COWLES FOUNDATION FOR RESEARCH IN ECONOMICS

AT YALE UNIVERSITY

Box 2125, Yale Station

New Haven, Connecticut 06520

COWLES FOUNDATION DISCUSSION PAPER NO. 796

WEAK CONVERGENCE TO THE MATRIX STOCHASTIC INTEGRAL $\int_0^1 BdB'$

by

Peter C. B. Phillips

July 1986

WEAK CONVERGENCE TO THE MATRIX STOCHASTIC INTEGRAL $\int_0^1 BdB'$

by

P.C.B. Phillips*

Cowles Foundation for Research in Economics

Yale University

ABSTRACT

The asymptotic theory of regression with integrated processes of the ARIMA type frequently involves weak convergence to stochastic integrals of the form $\int_0^1 WdW$, where $W(r)$ is standard Brownian motion. In multiple regressions and vector autoregressions with vector ARIMA processes the theory involves weak convergence to matrix stochastic integrals of the form $\int_0^1 BdB'$, where $B(r)$ is vector Brownian motion with non scalar covariance matrix. This paper studies the weak convergence of sample covariance matrices to $\int_0^1 BdB'$ under quite general conditions. The theory is applied to vector autoregressions with integrated processes.

KEY WORDS: Integrated process; Invariance Principle; Near integrated time series; Stochastic integral; vector autoregression; Weak convergence.

June, 1986

* This paper was written while the author was a Monash University Visiting Professor. My thanks go to the members of the Department of Econometrics and Operations Research for their hospitality and to Monash University for its support during this visit.

1. INTRODUCTION

Let $\{y_t\}_0^\infty$ be a multiple (n-vector) time series generated by:

(1) $\quad y_t = Ay_{t-1} + u_t \quad ; \quad t = 1, 2, \ldots$

(2) $\quad A = I_n \quad ;$

(3) $\quad y_0$ = random with a certain fixed distribution.

Under very general conditions on the sequence of innovations, $\{u_t\}_1^\infty$, in (1), $t^{-1/2} y_t$ converges almost surely to standardized vector Brownian motion $t^{-1/2} B(t)$ on $C^n[0,\infty]$. The covariance matrix, Ω, of $B(t)$ depends on the serial covariance properties of $\{u_t\}_1^\infty$. If the sequence $\{u_t\}_1^\infty$ is stationary with spectral density matrix $f_{uu}(\lambda)$ satisfying $f_{uu}(0) > 0$ (">" here signifies positive definite) then $\Omega = 2\pi f_{uu}(0)$. Strong invariance principles of this type have been proved recently be Berkes and Philipp (1981) and by Eberlain (1986).

Weak invariance principles follow directly from these strong convergence results as shown by Philipp and Stout (1975). In this case, it is usual to define the partial sum process $S_t = \Sigma_1^t u_j$ and construct the following random element of $D^n[0,1]$:

$$X_T(r) = T^{-1/2} S_{[Tr]} = T^{-1/2} S_{j-1}; \quad (j-1)/T < r < j/T$$

Then as $T \uparrow \infty$

(4) $\quad X_T(r) \underset{d}{\rightarrow} B(r)$

where $B(r)$ is vector Brownian motion on $C^n[0,1]$ with covariance matrix Ω. In (4) we use the symbol "$\underset{d}{\rightarrow}$" to signify weak convergence of the associated probability measures. Billingsley (1968) provides an extensive discussion of such weak invariance principles in the scalar case ($n=1$) and gives many useful applications.

One major time series application of (4) is to the theory of regression for integrated processes. If $\{u_t\}_1^\infty$ is generated by a linear process such as a finite order stationary and invertible vector ARMA model then y_t is known as an integrated process of order one (Box and Jenkins (1976)). We are often interested in the asymptotic behaviour of statistics from linear least squares regressions with integrated processes. Thus, from the first order vector autoregression of y_t on y_{t-1} in (1) we obtain the regression coefficient matrix

$$\hat{A} = (\sum_1^T y_t y_{t-1}')(\sum_1^T y_{t-1}y_{t-1}')^{-1} .$$

Here, $\hat{A}$ is a simple function of the sample moments of y_t. To the extent that y_t behaves asymptotically like vector Brownian motion, we might expect the asymptotic behaviour of $\hat{A}$ to be described by a corresponding functional of Brownian motion.

To be more precise consider standardized deviations of $\hat{A}$ about I_n:

$$(5) \qquad T(\hat{A} - I) = (T^{-1} \sum_1^T u_t y_{t-1}')(T^{-2} \sum_1^T y_{t-1}y_{t-1}')^{-1} .$$

By simple calculations we may write the sample second moment $T^{-2} \sum_1^T y_{t-1}y_{t-1}'$ as a quadratic functional of the random element $X_T(r)$, at least up to a term of $o_p(1)$. That is,

$$T^{-2} \sum_1^T y_{t-1}y_{t-1}' \; = \; \int_0^1 X_T(r)X_T(r)'dr + o_p(1)$$

Result (4) and the continuous mapping theorem then establish that

$$(6) \qquad T^{-2}\sum_{1}^{T} y_{t-1}y'_{t-1} \underset{d}{\rightarrow} \int_0^1 B(r)\, B(r)'dr$$

as $T \uparrow \infty$.

In a similar way, we might expect that

$$(7) \qquad T^{-1}\sum_{1}^{T} u_t y'_{t-1} \underset{d}{\rightarrow} \int_0^1 dB(r)\, B(r)'$$

at least when $\{u_t\}_1^\infty$ is a sequence of square integrable martingale differences. However, unlike (6), (7) cannot be obtained by a simple application of (4) and the continuous mapping theorem. The reason is that we cannot write the sample covariance matrix $T^{-1}\sum_1^T u_t y_{t-1}$ as a continuous functional of the random element $X_T(r)$. Moreover, the limit process $\int_0^1 dBB'$ (we shall sometimes suppress the argument of integration in integrals of this type) is a matrix stochastic integral and, since $B(r)$ is almost surely (vector Wiener measure) of unbounded variation, $\int_0^1 dBB'$ cannot be considered as the (mean square) limit of a Riemann Stieltjes sum. Furthermore, when the innovations u_t are not martingale differences $E(u_t y'_{t-1}) \neq 0$, in general, and there is no reason to expect (6) to hold.

In the scalar case ($n = 1$, $A = a$, $\Omega = \omega^2$) the problems described in the previous paragraph are easily resolved. We simply write

$$S_T^2 = \sum_1^T u_t^2 + 2\sum_2^T \left(\sum_1^{t-1} u_s\right)u_t$$

and, then, under quite general conditions, as $T \uparrow \infty$

$$(8) \quad T^{-1}\sum_1^T y_{t-1}u_t = \tfrac{1}{2}\left\{(T^{-1/2}X_T(1))^2 - T^{-1}\sum_1^T u_t^2\right\} + o_p(1) \underset{d}{\rightarrow} (1/2)\{\omega^2 W(1)^2 - \omega_0^2\}$$

where $W(r)$ denotes standard Brownian motion on $C[0,1]$ and where $\omega_0^2 = \lim_{T\to\infty} T^{-1} \sum_1^T E(u_t^2)$. Here $T^{-1} \sum_1^T u_t^2 \to \omega_0^2$ a.s. by a suitable strong law for weakly dependent time series (e.g. McLeish (1975)). In view of the formula $\int_0^1 WdW = (1/2)(W(1)^2 - 1)$ and the fact that $\omega W(r) \equiv B(r)$ (here the symbol "$\equiv$" signifies "has the same distribution as") we deduce that:

$$T^{-1} \sum_1^T y_{t-1}u_t \underset{d}{\to} \int_0^1 BdB + (1/2)(\omega^2 - \omega_0^2) \tag{9}$$

This reduces to the formula suggested above in (7) only when $\omega^2 = \omega_0^2$.

Thus, in the scalar case, we obtain the following limit law for the autoregressive coefficient:

$$T(\hat{a} - 1) \underset{d}{\to} \{\int_0^1 BdB + (1/2)(\omega^2 - \omega_0^2)\}/\{\int_0^1 B(r)^2dr\} \tag{10}$$

(10) is proved in Phillips (1986a) and it generalizes the simple formula, $\int_0^1 BdB/\int_0^1 B^2dr$, that was first suggested by White (1958) for the case where the innovation sequence is iid $N(0,\omega^2)$.

When $n > 1$ the argument that was used above to deduce (9) no longer applies. In fact, partial summation of the outer product S_TS_T' yields:

$$S_TS_T' = \sum_1^T u_tu_t' + \sum_2^{T-1} (\sum_1^{t-1} u_s)u_t' + \sum_2^{t-1} u_t(\sum_1^{t-1} u_s)'$$

so that, in place of (8), we now obtain:

$$\begin{aligned} T^{-1} \sum_1^T (y_{t-1}u_t' + u_ty_{t-1}') &= X_T(1)X_T(1)' - T^{-1} \sum_1^T u_tu_t' + o_p(1) \\ &\underset{d}{\to} B(1)B(1)' - \Omega_0 \end{aligned} \tag{11}$$

where $\Omega_0 = \lim_{T\to\infty} T^{-1} \sum_1^T E(u_t u_t')$. Determination of the limit law of the matrix $T^{-1} \sum_1^T y_{t-1} u_t'$ is not possible from (11), although the joint limiting distribution of its diagonal elements may be deduced. However, the latter is insufficient for many problems of central interest, such as the limiting distribution of the regression coefficients (5).

The purpose of the present paper is to obtain the matrix analogue of (9) directly. Our approach permits a wide class of possible innovation sequences and our main result is directly applicable to the study of regression statistics such as (5). It should also be useful in other contexts where weak convergence to the matrix stochastic integral $\int_0^1 BdB'$ is needed.

2. MAIN RESULTS

We shall require $\{u_t\}_1^\infty$ to satisfy conditions which are sufficient to ensure the validity of (4). In particular, we impose:

ASSUMPTION 2.1

(a) $E(u_t) = 0$ all t;

(b) $\sup_{i,t} E|u_{it}|^{\beta+\varepsilon} < \infty$ for some $\beta > 2$ and $\varepsilon > 0$;

(c) $\Omega = \lim_{T\to\infty} T^{-1} E(S_T S_T')$ exists and is positive definite;

(d) $\{u_t\}_1^\infty$ is strong mixing with mixing numbers α_m that satisfy

$$\sum_1^\infty \alpha_m^{1-2/\beta} < \infty \tag{12}$$

If $\{u_t\}_1^\infty$ is weakly stationary then (c) is, in fact, implied by the mixing condition (d) (theorem 18.5.3 of Ibragimov and Linnik (1971)). In this case, we obtain:

(13) $$\Omega = E(u_1u_1') + \sum_{k=2}^{\infty} E(u_1u_k') + \sum_{k=2}^{\infty} E(u_ku_1')$$

$$= \Omega_0 + \Omega_1 + \Omega_1', \quad \text{say.}$$

Under Assumption 2.1 we have (Herrndorf (1984), Phillips (1986b)):

LEMMA 2.2 *If* $\{u_t\}_1^{\infty}$ *is a sequence of random n-vectors that satisfy Assumption 2.1 then as* $T \to \infty$ $X_T(r) \xrightarrow{d} B(r)$, *vector Brownian motion with covariance matrix* Ω.

It is convenient to introduce a multiple ($n\times 1$) time series $\{z_t(x)\}_1^{\infty}$ generated by the model

(14) $$z_t(x) = Fz_{t-1}(x) + u_t \quad ; \quad t = 1, 2, \ldots$$

(15) $$F = \exp\{(x/T)G\}$$

(16) $$z_0(x) = y_0.$$

Here x is a scalar and G is an arbitrary $n\times n$ matrix. When $x = 0$ the model is equivalent to (1) - (3). Note that as $T \to \infty$, $F \to I_n$ so that for fixed $x \neq 0$ $z_t(x)$ behaves, at least asymptotically, like an integrated process. Such processes were introduced in Phillips (1986c) and were called "near integrated time series".

Back substitution in (14) yields

(17) $$z_t(x) = \sum_{j=1}^{t} \exp\{((t-j)x/T)G\}u_j + \exp\{(tx/T)G\}y_0$$

Define

$$(18)\qquad \dot{z}_t = (d/dx)z_t(x) = G\sum_1^t \exp\{((t-j)x/T)G\}((t-j)/T)u_j$$
$$+\ G(t/T)\exp\{(tx/T)G\}y_0$$

We now consider the asymptotic behaviour of sample moments of these processes.

<u>LEMMA 2.3</u> <u>If</u> $\{u_t\}_1^\infty$ <u>satisfies Assumption 2.1 and</u> $\{z_t(x)\}_1^\infty$ <u>is a near integrated time series generated by</u> (14) - (16) <u>then as</u> $T \to \infty$

$$(a)\qquad T^{-1}\dot{z}_T(x)z_T(x) \underset{d}{\to} GL_G(1,x)K_G(1,x)';$$

$$(b)\qquad T^{-2}\sum_1^T \dot{z}_t(x)z_t(x)' \underset{d}{\to} G\int_0^1 L_G(r,x)K_G(r,x)'dr;$$

$$(c)\qquad T^{-2}\sum_1^T z_t(x)z_t(x)' \underset{d}{\to} \int_0^1 K_G(r,x)K_G(r,x)'dr$$

<u>where</u>

$$K_G(r,x) = \int_0^r \exp\{(r-s)xG\}dB(s),$$

$$L_G(r,x) = \int_0^r \exp\{(r-s)xG\}(r-s)dB(s).$$

We also need:

<u>LEMMA 2.4</u> <u>If</u> $B(r)$ <u>is vector Brownian motion with covariance matrix</u> Ω <u>and</u> $J_C(r) = \int_0^r \exp\{(r-s)C\}dB(s)$ <u>then</u>

$$(19)\qquad J_C(1)J_C(1)' = \Omega + C\int_0^1 J_C(r)J_C(r)'dr + \int_0^1 J_C(r)J_C(r)'drC'$$
$$+ \int_0^1 J_C(r)dB(r)' + \int_0^1 dB(r)J_C(r)'$$

<u>for any</u> $n\times n$ <u>matrix</u> C.

LEMMA 2.5 *If* $\{u_t\}_1^\infty$ *satisfies Assumption* 2.1 *and* $\{z_t(x)\}_1^\infty$ *is a near integrated process generated by* (14) - (16) *then as* $T \to \infty$:

(a) $$T^{-1}\sum_1^T \{\dot{z}_{t-1}(x)u_t' + u_t\dot{z}_{t-1}(x)'\} \to_d G\int_0^1 L_G(r,x)dB(r)' + \int_0^1 [dB(r)L_G(r,x)']G'$$

(b) $$T^{-1}\sum_1^T \dot{z}_{t-1}(x)u_t' \to_d G\int_0^1 L_G(r,x)dB(r)'$$

(c) $$T^{-1}\sum_1^T z_{t-1}(h)u_t' - T^{-1}\sum_1^T y_{t-1}u_t' \to_d \int_0^1 K_G(r,h)dB(r)' - \int_0^1 B(r)dB(r)' .$$

We are now in a position to establish our main result:

THEOREM 2.6 *If* $\{u_t\}_1^\infty$ *is weakly stationary and satisfies Assumption* 2.1 *and if* $\{y_t\}_0^\infty$ *is generated by* (1) - (3), *then as* $T \to \infty$:

(a) $$T^{-1}\sum_1^T y_{t-1}u_t' \to_d \int_0^1 B(r)dB(r)' + \Omega_1,$$

(b) $$T^{-1}\sum_1^T z_{t-1}(h)u_t' \to_d \int_0^1 K_G(r,h)dB(r)' + \Omega_1$$

where

$$\Omega_1 = \lim_{T\to\infty} T^{-1}\sum_1^T E(y_{t-1}u_t') = \sum_{k=2}^\infty E(u_1u_k') .$$

COROLLARY 2.7 *If* $\{u_t\}_1^\infty$ *is a sequence of stationary martingale differences that satisfy Assumption* 2.1 *and if* $\{y_t\}_0^\infty$ *is generated by* (1) - (3) *then as* $T \to \infty$

$$T^{-1}\sum_1^T y_{t-1}u_t' \to_d \int_0^1 B(r)d\,B(r)' .$$

Theorem 2.6 may be extended to include sequences $\{u_t\}_1^{\infty}$ which are not weakly stationary with some strengthening of the moment and mixing conditions (b) and (d) of Assumption 2.1. The details are not given here since the case of predominant interest is that of weakly stationary innovations in (1).

We may now deduce the relevant asymptotics for regression statistics such as (5). In particular, we have:

THEOREM 2.8 If the conditions of theorem 2.6 hold then as $T \to \infty$:

$$T(\hat{A} - I) \underset{d}{\to} \{\int_0^1 B(r)dB(r)' + \Omega_1\}'\{\int_0^1 B(r)B(r)'dr\}^{-1} \tag{20}$$

Note that in the scalar case (setting $\Omega_1 = \omega_1$) we have $\omega^2 = \omega_0^2 + 2\omega_1$, so that (20) reduces, as we would expect, to the earlier formula (10).

3. PROOFS

Proof of Lemma 2.2 See Herrndorf (1984) for the case $n = 1$ and Phillips (1986b) for the case $n > 1$.

Proof of Lemma 2.3 To prove (a) we note that

$$\begin{aligned}
T^{-1/2} z_T(x) &= \sum_{j=1}^{T} \exp\{((1-j/T)x)G\} \int_{(j-1)/T}^{j/T} dX_T(s) + O_p(T^{-1/2}) \\
&= \sum_{j=1}^{T} \int_{(j-1)/T}^{j/T} \exp\{(1-s)xG\}dX_T(s) + O_p(T^{-1/2}) \\
&= \int_0^1 \exp\{(1-s)xG\}dX_T(s) + O_p(T^{-1/2}) \\
&\underset{d}{\to} \int_0^1 \exp\{(1-s)xG\}dB(s) \qquad \text{as } T \to \infty
\end{aligned}$$

in view of Lemma 2.2 and the continuous mapping theorem. In a similar way we find that

$$T^{-1/2}\dot{z}_t(x) \underset{d}{\to} G\int_0^1 \exp\{(1-s)xG\}(1-s)dB(s)$$

and result (a) follows directly. To prove (b) we write

$$\begin{aligned}
T^{-2}\sum_1^T \dot{z}_t z_t' &= T^{-2}\sum_{i=1}^T \Big[G\sum_{j=1}^i \exp\{((i-j)/T)xG\}((i-j)/T)u_j\Big] \\
&\qquad \Big[\sum_{k=1}^i u_k' \exp\{((i-k)/T)xG'\}\Big] + o_p(T^{-1/2}) \\
&= \sum_{i=1}^T \int_{(i-1)/T}^{i/T} dr\Big[G\sum_{j=1}^i \int_{(j-1)/T}^{j/T} \exp\{(r-s)xG\}(r-s)dX_T(s)\Big] \\
&\qquad \cdot\Big[\sum_{k=1}^i \int_{(k-1)/T}^{k/T} dX_T(t)' \exp\{(r-t)xG'\}\Big] + o_p(T^{-1/2}) \\
&= \int_0^1\int_0^r\int_0^r G\exp\{(r-s)xG\}(r-s)dX_T(s)dX_T(t)' \exp\{(r-t)xG'\}dr + o_p(T^{-1/2}) \\
&\underset{d}{\to} \int_0^1 \Big[G\int_0^r \exp\{(r-s)xG\}(r-s)dB(s)\Big]\Big[\int_0^r dB(t)' \exp\{(r-t)xG'\}\Big]dr \\
&= G\int_0^1 L_G(r,x)K_G(r,x)'dr
\end{aligned}$$

as required. Part (c) follows in a similar fashion.

<u>Proof of Lemma 2.4</u> First define $\xi(r) = \int_0^r \exp\{-sC)dB(s)$ and note that $J_C(r) = \exp(rC)\xi(r)$. Now by the multivariate Ito formula for stochastic differentiation we have:

$$d\{\xi(r)\xi(r)'\} = d\xi(r)\xi(r)' + \xi(r)d\xi(r)' + \exp\{-rC)\Omega\exp\{-rC')dr\ .$$

Hence

$$\begin{aligned}
&\int_0^1 \big[\exp(rC)d\{\xi(r)\xi(r)'\}\exp(rC')\big] \\
&= \int_0^1 dB(r)J_C(r)' + \int_0^1 J_C(r)dB(r)' + \Omega
\end{aligned}$$

leading to the result as stated.

Proof of Lemma 2.5 From (14) we obtain

$$z_t(x)z_t(x)' - z_{t-1}(x)z_{t-1}(x)' = (xG)T^{-1}z_{t-1}(x)z_{t-1}(x)'$$

$$+ T^{-1}z_{t-1}(x)z_{t-1}(x)'(xG') + z_{t-1}(x)u_t'$$

$$+ u_t z_{t-1}(x)' + u_t u_t' + 0_p(T^{-1})$$

and averaging over t we find

$$T^{-1}z_T(x)z_T(x)' = (xG)T^{-2}\sum_1^T z_{t-1}(x)z_{t-1}(x)'$$

$$+ T^{-2}\sum_1^T z_{t-1}(x)z_{t-1}(x)'(xG) + T^{-1}\sum_1^T z_{t-1}(x)u_t'$$

$$+ T^{-1}\sum_1^T u_t z_{t-1}(x)' + T^{-1}\sum_1^T u_t u_t' + 0_p(T^{-1})$$

Differentiating with respect to x yields:

$$(21)\qquad T^{-1}\dot{z}_T z_T' + T^{-1}z_T\dot{z}_T' = xGT^{-2}\sum_1^T(\dot{z}_{t-1}z_{t-1}' + z_{t-1}\dot{z}_{t-1}')$$

$$+ T^{-2}\sum_1^T(\dot{z}_{t-1}z_{t-1}' + z_{t-1}\dot{z}_{t-1}')(xG') + G(T^{-2}\sum_1^T z_{t-1}z_{t-1}')$$

$$+ (T^{-2}\sum_1^T z_{t-1}z_{t-1}')G' + T^{-1}\sum_1^T(\dot{z}_{t-1}u_t' + u_t\dot{z}_{t-1}') + 0_p(T^{-1})$$

From Lemma 2.3 and (21) we now deduce that

$$(22)\qquad T^{-1}\sum_1^T(\dot{z}_{t-1}u_t' + u_t\dot{z}_{t-1}') \underset{d}{\rightarrow} GL_G(1,x)K_G(1,x)'$$

$$+ K_G(1,x)L_G(1,x)'G' - xG\{G\int_0^1 L_G(r,x)K_G(r,x)'dr$$

$$+ \int_0^1 K_G(r,x)L_G(r,x)'drG'\} - \{G\int_0^1 L_G(r,x)K_G(r,x)'dx$$

$$+ \int_0^1 K_G(r,x)L_G(r,x)'drG'\}xG - G\int_0^1 K_G(r,x)K_G(r,x)'dr$$

$$- \int_0^1 K_G(r,x)K_G(r,x)'drG'$$

Now let $C = xG$ in (19) and differentiating (19) we have (noting that $J_{xG}(r) = K_G(r,x)$ and $(d/dx)J_{xG}(r) = GL_G(r,x)$):

$$(23) \qquad GL_G(1,x)K_G(1,x)' + K_G(1,x)L_G(1,x)'G'$$

$$= G\int_0^1 K_G(r,x)K_G(r,x)'dr + \int_0^1 K_G(r,x)K_G(r,x)'drG'$$

$$+ xG\{G\int_0^1 L_G(r,x)K_G(r,x)'dr + \int_0^1 K_G(r,x)L_G(r,x)drG'\}$$

$$+ \{G\int_0^1 L_G(r,s)K_G(r,x)'dr + \int_0^1 K_G(r,x)L_G(r,x)drG'\}(xG')$$

$$+ G\int_0^1 L_G(r,x)dB(r)' + \int_0^1 [dB(r)L_G(r,x)']G'$$

It follows from (22) and (23) that

$$T^{-1}\sum_1^T (\dot{z}_{t-1}u_t' + u_t\dot{z}_{t-1}) \to_d G\int_0^1 L_G(r,x)dB(r)'$$

$$+ \int_0^1 [dB(r)L_G(r,x)']G'$$

as required for part (a).

To prove part (b) we note first from (18) that $\dot{z}_t = Gw_t$ where

$$w_t = \sum_1^t \exp\{((t-j)x/T)G\}(t-j)/Tu_j + (t/T)\exp\{(tx/T)G\}y_o.$$

Thus, from part (a) we have

$$G(T^{-1}\sum_1^T w_{t-1}u_t') + (T^{-1}\sum_1^T u_t w_{t-1}')G'$$

$$\underset{d}{\to} G\int_0^1 L_G(r,x)dB(r)' + \int_0^1 dB(r)L_G(r,x)'G'$$

It follows that

$$\text{(24)} \qquad tr\{G(T^{-1}\sum_1^T w_{t-1}u_t')\} \underset{d}{\to} tr\{G\int_0^1 L_G(r,x)dB(r)'\}$$

Since (24) holds for all matrices G we deduce that

$$T^{-1}\sum_1^T w_{t-1}u_t' \to \int_0^1 L_G(r,x)dB(r)'$$

Result (b) follows directly.

To prove (c) we integrate with respect to x over the interval [0,h]. We have

$$T^{-1}\sum_1^T\int_0^h \dot{z}_{t-1}(x)u_t'dx = T^{-1}\sum_1^T z_{t-1}(h)u_t' - T^{-1}\sum_1^T y_{t-1}u_t'$$

and

$$\int_0^h G\int_0^1 L_G(r,x)dB(r)'dx = \int_0^1\int_0^h GL_G(r,x)dxdB(r)'$$

$$= \int_0^1 K_G(r,h)dB(r)' - \int_0^1 B(r)dB(r)'$$

Part (c) now follows from (b) and the continuous mapping theorem.

Proof of Theorem 2.6 We work from part (c) of Lemma 2.5. First let $G = fI_n$ for some $f < 0$ and write

$$T^{-1}\sum_{1}^{T} z_{t-1}(h)u_t' = e^{-hf/T}T^{-1}\sum_{2}^{T}\left(\sum_{1}^{t-1} e^{(t-j)hf/T}u_j\right)u_t'$$

(25) $$= e^{-hf/T}\sum_{s=1}^{T-1} e^{shf/T}\left(T^{-1}\sum_{t=s+1}^{T} u_{t-s}u_t'\right) .$$

Now let $h = T/M$. We shall allow $M \uparrow \infty$ as $T \uparrow \infty$ in such a way that $M/T \downarrow 0$ (and, thus, $h \uparrow \infty$). (25) becomes:

$$e^{-f/M}\sum_{s=1}^{T-1} e^{sf/M}\left(T^{-1}\sum_{t=s+1}^{T} u_{t-s}u_t'\right)$$

But $e^{-f/M} \to 1$ as $M \uparrow \infty$ and

(26) $$\sum_{s=1}^{T-1} e^{sf/M}\left(T^{-1}\sum_{t=s+1}^{T} u_{t-s}u_t'\right) \to_p \Omega_1$$

In fact, (26) is simply the Abel estimate of the component Ω_1 of the scaled spectral density matrix $\Omega = 2\pi f_{uu}(0)$ at the origin (see, e.g. Hannan (1970, p. 279)). We deduce that

$$T^{-1}\sum_{1}^{T} z_{t-1}(1)u_t' - T^{-1}\sum_{1}^{T} y_{t-1}u_t'$$

has the same asymptotic distribution as $T \uparrow \infty$ (with $h = T/M \uparrow \infty$) as

(27) $$\Omega_1 - T^{-1}\sum_{1}^{T} y_{t-1}u_t' .$$

Now consider

$$K_G(r,h) = \int_0^r e^{(r-s)hf}dB(s) \equiv N\left(0, \int_0^r e^{(r-s)hf}ds\,\Omega\right)$$

$$\equiv N(0, ((e^{2rhf} - 1)/2hf)\Omega)$$

Since $f < 0$ we deduce that

$$K_G(r,h) \to_p 0$$

as $h \uparrow \infty$. We may also show that

$$(28) \qquad \int_0^1 K_G(r,h)dB(r)' \underset{p}{\to} 0.$$

Now part (c) of Lemma 2.5 holds for all h, so that combining (27) and (28) with part (c) we obtain

$$T^{-1}\sum_1^T y_{t-1}u_t' - \Omega_1 \underset{d}{\to} \int_0^1 B(r)dB(r)'$$

giving result (a) as stated. Note also that

$$\begin{aligned}\lim_{T\to\infty}T^{-1}\sum_1^T y_{t-1}u_t' &= \lim_{T\to\infty}T^{-1}\sum_2^T\sum_{s=1}^{t-1} E(u_{t-s}u_t') \\ &= \lim_{T\to\infty}\sum_{s=1}^{T-1}(1-s/T)E(u_1u'_{s+1}) \\ &= \sum_{k=2}^{\infty} E(u_1u_k') .\end{aligned}$$

Part (b) of the Theorem follows directly from part (a) and part (c) of Lemma 2.5.

Proof of Theorem 2.8 This follows as a consequence of Theorem 2.6, (6) and the continuous mapping theorem.

REFERENCES

Berkes, I. and W. Philipp (1981) "Approximation theorems for independent and weakly dependent random variables", Annals of Probability, 7, 29-54.

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

Box, G.E.D. and G.M. Jenkins (1976) Time Series Analysis: Forecasting and Control, San Francisco: Holden Day.

Eberlain, E. (1986) "On strong invariance principles under dependence assumptions", Annals of Probability, 14, 260-270.

Hannan, E.J. (1970) Multiple Time Series New York: Wiley.

Ibragimov, I.A. and Y.V. Linnik (1971) "Independent and stationary sequences of random variables", Groningen: Wolfen-Noordhoff.

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

Philipp, W. and W. Stout (1975) "Almost sure invariance principles for weakly dependent random variables", Memoirs of the American Mathematical Society, 161, 1-140.

Phillips, P.C.B. (1986a) "Time series regression with a unit root", Econometrica (forthcoming).

_____ (1986b) "Asymptotic expansions in nonstationary vector autoregressions", Econometric Theory (forthcoming).

_______ (1986c) "Regression theory for near-integrated time series", Cowles Foundation Discussion Paper, # 781, Yale University.

White, J.S. (1958) "The limiting distribution of the serial correlation coefficient in the explosive case", Annals of Mathematical Statistics, 29, 1188-1197.