables under the NED assumption. Hence U and W, having independent increments, are martingales with respect to the same filtration. We seek conditions under which the weak limit of G_n, after centering, is the Itô integral $\int_0^1 U dW'$.

The only existing results of this type allowing serial correlation appear to be those of Phillips (1988b) and Hansen (1992). Phillips (1988b) assumes linear processes with i.i.d. innovations. Hansen (1992) assumes adapted strong-mixing processes. Our result contains these forms of dependence as special cases and,

in general, dominates them in terms of size conditions.[4] The assumptions also do not require U_{nt} and W_{nt} to be adapted sequences. These can depend on events of the infinitely far future, provided the dependence is damped at such a rate that the limit processes are martingales with respect to the same filtration. This result holds by the application of the same type of blocking argument that allows the CLT to be proved under dependence.

Our theorem holds under essentially the same conditions as specified previously for the FCLT. As before, it need not be the case that U and W are Brownian motion, but this property will hold for the leading case of $X_{nt} = n^{-1/2}X_t$ where $\lim_{n\to\infty} n^{-1}E(\sum_{t=1}^{n} X_t)(\sum_{t=1}^{n} X_t)' = \Omega$ (finite, positive semidefinite). In the theorem it is useful to specify the joint convergence of the triple $(U_n, W_n, G_n - \Lambda_n^{UW})$, so that the result may be used subsequently to construct the limiting distributions of the statistics familiar in unit root testing and cointegration theory, involving Brownian functionals, by applications of the continuous mapping theorem.

THEOREM 4.1. *Let the conditions of Theorem* 3.2 *hold for* $X_{nt} = (U_{nt}', W_{nt}')'$ *and* $K_n(\xi) = [n\xi] + 1$. *Then*

$$(U_n, W_n, G_n - \Lambda_n^{UW}) \xrightarrow{d} \left(U, W, \int_0^1 U dW'\right), \tag{4.2}$$

where U and W are a.s. continuous Gaussian processes having independent increments and

$$\Lambda_n^{UW} = \sum_{t=1}^{n} \sum_{s=t+1}^{n} EU_{nt} W_{ns}'. \tag{4.3}$$

To establish this result we adopt the approach of Chan and Wei (1988, Theorem 2.4(ii)). Under the stated conditions, $(U_n, W_n) \xrightarrow{d} (U, W)$. Therefore, by the Skorokhod representation theorem (Skorokhod, 1956) there exist random processes (U^n, W^n) having the same distribution as (U_n, W_n) but that converge almost surely to (U, W), and these processes are used to construct an approximation to the integral. Letting G^n denote the counterpart of G_n for these variables, the joint distribution of $(U^n, W^n, G^n - \Lambda_n^{UW})$ is the same as that specified in (4.2). To prove the theorem it is sufficient to show

$$G^n - \Lambda_n^{UW} \xrightarrow{p} \int_0^1 U\, dW', \tag{4.4}$$

because the joint convergence follows as a case of Theorem 29.16 of Davidson (1994) in which $G^n - \Lambda_n^{UW}$ is mapped into an (a.s. constant) element of $C[0,1]^{pq}$. Moreover, we can consider the case of scalar U_n and W_n without loss of generality, because the general case then follows by applying Theorem 30.14 of Davidson (1994).

The convergence in (4.4) is shown in two steps, using a blocking argument. Let k_n (the number of blocks) be a nondecreasing sequence such that $k_n \to \infty$ as $n \to \infty$ and $\lim_{n\to\infty} k_n(1/n + \delta_n^2) = 0$, where δ_n is the uniform distance between (U^n, W^n) and (U, W) except on a set of arbitrarily small probability, and let

$$G_n^* = \sum_{j=1}^{k_n} U_n(\xi_{j-1})(W_n(\xi_j) - W_n(\xi_{j-1})), \tag{4.5}$$

where $\xi_j = j/k_n$. Also, let G^{n*} denote the counterpart of G_n^* for (U^n, W^n). The first step, based on Chan and Wei (1988), is to show that

$$\left| G^{n*} - \int_0^1 U\,dW \right| \xrightarrow{p} 0. \tag{4.6}$$

This proof can follow that of Theorem 30.13 of Davidson (1994), line for line up to equation (30.78), with appropriate changes of notation.

Because the distributions of G_n^* and G^{n*} are the same, it suffices for the second step to show that

$$|G_n - G_n^* - \Lambda_n^{UW}| \xrightarrow{p} 0. \tag{4.7}$$

Noting that

$$G_n - G_n^* - \Lambda_n^{UW} = A_n - B_n, \tag{4.8}$$

where

$$A_n = \sum_{j=1}^{k_n} \sum_{t=n_{j-1}+1}^{n_j-1} \sum_{m=0}^{t-n_{j-1}-1} (U_{n,t-m} W_{n,t+1} - EU_{n,t-m} W_{n,t+1}) \tag{4.9}$$

and

$$B_n = \sum_{j=1}^{k_n} \sum_{t=n_{j-1}}^{n_j-1} \sum_{m=t-n_{j-1}}^{t-1} EU_{n,t-m} W_{n,t+1}, \tag{4.10}$$

where $n_j = [n\xi_j]$, for $j = 1, \ldots, k_n$, the proof of (4.7) is completed by showing $A_n \xrightarrow{p} 0$ and $B_n \to 0$. These arguments, which are fairly lengthy, are given in Appendix C.

5. CONCLUSION

This paper has given new weak convergence results that place the asymptotics underlying the theory of cointegrating regressions on virtually the same footing as standard asymptotics. We prove the FCLT under conditions similar to the best ones known to us for the ordinary CLT, from the point of view of the amount of dependence and heterogeneity permitted in the underlying random

processes. We also show that stochastic integral convergence holds under effectively the same conditions, something that has not been demonstrated previously to our knowledge.

NOTES

1. See Hansen (1992). Davidson's Theorem 30.13 (1994) is corrected in the 1997 reprint of the work by revising the conditions. The present paper shows that Davidson's original theorem is correct as stated, even though the proof contains an error.

2. In this paper, $\xrightarrow{d}$ denotes convergence in distribution, whereas $\xrightarrow{p}$ denotes convergence in probability.

3. Summations over an empty index set are defined as zero here.

4. It can be difficult to determine whether one set of conditions actually contains another. However, we note that if in the Phillips (1988b) model the linear MA coefficients are 1-summable (see Phillips, 1988b, p. 530), the process is L_2-NED on the i.i.d. forcing variables of size $-\frac{3}{2}$.

REFERENCES

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

Billingsley, P. (1968) *Convergence of Probability Measures.* New York: Wiley.

Chan, N.H. & C.Z. Wei (1988) Limiting distributions of least-squares estimates of unstable autoregressive processes. *Annals of Statistics* 16, 367–401.

Davidson, J. (1992) A central limit theorem for globally nonstationary near-epoch dependent functions of mixing processes. *Econometric Theory* 8, 313–334.

Davidson, J. (1993) The central limit theorem for globally nonstationary near-epoch dependent functions of mixing processes: The asymptotically degenerate case. *Econometric Theory* 9, 402–412.

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

Davidson, J. (2000) When Is a Time Series $I(0)$? Evaluating the Memory Properties of Nonlinear Dynamic Models. Working paper (revised).

Davidson, J. & R.M. De Jong (1997) Strong laws of large numbers for dependent heterogeneous processes: A synthesis of new and recent results. *Econometric Reviews* 16 (3), 251–279.

De Jong, R.M. (1997) Central limit theorems for dependent heterogeneous random variables. *Econometric Theory* 13, 353–367.

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

Hansen, B.E. (1992) Convergence to stochastic integrals for dependent heterogeneous processes. *Econometric Theory* 8, 489–500.

Jeganathan, P. (1991) On the asymptotic behavior of least-squares estimators in AR time series models with roots near the unit circle. *Econometric Theory* 7, 269–306.

Johansen, S. (1988) Statistical analysis of cointegrated vectors. *Journal of Economic Dynamics and Control* 12, 231–254.

Johansen, S. (1991) Estimation and hypothesis testing of cointegrated vectors in Gaussian vector autoregressive models. *Econometrica* 59, 1551–1580.

McLeish, D.L. (1975a) A maximal inequality and dependent strong laws. *Annals of Probability* 3, 829–839.

McLeish, D.L. (1975b) Invariance principles for dependent variables. *Zeitschrift für Wahrscheinlichskeitstheorie und Verwandte Gebiete* 33, 165–178.

McLeish, D.L. (1977) On the invariance principle for nonstationary mixingales. *Annals of Probability* 5, 616–621.

Park, J.Y. & P.C.B. Phillips (1988) Statistical inference in regressions with integrated processes, part 1. *Econometric Theory* 4, 468–497.

Park, J.Y. & P.C.B. Phillips (1989) Statistical inference in regressions with integrated processes, part 2. *Econometric Theory* 5, 95–131.

Phillips, P.C.B. (1986) Understanding spurious regressions in econometrics. *Journal of Econometrics* 33, 311–340.

Phillips, P.C.B. (1987) Time series regression with a unit root. *Econometrica* 55, 277–301.

Phillips, P.C.B. (1988a) Weak convergence to the stochastic integral $\int_0^1 BdB'$. *Journal of Multivariate Analysis* 24, 252–264.

Phillips, P.C.B. (1988b) Weak convergence of sample covariance matrices to stochastic integrals via martingale approximations. *Econometric Theory* 4, 528–533.

Phillips, P.C.B. & S.N. Durlauf (1986) Multiple time series regression with integrated processes. *Review of Economic Studies* LIII (4), 473–496.

Phillips, P.C.B. & V. Solo (1992) Asymptotics for linear processes. *Annals of Statistics* 20, 971–1001.

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

Skorokhod, A.V. (1956) Limit theorems for stochastic processes. *Theory of Probability and Its Applications* 1, 261–290.

Strasser, H. (1986) Martingale difference arrays and stochastic integrals. *Probability Theory and Related Fields* 72, 83–98.

Wooldridge, J.M. & H. White (1988) Some invariance principles and central limit theorems for dependent heterogeneous processes. *Econometric Theory* 4, 210–230.

APPENDIX A: TECHNICAL LEMMAS FOR MIXINGALES

An important tool for obtaining our results is the mixingale property. The L_2-mixingales were introduced by McLeish (1975a), and the extension to L_p-mixingales, $p \geq 1$, by Andrews (1988). Let $\mathcal{G}_{nt}$ denote an array of σ-fields, increasing in t for each n.

DEFINITION 2. *$\{X_{nt}, \mathcal{G}_{nt}\}$ is called an L_p-mixingale if for $m \geq 0$,*

$$\|X_{nt} - E(X_{nt}|\mathcal{G}_{n,t+m})\|_p \leq a_{nt}\psi(m+1), \tag{A.1}$$

$$\|E(X_{nt}|\mathcal{G}_{n,t-m})\|_p \leq a_{nt}\psi(m), \tag{A.2}$$

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

The notation here and in the rest of the paper is as in Davidson (1992, 1993) and De Jong (1997). The a_{nt} are referred to as the mixingale magnitude indices, and X_{nt} is called a mixingale of size $-\lambda$ if $\psi(m)$ is of size $-\lambda$.

Under integrability conditions, random variables that are NED on a mixing process are known to be mixingales, and in particular we have the following standard result (see, e.g., Davidson, 1994, Corollary 17.6).

LEMMA A.1. *If X_{nt} satisfies parts* (a)–(c) *of Assumption* 1, *$\{X_{nt}, \mathcal{F}_{n,-\infty}^t\}$ is an L_2-mixingale of size $-\frac{1}{2}$ with mixingale magnitude indices c_{nt}.*

Results of this kind are often used implicitly in the sequel, where we proceed by showing that certain functions of the variables are NED on V_{nt} and applying the same type of argument. Although the mixingale assumption is not structured enough to yield weak convergence results without supplementary conditions, because for example it may not be preserved under transformations, it is useful at certain stages of the proofs. In addition to the various known mixingale properties documented in sources such as Davidson (1994), we make use here of the following results.

LEMMA A.2. *If $\{Y_{nj}, \mathcal{F}_{nj}\}$ is an L_1-mixingale with magnitude indices a_{nj} and*

$$\limsup_{n\to\infty} \sum_{j=1}^{k_n} a_{nj} < \infty \tag{A.3}$$

and for all q,

$$\sum_{j=1}^{k_n} (E(Y_{nj}|\mathcal{F}_{n,j-q}) - E(Y_{nj}|\mathcal{F}_{n,j-q-1})) \xrightarrow{p} 0, \tag{A.4}$$

then

$$\sum_{j=1}^{k_n} Y_{nj} \xrightarrow{p} 0. \tag{A.5}$$

Proof. Note that for all m,

$$\begin{aligned}\sum_{j=1}^{k_n} Y_{nj} &= \sum_{j=1}^{k_n} (Y_{nj} - E(Y_{nj}|\mathcal{F}_{n,j+m})) \\ &\quad + \sum_{q=-m+1}^{m} \sum_{j=1}^{k_n} (E(Y_{nj}|\mathcal{F}_{n,j+q}) - E(Y_{nj}|\mathcal{F}_{n,j+q-1})) \\ &\quad + \sum_{j=1}^{k_n} E(Y_{nj}|\mathcal{F}_{n,j-m}).\end{aligned} \tag{A.6}$$

The L_1-norm of the first and third terms is bounded by $\sum_{j=1}^{k_n} a_{nj}\psi(m)$, which can be made arbitrarily small by selecting a large value of m. The second term converges in probability to zero by the requirement of equation (A.4). See also Andrews (1988). ■

LEMMA A.3. *Let $\{X_{nt}, \mathcal{G}_{nt}\}$ and $\{Y_{nt}, \mathcal{G}_{nt}\}$ be triangular L_2-mixingale arrays of size $-\frac{1}{2}$ with mixingale magnitude indices a_{nt}^X and a_{nt}^Y, respectively, where $\sum_{t=1}^{n}(a_{nt}^X)^2 = O(1)$ and $\sum_{t=1}^{n}(a_{nt}^Y)^2 = O(1)$. If $\gamma_n \geq 1$ is an increasing integer-valued function of n with $\gamma_n \to \infty$ as $n \to \infty$, then*

$$\lim_{n\to\infty} \sum_{t=1}^{n} \sum_{s=1}^{n} |E(X_{nt}Y_{ns})| I(|t-s| > \gamma_n) = 0. \tag{A.7}$$

Proof. This is analogous to Lemma 4 of De Jong (1997). ■

LEMMA A.4. *If $\{X_{nt}, \mathcal{G}_{nt}\}$ and $\{Y_{nt}, \mathcal{G}_{nt}\}$ are L_2-mixingales with mixingale numbers $\psi^X(j)$ and $\psi^Y(j)$ and magnitude indices a_{nt}^X and a_{nt}^Y, then*

$$\left\| \sum_{t=1}^{n} \sum_{s=1}^{t} X_{nt} Y_{ns} \right\|_1 \leq C \left(\sum_{t=1}^{n} (a_{nt}^X)^2 \sum_{j=1}^{\infty} (\log j)^2 \psi^X(j)^2 \right)^{1/2} \times \left(\sum_{s=1}^{n} (a_{ns}^Y)^2 \sum_{j=1}^{\infty} (\log j)^2 \psi^Y(j)^2 \right)^{1/2}. \tag{A.8}$$

for $0 < C < \infty$.

Proof. Define

$$X_{ntl} = E(X_{nt} | \mathcal{G}_{n,t-l}) - E(X_{nt} | \mathcal{G}_{n,t-l-1}) \tag{A.9}$$

and

$$Y_{nsi} = E(Y_{ns} | \mathcal{G}_{n,s-i}) - E(Y_{ns} | \mathcal{G}_{n,s-i-1}) \tag{A.10}$$

and note that

$$\begin{aligned} \left\| \sum_{t=1}^{n} \sum_{s=1}^{t} X_{nt} Y_{ns} \right\|_1 &= \left\| \sum_{l=-\infty}^{\infty} \sum_{i=-\infty}^{\infty} \sum_{t=1}^{n} \sum_{s=1}^{t} X_{ntl} Y_{nsi} \right\|_1 \\ &\leq \sum_{l=-\infty}^{\infty} \sum_{i=-\infty}^{\infty} \left\| \sum_{t=1}^{n} \sum_{s=1}^{t} X_{ntl} Y_{nsi} I(t-l > s-i) \right\|_1 \\ &\quad + \sum_{l=-\infty}^{\infty} \sum_{i=-\infty}^{\infty} \left\| \sum_{t=1}^{n} \sum_{s=1}^{t} X_{ntl} Y_{nsi} I(t-l < s-i) \right\|_1 \\ &\quad + \sum_{l=-\infty}^{\infty} \sum_{i=-\infty}^{\infty} \left\| \sum_{t=1}^{n} \sum_{s=1}^{t} X_{ntl} Y_{nsi} I(t-l = s-i) \right\|_1 . \end{aligned} \tag{A.11}$$

Consider each of these three sets of terms. Note first that the sequence

$$\sum_{s=1}^{t} X_{ntl} Y_{nsi} I(t-l > s-i), \qquad 1 \leq t \leq n \tag{A.12}$$

is a martingale difference with respect to the $\mathcal{G}_{n,t-l}$ and, therefore, for some constants $C_1 > 0$ and $C_2 > 0$,

$$\begin{aligned} &\sum_{l=-\infty}^{\infty} \sum_{i=-\infty}^{\infty} \left\| \sum_{t=1}^{n} \sum_{s=1}^{t} X_{ntl} Y_{nsi} I(t-l > s-i) \right\|_1 \\ &\quad \leq C_1 \sum_{l=-\infty}^{\infty} \sum_{i=-\infty}^{\infty} E \left(\sum_{t=1}^{n} X_{ntl}^2 \max_{1 \leq t \leq n} \left(\sum_{s=1}^{t} Y_{nsi} \right)^2 \right)^{1/2} \\ &\quad \leq C_1 \sum_{l=-\infty}^{\infty} \left(\sum_{t=1}^{n} E X_{ntl}^2 \right)^{1/2} \sum_{i=-\infty}^{\infty} \left(E \max_{1 \leq t \leq n} \left(\sum_{s=1}^{t} Y_{nsi} \right)^2 \right)^{1/2} \\ &\quad \leq C_2 \sum_{l=-\infty}^{\infty} \left(\sum_{t=1}^{n} E X_{ntl}^2 \right)^{1/2} \sum_{i=-\infty}^{\infty} \left(\sum_{s=1}^{n} E Y_{nsi}^2 \right)^{1/2}, \end{aligned} \tag{A.13}$$

where these inequalities are, respectively, by the Burkholder, Cauchy–Schwarz, and Doob inequalities. A similar argument holds for the second set of terms in (A.11), noting that

$$\sum_{t=s}^{n} X_{ntl} Y_{nsi} I(t-l<s-i), \qquad 1 \le s \le n$$

is a martingale difference with respect to $\mathcal{G}_{n,s-i}$ and also that for each s,

$$\left|\sum_{t=s}^{n} X_{ntl}\right| \le 2 \max_{1\le s\le n} \left|\sum_{t=1}^{s} X_{ntl}\right|. \tag{A.14}$$

Finally, we have

$$\begin{aligned}
&\sum_{l=-\infty}^{\infty} \sum_{i=-\infty}^{\infty} \left\| \sum_{t=1}^{n} \sum_{s=1}^{t} X_{ntl} Y_{nsi} I(t-l=s-i) \right\|_1 \\
&\quad \le C_3 \sum_{l=-\infty}^{\infty} \sum_{i=-\infty}^{\infty} \sum_{t=1}^{n} \|X_{ntl}\|_2 \|Y_{n,t-l+i,i} I(1 \le t-l+i \le n)\|_2 \\
&\quad \le C_3 \sum_{l=-\infty}^{\infty} \left(\sum_{t=1}^{n} EX_{ntl}^2 \right)^{1/2} \\
&\qquad \times \sup_{-\infty\le l\le\infty} \sum_{i=-\infty}^{\infty} \left(\sum_{t=1}^{n} EY_{n,t-l+i,i}^2 I(1 \le t-l+i \le n) \right)^{1/2}
\end{aligned} \tag{A.15}$$

for $C_3 > 0$.

The majorant of (A.11) does not exceed the sum of (A.15) and two terms of the form (A.13). Now, the lemma follows by combining the mixingale assumption with the fact that

$$\begin{aligned}
EX_{ntl}^2 &= EE(X_{nt}|\mathcal{G}_{n,t-l})^2 - EE(X_{nt}|\mathcal{G}_{n,t-l-1})^2 \\
&= E(X_{nt} - E(X_{nt}|\mathcal{G}_{n,t-l-1}))^2 - E(X_{nt} - E(X_{nt}|\mathcal{G}_{n,t-l}))^2,
\end{aligned} \tag{A.16}$$

with similar equalities for Y_{nsi}, and the fact that for a monotone decreasing sequence $\{x_j, j \ge 1\}$ the relation

$$\sum_{j=1}^{\infty} (x_j^2 - x_{j+1}^2)^{1/2} \le C \left(\sum_{j=1}^{\infty} (\log j)^2 x_j^2 \right)^{1/2}, \tag{A.17}$$

for $C > 0$, holds by an argument similar to McLeish (1975a, Theorem 1.6). ■

APPENDIX B: PROOF OF THEOREM 3.1

Fix $\delta > 0$ and let $\xi_{\delta j} = j\delta/2$ for $j = 0,\ldots,[2/\delta]$. For any pair ξ, ξ' such that $|\xi - \xi'| < \delta/2$, let $j_\delta(\xi,\xi')$ denote the maximal value of j such that $\xi > \xi_{\delta j}$ and $\xi' > \xi_{\delta j}$ and note that $0 < \xi - \xi_{\delta j_\delta(\xi,\xi')} < \delta$ and $0 < \xi' - \xi_{\delta j_\delta(\xi,\xi')} < \delta$. In addition, define

$$\nu_n^2(\xi,\delta) = \sum_{t=K_n(\xi)+1}^{K_n(\min(\xi+\delta,1))} c_{nt}^2 \tag{B.1}$$

and let $X_n(\xi) = X_n(1)$ if $\xi > 1$. Then

$$\begin{aligned}
&P\Big(\sup_{\xi\in[0,1]}\sup_{\{\xi':|\xi-\xi'|<\delta/2\}}|X_n(\xi)-X_n(\xi')|>\varepsilon\Big)\\
&\quad= P\Big(\sup_{\xi\in[0,1]}\sup_{\{\xi':|\xi-\xi'|<\delta/2\}}|X_n(\xi)-X_n(\xi_{\delta j_\delta(\xi,\xi')})+X_n(\xi_{\delta j_\delta(\xi,\xi')})-X_n(\xi')|>\varepsilon\Big)\\
&\quad\le 2P\Big(\max_{j=0,\ldots,[2/\delta]}\sup_{\{\xi:0<\xi-\xi_{\delta j}<\delta\}}|X_n(\xi)-X_n(\xi_{\delta j})|>\varepsilon/2\Big)\\
&\quad\le 2\sum_{j=0}^{[2/\delta]}P\Big(\sup_{\{\xi:0<\xi-\xi_{\delta j}<\delta\}}|X_n(\xi)-X_n(\xi_{\delta j})|>\varepsilon/2\Big)\\
&\quad= 2\sum_{j=0}^{[2/\delta]}P\Big(\Big(\sup_{\{\xi:0<\xi-\xi_{\delta j}<\delta\}}|X_n(\xi)-X_n(\xi_{\delta j})|\Big)^2\\
&\qquad\qquad\times I\Big(\Big(\sup_{\{\xi:0<\xi-\xi_{\delta j}<\delta\}}|X_n(\xi)-X_n(\xi_{\delta j})|\Big)^2>\varepsilon^2/4\Big)>\varepsilon^2/4\Big)\\
&\quad\le 2\sum_{j=0}^{[2/\delta]}4\varepsilon^{-2}E\Big(\sup_{\{\xi:0<\xi-\xi_{\delta j}<\delta\}}|X_n(\xi)-X_n(\xi_{\delta j})|\Big)^2\\
&\qquad\times I\Big(\Big(\sup_{\{\xi:0<\xi-\xi_{\delta j}<\delta\}}|X_n(\xi)-X_n(\xi_{\delta j})|\Big)^2>\varepsilon^2/4\Big)\\
&\quad= 2\sum_{j=0}^{[2/\delta]}\nu_n^2(\xi_{\delta j},\delta)4\varepsilon^{-2}E\Big(\sup_{\{\xi:0<\xi-\xi_{\delta j}<\delta\}}|X_n(\xi)-X_n(\xi_{\delta j})|/\nu_n(\xi_{\delta j},\delta)\Big)^2\\
&\qquad\times I\Big(\Big(\sup_{\{\xi:0<\xi-\xi_{\delta j}<\delta\}}|X_n(\xi)-X_n(\xi_{\delta j})|^2/\nu_n^2(\xi_{\delta j},\delta)>\varepsilon^2/(4\nu_n(\xi_{\delta j},\delta)^2)\Big)\Big)\\
&\quad\le 4\Big(\sum_{t=1}^{n}c_{nt}^2\Big)\max_{j=0,\ldots,[2/\delta]}4\varepsilon^{-2}E\Big(\sup_{\{\xi:0<\xi-\xi_{\delta j}<\delta\}}|X_n(\xi)-X_n(\xi_{\delta j})|/\nu_n(\xi_{\delta j},\delta)\Big)^2\\
&\qquad\times I\Big(\Big(\sup_{\{\xi:0<\xi-\xi_{\delta j}<\delta\}}|X_n(\xi)-X_n(\xi_{\delta j})|^2/\nu_n(\xi_{\delta j},\delta)^2\Big)>\varepsilon^2/(4\nu_n(\xi_{\delta j},\delta)^2)\Big)\\
&\quad\le C\varepsilon^{-2}\max_{j=0,\ldots,[2/\delta]}E\Big(\sup_{\{\xi:0<\xi-\xi_{\delta j}<\delta\}}|X_n(\xi)-X_n(\xi_{\delta j})|/\nu_n(\xi_{\delta j},\delta)\Big)^2\\
&\qquad\times I\Big(\Big(\sup_{\{\xi:0<\xi-\xi_{\delta j}<\delta\}}|X_n(\xi)-X_n(\xi_{\delta j})|^2/\nu_n(\xi_{\delta j},\delta)^2\Big)>\varepsilon^2\\
&\qquad\qquad\times\Big(4\max_{j=0,\ldots,[2/\delta]}\nu_n^2(\xi_{\delta j},\delta)\Big)^{-1}\Big)
\end{aligned}\tag{B.2}$$

for some finite constant $C > 0$. The second inequality follows from subadditivity, the third inequality is Markov's, and the remaining steps follow from the assumptions.

Next, note that by the mixingale property (Lemma A.1) and Corollary 16.14 of Davidson (1994), the sequence

$$Y_n(\delta, \xi') = \sup_{\{\xi : 0 < \xi - \xi' < \delta\}} |X_n(\xi) - X_n(\xi')| / \nu_n(\xi', \delta) \tag{B.3}$$

is uniformly square-integrable. Moreover, Assumption 1(b) implies that this property is independent of the segment of the data sequence represented by Y_n. (Compare McLeish, 1975b, Lemma 6.5; and McLeish, 1977, proof of Theorem 2.4.) In other words,

$$\begin{aligned}
&\limsup_{n\to\infty} \max_{j=0,\ldots,[2/\delta]} EY_n(\delta, \xi_{\delta j})^2 I(|Y_n(\delta, \xi_{\delta j})| > K) \\
&\quad = \max_{j=0,\ldots,[2/\delta]} \limsup_{n\to\infty} EY_n(\delta, \xi_{\delta j})^2 I(|Y_n(\delta, \xi_{\delta j})| > K) \\
&\quad = f(K),
\end{aligned} \tag{B.4}$$

where $f(K)$ does not depend on δ and $f(K) \to 0$ as $K \to \infty$.

Because δ is arbitrary, it follows by (3.5), (B.2), and (B.4) that $X_n(\xi)$ is stochastically equicontinuous on [0,1] if

$$\lim_{\delta\to 0} \limsup_{n\to\infty} \max_{j=0,\ldots,[2/\delta]} \nu_n^2(\xi_{\delta j}, \delta) = 0. \tag{B.5}$$

Because the max and the lim sup in equation (B.5) can similarly be interchanged, this holds by the assumption of equation (3.3).

Next, we show that $X(\xi)$ has independent increments. In view of the Gaussianity, it suffices to show that for any set $\{\xi_1, \ldots, \xi_k : 0 < \xi_1 < \xi_2 < \cdots < \xi_k < 1\}$ and all $i < j$, $X(\xi_i) - X(\xi_{i-1})$ and $X(\xi_j) - X(\xi_{j-1})$ are uncorrelated. This follows because

$$\begin{aligned}
&E(X(\xi_i) - X(\xi_{i-1}))(X(\xi_j) - X(\xi_{j-1})) \\
&\quad = \lim_{n\to\infty} E(X_n(\xi_i) - X_n(\xi_{i-1}))(X_n(\xi_j) - X_n(\xi_{j-1})),
\end{aligned} \tag{B.6}$$

where for any fixed $\delta > 0$,

$$\begin{aligned}
&|E(X_n(\xi_i) - X_n(\xi_{i-1}))(X_n(\xi_j) - X_n(\xi_{j-1}))| \\
&\quad \le \left| \sum_{t=K_n(\xi_{i-1})+1}^{K_n(\xi_i)} \sum_{s=K_n(\xi_{j-1})+1}^{K_n(\xi_{j-1}+\delta)} EX_{nt} X_{ns} \right| + \left| \sum_{t=K_n(\xi_{i-1})+1}^{K_n(\xi_i)} \sum_{s=K_n(\xi_{j-1}+\delta)+1}^{K_n(\xi_j)} EX_{nt} X_{ns} \right| \\
&\quad \le \left\| \sum_{t=K_n(\xi_{i-1})+1}^{K_n(\xi_i)} X_{nt} \right\|_2 \left\| \sum_{s=K_n(\xi_{j-1})+1}^{K_n(\xi_{j-1}+\delta)} X_{ns} \right\|_2 \\
&\qquad + \sum_{t=1}^{K_n(1)} \sum_{s=1}^{K_n(1)} |EX_{nt} X_{ns}| I(|s-t| \ge K_n(\xi_{j-1} + \delta) - K_n(\xi_i)).
\end{aligned} \tag{B.7}$$

Note that

$$\lim_{n\to\infty} \sum_{t=1}^{K_n(1)} \sum_{s=1}^{K_n(1)} |EX_{nt}X_{ns}| I(|s-t| \geq K_n(\xi_{j-1}+\delta) - K_n(\xi_i)) = 0 \tag{B.8}$$

for $\delta > 0$, by Lemma A.3 and the requirement that $K_n(\xi) - K_n(\xi') \to \infty$ for all $\xi > \xi'$, and also that

$$\lim_{\delta\to 0} \lim_{n\to\infty} \left\| \sum_{s=K_n(\xi_{j-1})+1}^{K_n(\xi_{j-1}+\delta)} X_{ns} \right\|_2 = 0$$

by the assumption in equation (3.3). Because δ is arbitrary in (B.8), it follows that

$$E(X(\xi_i) - X(\xi_{i-1}))(X(\xi_j) - X(\xi_{j-1})) = 0.$$

This completes the proof. ■

APPENDIX C: PROOF OF THEOREM 4.1

First, write

$$B_n = \sum_{j=1}^{k_n} \sum_{t=n_{j-1}}^{n_j-1} \sum_{m=t-n_{j-1}}^{t-1} EU_{n,t-m}W_{n,t+1}I(m \leq q_n)$$
$$+ \sum_{j=1}^{k_n} \sum_{t=n_{j-1}}^{n_j-1} \sum_{m=t-n_{j-1}}^{t-1} EU_{n,t-m}W_{n,t+1}I(m > q_n), \tag{C.1}$$

where q_n is a nondecreasing sequence such that $q_n \to \infty$ as $n \to \infty$. We define c_{nt}^U and c_{nt}^W as the constants with respect to which Assumption 1 holds for U_{nt} and W_{nt}, respectively. Note that because

$$\lim_{n\to\infty} \max_{1\leq t\leq n} \max\{(c_{nt}^U)^2, (c_{nt}^W)^2\} = 0 \tag{C.2}$$

by Assumption 1(d), it is possible to choose k_n and q_n such that

$$\lim_{n\to\infty} k_n q_n^2 \max_{1\leq t\leq n} \max\{(c_{nt}^U)^2, (c_{nt}^W)^2\} = 0. \tag{C.3}$$

The second term in (C.1) converges to zero, because its absolute value is bounded by

$$\sum_{t=1}^{n} \sum_{m=1}^{t-1} |EU_{n,t-m} W_{nt}| I(m > q_n) = o(1) \tag{C.4}$$

by Lemma A.3. For the first term in (C.1), note that

$$\begin{aligned}
&\sum_{j=1}^{k_n} \sum_{t=n_{j-1}}^{n_j-1} \sum_{m=t-n_{j-1}+1}^{t-1} |EU_{n,t-m} W_{n,t+1} I(m \le q_n)| \\
&\quad \le \sum_{j=1}^{k_n} \sum_{t=n_{j-1}}^{n_{j-1}+q_n} \sum_{m=t-n_{j-1}+1}^{t-1} |EU_{n,t-m} W_{n,t+1} I(m \le q_n)| \\
&\quad = O(k_n q_n^2 \max_{1\le t\le n} \max\{(c_{nt}^U)^2, (c_{nt}^W)^2\}) \\
&\quad = o(1)
\end{aligned} \tag{C.5}$$

by (C.3), noting that values of t exceeding $n_{j-1} + q_n$ contribute zero to the sum.

To show that $A_n \xrightarrow{p} 0$, first define

$$h(a,x) = xI(|x| \le a) + aI(x > a) - aI(x < -a) \tag{C.6}$$

and

$$g(a,x) = (x-a)I(x > a) + (x+a)I(x < -a) \tag{C.7}$$

and note that $x = g(a,x) + h(a,x)$. For some $K > 0$ to be chosen, define

$$\tilde{U}_{nt} = g(Kc_{nt}^U, U_{nt}) - Eg(Kc_{nt}^U, U_{nt}) \quad \text{and} \quad \tilde{U}_{nt}^m = E(\tilde{U}_{nt} | \mathcal{G}_{n,t-m}) \tag{C.8}$$

and

$$\bar{U}_{nt} = h(Kc_{nt}^U, U_{nt}) - Eh(Kc_{nt}^U, U_{nt}) \quad \text{and} \quad \bar{U}_{nt}^m = E(\bar{U}_{nt} | \mathcal{G}_{n,t-m}), \tag{C.9}$$

where $\mathcal{G}_{nt} = \sigma(V_{nt}, V_{n,t-1}, \ldots)$. Note that $U_{nt} = \tilde{U}_{nt} + \bar{U}_{nt}$. Also note that g and h are Lipschitz functions and therefore $\tilde{U}_{nt}$ and $\bar{U}_{nt}$ are L_2-NED on V_{nt} for all K, with NED magnitude indices c_{nt}^U and NED numbers $\nu(m)$. (See Davidson, 1994, Theorem 17.12.) Therefore, $\tilde{U}_{nt}$ is also an L_2-mixingale of size $-\frac{1}{2}$ with mixingale magnitude indices c_{nt}^U, implying that

$$\|\tilde{U}_{nt}^m\|_2 \le Cm^{-1/2-\mu} c_{nt}^U \tag{C.10}$$

for $\mu > 0$ and $C > 0$ and also, by Assumption 1(b), that

$$\|\tilde{U}_{nt}^m\|_2 \le c_{nt}^U \sup_{t,n} \|U_{nt}/c_{nt}^U I(|U_{nt}|/c_{nt}^U > K)\|_2 \le c_{nt}^U f(K) \tag{C.11}$$

for some $f(K)$ not depending on t or n, where $f(K) \to 0$ as $K \to \infty$. These inequalities further imply that

$$\|\tilde{U}_{nt}^m\|_2 \leq (Cm^{-1/2-\mu}c_{nt}^U)^{1-\mu}(c_{nt}^U f(K))^\mu = C'c_{nt}^U f(K)^\mu m^{-1/2-\mu/2+\mu^2} \tag{C.12}$$

for $C' > 0$. Therefore, $\tilde{U}_{nt}$ is an L_2-mixingale of size $-\frac{1}{2}$ with mixingale magnitude indices $c_{nt}^U f(K)^\mu$ for some small enough $\mu > 0$. Similarly, we may decompose W_{nt} into $\overline{W}_{nt}$ and $\widetilde{W}_{nt}$, having the same properties with respect to constants c_{nt}^W.

Note that

$$\begin{aligned} A_n = \sum_{j=1}^{k_n} \sum_{t=p_j}^{n_j-1} \sum_{s=p_j}^{t} (&\tilde{U}_{ns}\widetilde{W}_{n,t+1} - E\tilde{U}_{ns}\widetilde{W}_{n,t+1} + \tilde{U}_{ns}\overline{W}_{n,t+1} - E\tilde{U}_{ns}\overline{W}_{n,t+1} \\ &+ \bar{U}_{ns}\widetilde{W}_{n,t+1} - E\bar{U}_{ns}\widetilde{W}_{n,t+1} + \bar{U}_{ns}\overline{W}_{n,t+1} - E\bar{U}_{ns}\overline{W}_{n,t+1}), \end{aligned} \tag{C.13}$$

where for economy of notation we henceforth use the symbol p_j to denote $n_{j-1} + 1$. Consider the four sums of terms corresponding to this decomposition. It follows by Lemma A.4 that the L_1-norms of all these sums except those involving $\bar{U}_{ns}\overline{W}_{n,t+1}$ are of order

$$O\left(\sum_{j=1}^{k_n}\left(\sum_{s=p_j}^{n_j}(c_{ns}^U)^2 \sum_{t=p_j}^{n_j}(c_{nt}^W)^2\right)^{1/2} f(K)^\mu\right) = O(f(K)^\mu), \tag{C.14}$$

where the equality in (C.14) is by assumption. By choosing a large enough K, the limsups of the corresponding components of A_n can be made as small as desired. Accordingly, let the remaining component be defined as

$$\bar{A}_n = \sum_{j=1}^{k_n} \sum_{t=p_j}^{n_j-1} \sum_{s=p_j}^{t} \bar{U}_{ns}\overline{W}_{n,t+1} - E(\bar{U}_{ns}\overline{W}_{n,t+1}), \tag{C.15}$$

and we complete the proof by showing that for all $K > 0$, $\bar{A}_n \xrightarrow{p} 0$, by an application of Lemma A.2.

First write

$$\bar{A}_n = \sum_{j=1}^{k_n} Y_{nj}, \tag{C.16}$$

where

$$Y_{nj} = \sum_{t=p_j}^{n_j-1} \sum_{s=p_j}^{t} (\bar{U}_{ns}\overline{W}_{n,t+1} - E\bar{U}_{ns}\overline{W}_{n,t+1}). \tag{C.17}$$

Define $\mathcal{F}_{nj} = \sigma(V_{n,n_j}, V_{n,n_j-1}, \ldots)$ and $\mathcal{H}_{n,j-m}^{j+m} = \sigma(V_{n,n_{j-m-1}+1}, \ldots, V_{n,n_{j+m}})$ and for brevity of notation let E_{j-m}^{j+m} denote $E(\cdot|\mathcal{H}_{n,j-m}^{j+m})$. Then, note that for $m > 0$ there exist positive constants C_1, C_2, and C_3 such that

$$
\begin{aligned}
\|Y_{nj} - E_{j-m}^{j+m} Y_{nj}\|_1 &= \left\| \sum_{t=p_j}^{n_j-1} \sum_{s=p_j}^{t} (\bar{U}_{ns} \overline{W}_{n,t+1} - E_{j-m}^{j+m} \bar{U}_{ns} \overline{W}_{n,t+1}) \right\|_1 \\
&\le \left\| \sum_{t=p_j}^{n_j-1} \sum_{s=p_j}^{t} \bar{U}_{ns} (\overline{W}_{n,t+1} - E_{j-m}^{j+m} \overline{W}_{n,t+1}) \right\|_1 \\
&\quad + \left\| \sum_{t=p_j}^{n_j-1} \sum_{s=p_j}^{t} E_{j-m}^{j+m} \overline{W}_{n,t+1} (\bar{U}_{ns} - E_{j-m}^{j+m} \bar{U}_{ns}) \right\|_1 \\
&\quad + \left\| \sum_{t=p_j}^{n_j-1} \sum_{s=p_j}^{t} E_{j-m}^{j+m} \overline{W}_{n,t+1} E_{j-m}^{j+m} \bar{U}_{ns} - E_{j-m}^{j+m} \overline{W}_{n,t+1} \bar{U}_{ns} \right\|_1 \\
&\le \left\| \sum_{t=p_j}^{n_j-1} (\overline{W}_{n,t+1} - E_{j-m}^{j+m} \overline{W}_{n,t+1}) \sum_{s=p_j}^{t} \bar{U}_{ns} \right\|_1 \\
&\quad + \left\| \sum_{t=p_j}^{n_j-1} \overline{W}_{n,t+1} \sum_{s=p_j}^{t} (\bar{U}_{ns} - E_{j-m}^{j+m} \bar{U}_{ns}) \right\|_1 \\
&\quad + \left\| \sum_{t=p_j}^{n_j-1} E_{j-m}^{j+m} \overline{W}_{n,t+1} \sum_{s=p_j}^{t} (\bar{U}_{ns} - E_{j-m}^{j+m} \bar{U}_{ns}) \right\|_1 \\
&\le \sum_{t=p_j}^{n_j-1} \|\overline{W}_{n,t+1} - E_{j-m}^{j+m} \overline{W}_{n,t+1}\|_2 \left\| \sum_{s=p_j}^{t} \bar{U}_{ns} \right\|_2 \\
&\quad + \sum_{s=p_j}^{n_j-1} \|\bar{U}_{ns} - E_{j-m}^{j+m} \bar{U}_{ns}\|_2 \left(\left\| \sum_{t=p_j}^{n_j-1} \overline{W}_{n,t+1} \right\|_2 + \left\| \sum_{t=p_j}^{n_j-1} E_{j-m}^{j+m} \overline{W}_{n,t+1} \right\|_2 \right) \\
&\le C_1 \left(\sum_{t=p_j}^{n_j} c_{nt}^W \right) (mn/k_n)^{-1/2-\varepsilon} \left(\sum_{s=p_j}^{n_j} (c_{ns}^U)^2 \right)^{1/2} \\
&\quad + C_2 \left(\sum_{s=p_j}^{n_j} c_{ns}^U \right) (mn/k_n)^{-1/2-\varepsilon} \left(\sum_{t=p_j}^{n_j} (c_{nt}^W)^2 \right)^{1/2} \\
&\le C_3 m^{-1/2-\varepsilon} (n/k_n)^{-\varepsilon} \left(\sum_{s=p_j}^{n_j} (c_{ns}^U)^2 \sum_{t=p_j}^{n_j} (c_{nt}^W)^2 \right)^{1/2} \qquad \textbf{(C.18)}
\end{aligned}
$$

for some $\varepsilon > 0$. The first inequality follows from rearranging the terms and the norm inequality; the second inequality uses iterated expectations; the third is the Cauchy–Schwarz inequality and rearranging of terms; the fourth uses the NED definition, Theorem 1.6 of McLeish (1975a) (see also Davidson, 1994, Theorem 16.9), and the size

assumptions; and the fifth is obtained using Jensen's inequality. For the case $m = 0$, all except the two final steps of (C.18) hold, but for this case we have

$$\|Y_{nj} - E(Y_{nj}|\mathcal{H}^j_{n,j})\|_1 \leq C_4 \left(\sum_{s=p_j}^{n_j} (c^U_{ns})^2 \sum_{t=p_j}^{n_j} (c^W_{nt})^2 \right)^{1/2} \tag{C.19}$$

for $C_4 > 0$, using some of the same arguments as before. We have therefore established that Y_{nj} is L_1-NED of size $-\frac{1}{2}$, on a mixing process. Because it also possesses all its moments, it follows by Corollary 17.6 of Davidson (1994) that $\{Y_{nj}, \mathcal{F}_{nj}\}$ is also an L_1-mixingale of size $-\frac{1}{2}$, with respect to constants

$$a_{nj} = \left(\sum_{s=p_j}^{n_j} (c^U_{ns})^2 \sum_{t=p_j}^{n_j} (c^W_{nt})^2 \right)^{1/2}. \tag{C.20}$$

Note that $\limsup_{n\to\infty} \sum_{j=1}^{k_n} a_{nj} < \infty$, because by assumption,

$$\max \left\{ \limsup_{n\to\infty} \sum_{t=1}^{n} (c^U_{nt})^2, \limsup_{n\to\infty} \sum_{t=1}^{n} (c^W_{nt})^2 \right\} < \infty. \tag{C.21}$$

By Lemma A.2, the proof is therefore complete if we can show that for all q,

$$\sum_{j=1}^{k_n} (E(Y_{nj}|\mathcal{F}_{n,j-q}) - E(Y_{nj}|\mathcal{F}_{n,j-q-1})) \xrightarrow{p} 0. \tag{C.22}$$

We next write

$$\bar{U}_{nt} = (\bar{U}_{nt} - \bar{U}^m_{nt}) + \bar{U}^{-m}_{nt} + (\bar{U}^m_{nt} - \bar{U}^{-m}_{nt}), \tag{C.23}$$

and letting $\psi(m)$ denote the mixingale numbers relating to $\bar{U}_{nt}$, note that $\{\bar{U}_{nt} - \bar{U}^m_{nt}, \mathcal{G}_{nt}\}$ and $\{\bar{U}^{-m}_{nt}, \mathcal{G}_{nt}\}$ are L_2-mixingales with mixingale numbers equal to $\psi(m)$ for $l \leq m$ and $\psi(l)$ for $l > m$. Therefore, by Lemma A.4 and the assumptions,

$$\begin{aligned}
\limsup_{n\to\infty} \Bigg\| & \sum_{j=1}^{k_n} \sum_{t=p_j}^{n_j-1} \sum_{s=p_j}^{t} (E(\bar{U}_{ns}\bar{W}_{n,t+1}|\mathcal{F}_{n,j-q}) - E(\bar{U}_{ns}\bar{W}_{n,t+1}|\mathcal{F}_{n,j-q-1})) \\
& - \sum_{j=1}^{k_n} \sum_{t=p_j}^{n_j-1} \sum_{s=p_j}^{t} (E((\bar{U}^m_{ns} - \bar{U}^{-m}_{ns})\bar{W}_{n,t+1}|\mathcal{F}_{n,j-q}) \\
& - E((\bar{U}^m_{ns} - \bar{U}^{-m}_{ns})\bar{W}_{n,t+1}|\mathcal{F}_{n,j-q-1}))\Bigg\|_1 \\
& \leq C \left((\psi(m)^2 \sum_{l=1}^{m} (\log l)^2 + \sum_{l=m+1}^{\infty} \psi(l)^2 (\log(l))^2 \right)^{1/2} \\
& = O(m^{-\varepsilon})
\end{aligned} \tag{C.24}$$

for some $C > 0$ and $\varepsilon > 0$. Therefore by choosing m large enough, the difference between the expressions can be made negligible. A similar argument can be used to replace $\bar{W}_{n,t+1}$ by $\bar{W}^m_{n,t+1} - \bar{W}^{-m}_{n,t+1}$ in the last expression, and therefore it remains to show that for all q, K, and m,

$$\sum_{j=1}^{k_n}\sum_{t=p_j}^{n_j-1}\sum_{s=p_j}^{t}(E((\bar{U}_{ns}^{m}-\bar{U}_{ns}^{-m})(\overline{W}_{n,t+1}^{m}-\overline{W}_{n,t+1}^{-m})|\mathcal{F}_{n,j-q})$$

$$-E((\bar{U}_{ns}^{m}-\bar{U}_{ns}^{-m})(\overline{W}_{n,t+1}^{m}-\overline{W}_{n,t+1}^{-m})|\mathcal{F}_{n,j-q-1}))\xrightarrow{p}0. \tag{C.25}$$

Noting that

$$\bar{U}_{nt}^{m}-\bar{U}_{nt}^{-m}=\sum_{h=-m}^{m-1}(\bar{U}_{nt}^{h+1}-\bar{U}_{nt}^{h}) \tag{C.26}$$

and

$$\overline{W}_{n,t+1}^{m}-\overline{W}_{n,t+1}^{-m}=\sum_{h=-m}^{m-1}(\overline{W}_{n,t+1}^{h+1}-\overline{W}_{n,t+1}^{h}), \tag{C.27}$$

it follows that this result holds if for all q, K, h, and l,

$$\sum_{j=1}^{k_n}\sum_{t=p_j}^{n_j-1}\sum_{s=p_j}^{t}(E((\bar{U}_{ns}^{h}-\bar{U}_{ns}^{h+1})(\overline{W}_{n,t+1}^{l}-\overline{W}_{h,t+1}^{l+1})|\mathcal{F}_{n,j-q})$$

$$-E((\bar{U}_{ns}^{h}-\bar{U}_{ns}^{h+1})(\overline{W}_{n,t+1}^{l}-\overline{W}_{n,t+1}^{l+1})|\mathcal{F}_{n,j-q-1}))$$

$$\equiv\sum_{j=1}^{k_n}(E(Z_{nj}|\mathcal{F}_{n,j-q})-E(Z_{nj}|\mathcal{F}_{n,j-q-1}))$$

$$\xrightarrow{p}0. \tag{C.28}$$

Because the terms of (C.28) are uncorrelated, the latter statement is true if for all q, K, h, and l,

$$\lim_{n\to\infty}\sum_{j=1}^{k_n}EZ_{nj}^{2}=0. \tag{C.29}$$

However, note that

$$\sum_{j=1}^{k_n}EZ_{nj}^{2}=\sum_{j=1}^{k_n}\sum_{t_1=p_j}^{n_j-1}\sum_{t_2=p_j}^{n_j-1}\sum_{s_1=p_j}^{t_1}\sum_{s_2=p_j}^{t_2}E(\bar{U}_{ns_1}^{h_1}-\bar{U}_{ns_1}^{h_1+1})(\overline{W}_{n,t_1+1}^{l_1}-\overline{W}_{n,t_1+1}^{l_1+1})$$

$$\times(\bar{U}_{ns_2}^{h_2}-\bar{U}_{ns_2}^{h_2+1})(\overline{W}_{n,t_2+1}^{l_2}-\overline{W}_{n,t_2+1}^{l_2+1}). \tag{C.30}$$

Consider, as representative, the terms for which

$$s_1-h_1\le t_1+1-l_1\le s_2-h_2\le t_2+1-l_2. \tag{C.31}$$

The other cases are treated identically. First, note that by the martingale difference property of the four terms in equation (C.30), the terms in that equation are zero unless $s_2-h_2=t_2+1-l_2$. Therefore, for the terms that satisfy the preceding restriction, we have, applying Lemma A.4 once again,

$$
\begin{aligned}
&\Bigg\| \sum_{j=1}^{k_n} \sum_{t_1=p_j}^{n_j-1} \sum_{s_1=p_j}^{t_1} (\bar{U}_{ns_1}^{h_1} - \bar{U}_{ns_1}^{h_1+1})(\bar{W}_{n,t_1+1}^{l_1} - \bar{W}_{n,t_1+1}^{l_1+1}) \\
&\quad \times \sum_{t_2=p_j}^{n_j-1} I(1 \le t_2 + 1 - l_2 + h_2 \le n)(\bar{U}_{n,t_2-l_2+h_2}^{h_2} - \bar{U}_{n,t_2-l_2+h_2}^{h_2+1})(\bar{W}_{n,t_2+1}^{l_2} - \bar{W}_{n,t_2+1}^{l_2+1}) \Bigg\|_1 \\
&\qquad \le \sum_{j=1}^{k_n} \Bigg\| \sum_{t_1=p_j}^{n_j-1} \sum_{s_1=p_j}^{t_1} (\bar{U}_{ns_1}^{h_1} - \bar{U}_{ns_1}^{h_1+1})(\bar{W}_{n,t_1+1}^{l_1} - \bar{W}_{n,t_1+1}^{l_1+1}) \Bigg\|_1 \\
&\qquad\quad \times \Bigg\| \sum_{t_2=p_j}^{n_j-1} I(1 \le t_2 + 1 - l_2 + h_2 \le n)(\bar{U}_{n,t_2+1-l_2+h_2}^{h_2} - \bar{U}_{n,t_2-l_2+h_2}^{h_2+1}) \\
&\qquad\qquad \times (\bar{W}_{n,t_2+1}^{l_2} - \bar{W}_{n,t_2+1}^{l_2+1}) \Bigg\|_\infty \\
&\qquad = O\Bigg(\sum_{j=1}^{k_n} \Bigg(\sum_{s=p_j}^{n_j} (c_{ns}^U)^2 \sum_{t=p_j}^{n_j} (c_{nt}^W)^2 \Bigg)^{1/2} \\
&\qquad\qquad \times \sum_{t_2=p_j}^{n_j} I(1 \le t_2 + 1 - l_2 + h_2 \le n) c_{nt_2}^W c_{n,t_2+1-l_2+h_2}^U \Bigg) \\
&\qquad = O\Bigg(\sum_{j=1}^{k_n} \sum_{s=p_j}^{n_j} (c_{ns}^U)^2 \sum_{t=p_j}^{n_j} (c_{nt}^W)^2 \Bigg) \\
&\qquad = O\Bigg(\max_{1 \le j \le k_n} \Bigg(\sum_{s=p_j}^{n_j} (c_{ns}^U)^2 \sum_{t=p_j}^{n_j} (c_{nt}^W)^2 \Bigg)^{1/2} \Bigg) \\
&\qquad = o(1), \qquad \textbf{(C.32)}
\end{aligned}
$$

where the last equality follows from the assumption of equation (3.3). This completes the proof. ■