

COWLES FOUNDATION FOR RESEARCH IN ECONOMICS
AT YALE UNIVERSITY

Box 2125, Yale Station
New Haven, Connecticut 06520

COWLES FOUNDATION DISCUSSION PAPER NO. 872

Note: Cowles Foundation Discussion Papers are preliminary materials circulated to stimulate discussion and critical comment. Requests for single copies of a Paper will be filled by the Cowles Foundation within the limits of the supply. References in publications to Discussion Papers (other than acknowledgment that a writer had access to such unpublished material) should be cleared with the author to protect the tentative character of these papers.

SPECTRAL REGRESSION FOR COINTEGRATED TIME SERIES

by

P .C .B. Phillips

June 1988

SPECTRAL REGRESSION FOR COINTEGRATED TIME SERIES

by

P. C. B. Phillips*

Cowles Foundation for Research in Economics
Yale University

April 1988

* This paper was written while the author was a visiting professor at the University of Auckland. My thanks go to Glenna Ames for her skill and effort in keyboarding the manuscript of this paper and to the NSF for support under grant number SES 8519595.

0. ABSTRACT

This paper studies the use of spectral regression techniques in the context of cointegrated systems of multiple time series. Several alternatives are considered including efficient and band spectral methods as well as system and single equation techniques. It is shown that single equation spectral regressions suffer asymptotic bias and nuisance parameter problems that render these regressions impotent for inferential purposes. By contrast systems methods are shown to be covered by the LAMN asymptotic theory, bringing the advantages of asymptotic median unbiasedness, scale nuisance parameters and the convenience of asymptotic chi-squared tests. System spectral methods also have advantages over full system direct maximum likelihood in that they do not require complete specification of the error processes. Instead they offer a nonparametric treatment of regression errors which avoids certain methodological problems of dynamic specification and permits additional generality in the class of error processes. In addition, systems spectral estimation leads to simply computed explicit formulae and thereby eliminates the inconvenience of the nonlinear estimation that is typically required by maximum likelihood as, for example, in the case of vector ARMA errors. System based spectral techniques that restrict information use to a frequency band around the origin share the same advantages as fully efficient methods. Interestingly, spectral methods require no special modifications to deal with the regressor endogeneity that is a characteristic feature of cointegrated systems.

1. INTRODUCTION

Efficient techniques for estimating the coefficients in a multiple system of linear equations by spectral methods were introduced by Hannan (1963). These techniques, which are related to work by Whittle (1951) on Gaussian likelihood estimation, provide the basis for a regression analysis in the frequency domain. Their principal advantage is that they permit a nonparametric treatment of regression errors so that it is not necessary for an investigator to be explicit about the generating mechanism for the errors other than to assume stationarity. In addition, the techniques make it possible to focus attention in a regression on the most relevant frequency thereby offering a selective approach that has become known as band spectrum regression—see Hannan and Robinson (1973) and Engle (1974). They have also been extended to nonlinear models under conditions which parallel those of nonlinear regression theory—see Hannan (1971) and Robinson (1972). All of the above mentioned theory has been developed for models where the time series are stationary.

The objective of the present paper is to show how spectral methods may also be usefully employed in regressions for certain nonstationary time series. Indeed, their use may even be more appealing in this context than in regressions for stationary series. The model we have in mind is a multivariate system of cointegrated time series. Such systems have been the object of study in many recent papers (see Engle and Granger (1987), Phillips and Durlauf (1986), Stock (1987) and the special issues of the *Oxford Bulletin of Economics and Statistics* (1986) and the *Journal of Economic Dynamics and Control* (1988)).

Our approach here follows that of some other ongoing research by the author (Phillips (1988)). This research focuses attention on full information estimation of cointegrated systems and gives strong arguments for the use of full maximum likelihood estimation of the system in error correction model (ECM) format. It is shown that such estimation brings the problem of inference within the locally asymptotically mixed normal (LAMN) family of Jeganathan (1980, 1982). This means that the cointegrating coefficients are asymptotically median unbiased and symmetrically distributed, that an optimal theory of inference applies and that hypothesis tests may be conducted using standard asymptotic chi-squared tests.

The present paper shows that similar advantages are enjoyed by systems spectral methods. Moreover, these methods have the additional advantage over classical maximum likelihood that they permit a nonparametric treatment of the regression errors. In other words, full system specification and estimation (as in maximum likelihood) is not required. Indeed, the system spectral methods given here involve linear estimating equations and result in simply computed explicit formulae. These features mean that the methods avoid what can be awkward methodological problems of dynamic specification and they focus entirely on what is the central problem of cointegrating regression theory—the estimation of long run equilibrium relationships.

The following notation is used throughout the paper. The symbol " \Rightarrow " signifies weak convergence, the symbol " $=$ " signifies equality in distribution and the inequality " > 0 " signifies positive definite when applied to matrices. Stochastic processes such as the Brownian motion $W(r)$ on $[0,1]$ are frequently written as W to achieve notational

economy. Similarly, we write integrals with respect to Lebesgue measure such as $\int_0^1 W(s)ds$ more simply as $\int_0^1 W$. Vector Brownian motion with covariance matrix Ω is written "BM(Ω)". We use $\|A\|$ to represent the Euclidean norm $\text{tr}(A'A)^{1/2}$ of the matrix A , $[x]$ to denote the smallest integer $\leq x$ and $I(1)$ and $I(0)$ to signify time series that are integrated of order one and zero, respectively. All limits given in the paper are taken as the sample size $T \rightarrow \infty$ unless otherwise stated.

2. MODEL AND ESTIMATORS

Our model is the cointegrated system

$$(1) \quad y_{1t} = \beta' y_{2t} + u_{1t}$$

$$(2) \quad \Delta y_{2t} = u_{2t}$$

where

$$y_t = \begin{bmatrix} y_{1t} \\ y_{2t} \end{bmatrix} \begin{matrix} 1 \\ m \end{matrix} = I(1)$$

is an integrated n -vector process ($n = m+1$) and

$$u_t = \begin{bmatrix} u_{1t} \\ u_{2t} \end{bmatrix} \begin{matrix} 1 \\ m \end{matrix} = I(0)$$

is stationary with continuous spectral density matrix $f_{uu}(\lambda) > 0$ over $-\pi < \lambda \leq \pi$. As formulated above (1) is a single equation cointegrating regression with cointegrating vector $\alpha' = (1, -\beta')$. Our approach may

easily be extended to multiple equation cointegrating regressions in which case β is a matrix of coefficients.

We shall assume that the partial sum process $\underline{P}_t = \sum_1^t u_j$ satisfies the multivariate invariance principle

$$(3) \quad T^{-1/2} \underline{P}_{[Tr]} \Rightarrow \underline{B}(r) = BM(\underline{\Omega}), \quad 0 < r \leq 1$$

where $\underline{\Omega} = 2\pi f_{uu}(0)$. We decompose the "long run" covariance matrix $\underline{\Omega}$ as follows

$$\underline{\Omega} = \underline{\Sigma} + \underline{\Lambda} + \underline{\Lambda}'$$

where

$$\underline{\Sigma} = E(u_0 u_0'), \quad \underline{\Lambda} = \sum_{k=1}^{\infty} E(u_0 u_k')$$

and we define

$$\underline{\Delta} = \underline{\Sigma} + \underline{\Lambda}.$$

In addition to (3) we assume weak convergence of the stochastic process constructed from the sample covariance between \underline{P}_t and u_t , viz

$$(4) \quad T^{-1} \sum_1^{[Tr]} \underline{P}_t u_t' \Rightarrow \int_0^r B dB' + r \underline{\Delta}.$$

Explicit conditions under which (3) and (4) hold are discussed in earlier work and the reader is referred to Phillips (1987) for references and for a review. Suffice to say here they are general enough to include a wide class of weakly dependent processes $\{u_t\}$ under mild moment conditions.

It is convenient to partition the Brownian motion B and the

matrices $\underline{\Omega}$, $\underline{\Sigma}$, $\underline{\Lambda}$, $\underline{\Delta}$ conformably with the vector y_t . For example, we shall write

$$B = \begin{bmatrix} B_1 \\ B_2 \end{bmatrix}, \quad \underline{\Omega} = \begin{bmatrix} \omega_{11} & \omega'_{21} \\ \omega_{21} & \Omega_{22} \end{bmatrix}, \quad \underline{\Sigma} = \begin{bmatrix} \sigma_{11} & \sigma'_{21} \\ \sigma_{21} & \Sigma_{22} \end{bmatrix}$$

and so on. We also define $\omega_{-11.2} = \omega_{11} - \omega'_{21}\Omega_{22}^{-1}\omega_{21}$.

The cointegrated system (1) and (2) has the following ECM representation

$$(5) \quad \Delta y_t = \gamma \alpha' y_{t-1} + v_t$$

where

$$\gamma' = (-1, 0), \quad \alpha' = (1, -\beta')$$

$$v_t = \begin{bmatrix} 1 & \beta' \\ 0 & I \end{bmatrix} u_t = Du_t$$

(see Phillips (1988)). It is this system that we propose to estimate using spectral methods.

Note that the error process v_t in (5) is stationary with spectral matrix $f_{vv}(\lambda) = Df_{uu}(\lambda)D' > 0$. We write $\Omega = 2\pi f_{vv}(0)$, $B(r) = DB(r) = BM(\Omega)$, $P_t = DP_t$ and similarly write $\Sigma = D\underline{\Sigma}D'$, $\Lambda = DAD'$, $\Delta = D\underline{\Delta}D'$. These matrices and vectors are partitioned conformably with y_t just as their counterparts without the sub bar. Corresponding to (3) and (4) we now have

$$(6) \quad T^{-1/2}P_{\{Tr\}} \Rightarrow B(r),$$

$$(7) \quad T^{-1} \sum_1^{[Tr]} P_t v_t' \Rightarrow \int_0^r B dB + r \Delta .$$

We make use of the efficient method of estimation introduced by Hannan (1963) for linear systems and later extended by Hannan (1971) and Robinson (1972) to non linear regression equations. To this end we introduce the finite Fourier transforms

$$w_{\Delta}(\lambda) = (2\pi T)^{-1/2} \sum_{t=1}^T \Delta y_t e^{it\lambda}$$

$$w_{*}(\lambda) = (2\pi T)^{-1/2} \sum_{t=1}^T y_{*t} e^{it\lambda}$$

$$w_y(\lambda) = (2\pi T)^{-1/2} \sum_{t=1}^T y_{t-1} e^{it\lambda}$$

$$w_v(\lambda) = (2\pi T)^{-1/2} \sum_{t=1}^T v_t e^{it\lambda}$$

for $\lambda \in [-\pi, \pi]$, $y_{*t} = (y_{1t}, \Delta y_{2t})$ and we transform (4) accordingly as

$$(8) \quad w_{\Delta}(\lambda) = \gamma \alpha' w_y(\lambda) + w_v(\lambda) .$$

We partition $w_y(\lambda)$ conformably with y_t using the notation

$$w_y(\lambda)' = (w_1(\lambda), w_2(\lambda)') .$$

Following Hannan (1971) and Robinson (1972) a class of nonlinear weighted least squares estimates of α may be obtained by minimization of the Hermitian form

$$(9) \quad \sum_B \text{tr} \{ [w_{\Delta}(\lambda_s) - \gamma \alpha' w_y(\lambda_s)] [w_{\Delta}(\lambda_s) - \gamma \alpha' w_y(\lambda_s)]^* \Phi(\lambda_s) \}$$

with respect to α , where Φ is a given positive definite Hermitian matrix, $\lambda_s = 2\pi s/T$ and s is integral with values in the interval $-[T/2] < s \leq [T/2]$. The summation in (9) is over $\lambda_s \in B$ which is a subset of $(-\pi, \pi)$ such that if $\lambda \in B$ then $-\lambda \in B$ also. The use of B permits the restriction of the regression to a set of frequency bands in $(-\pi, \pi)$ and is inspired by the idea that the model, when formulated in the frequency domain as in (8), may be more appropriate for λ in certain bands than in others. In practice, therefore, the regression may be confined to what seem to be the relevant frequency bands and as such is known as band spectrum regression. The reader is referred to Hannan (1963, 1970), Hannan and Robinson (1973) and Robinson (1972, 1976) for further details and discussion and to Engle (1974) for an econometric application of these ideas.

In the present context, since the cointegrating vector α defines a long run relationship between the components of the time series one possibility would be to confine the regression to a band around the origin so that low frequency elements in the series are emphasized. In his application to a U.S. aggregate expenditure relationship Engle (1974) in one case eliminated high frequency elements from the regression, using the argument that these are associated with transitory components of the two variables expenditure and income in the regression.

In conventional spectral regression choice of the weight function $\Phi(\cdot)$ that appears in (9) involves only efficiency considerations. Indeed, the criterion (9) would be proportional to the exponent in the Gaussian likelihood of (8) if the $w_v(\lambda_s)$ were independent (complex) normal random vectors with covariance matrix $\Phi(\lambda_s)^{-1}$. In the Hannan

efficient procedure $\Phi(\cdot)$ is selected in such a way that this is achieved asymptotically. This approach, which originates in the work of Whittle (1951), relies on the fact that under rather general conditions on $\{v_t\}$ and for λ_s in a band around ω (so that $\lambda_s \rightarrow \omega$ as $T \rightarrow \infty$) we find

$$(10) \quad w_v(\lambda_s) \Rightarrow N^c(0, f_{vv}(\omega)), \quad \omega \neq 0, \pi$$

(for example, Brillinger (1974), Theorem 4.4.1) where " N^c " signifies the complex normal distribution. In designing an efficient procedure we may then select $\Phi(\lambda_s) = \hat{f}_{vv}(\omega)^{-1}$ for λ_s in a band centered on ω and for some suitable choice of consistent spectral estimate \hat{f}_{vv} . Details of the construction are given by Hannan (1963).

We envisage a straightforward application of these ideas in the present context. However, unlike the conventional spectral regression model the regressors y_{t-1} in (3) (and hence $w_y(\lambda)$ in (4)) are in general coherent with the errors v_t ($w_v(\lambda)$). The regressors are also $I(1)$ not $I(0)$ processes. These features of the present model make the choice of weight function Φ critical. As we shall show, a nonefficient choice of Φ induces a (second order) bias effect in estimation as well as a loss of efficiency. As a result, fully efficient procedures have much more to recommend them in the present application.

A simple way to estimate $f_{vv}(\lambda)$ is to use the residuals from an initial least squares regression on (1). Writing $\hat{v}_t = \Delta y_t - \gamma \hat{\alpha}' y_{t-1}$ we may now compute the smoothed periodogram estimate

$$(11) \quad \hat{f}_{vv}(\omega_j) = \frac{M}{T} \sum_{B_j} [w_{\Delta}(\lambda_s) - \gamma \hat{\alpha}' w_y(\lambda_s)] [w_{\Delta}(\lambda_s) - \gamma \hat{\alpha}' w_y(\lambda_s)]^*$$

where the summation is over

$$\lambda_s \in B_j = (\omega_j - \pi/2M < \lambda \leq \omega_j + \pi/2M)$$

that is, a frequency band of width π/M centered on

$$\omega_j = \frac{\pi j}{M}, \quad j = -M+1, \dots, M$$

for M integer. Setting $m = [T/M]$ we are now in effect averaging m neighboring periodogram ordinates around the frequency ω_j to obtain $\hat{f}_{vv}(\omega_j)$. As usual, we require $M \rightarrow \infty$ in such a way that $M/T \rightarrow 0$ (so that $m \rightarrow \infty$). In fact, it is convenient for the proofs to require that $M = o(T^{1/2})$, as in Hannan (1970, p. 489). Since $\hat{\alpha}$ is consistent (Phillips and Durlauf (1986), Stock (1987)) we find that when $\omega_j \rightarrow \omega$ we have $\hat{f}_{vv}(\omega_j) \rightarrow_p f_{vv}(\omega)$ as $T \rightarrow \infty$.

A further consideration is that, since $\gamma' = (-1, 0)$ is known by virtue of the construction of (5), nonlinear methods are not required. Indeed, minimization of (9) with the following choice of weight function

$$\Phi(\lambda_s) = \hat{f}_{vv}(\omega_j)^{-1} \quad \text{for all } \lambda_s \in B_j$$

leads directly to the estimator

$$(12) \quad \begin{aligned} \bar{\beta} &= - \left[\frac{1}{2M} \sum_{j=-M+1}^M \gamma' \hat{f}_{vv}^{-1}(\omega_j) \gamma \hat{f}_{22}(\omega_j) \right]^{-1} \left[\frac{1}{2M} \sum_{j=-M+1}^M \hat{f}_{2*}(\omega_j) \hat{f}_{vv}^{-1}(\omega_j) \gamma \right] \\ &= - \left[\frac{1}{2M} \sum_{j=-M+1}^M e' \hat{f}_{vv}^{-1}(\omega_j) e \hat{f}_{22}(\omega_j) \right]^{-1} \left[\frac{1}{2M} \sum_{j=-M+1}^M \hat{f}_{2*}(\omega_j) \hat{f}_{vv}^{-1}(\omega_j) e \right] \end{aligned}$$

where

$$(13) \quad \hat{f}_{22}(\omega_j) = \frac{1}{m} \sum_{B_j} w_2(\lambda_s) w_2(\lambda_s)^*$$

$$(14) \quad \hat{f}_{2*}(\omega_j) = \frac{1}{m} \sum_{B_j} w_2(\lambda_s) w_{*}(\lambda_s)^*$$

and $e' = -\gamma' = (1, 0, \dots, 0)$ is the first unit n -vector.

Since our focus of interest is the (long run) cointegrating vector $\alpha' = (1, -\beta')$ alternative estimators might be considered that are based on low frequency averages. One such possibility is

$$(15) \quad \tilde{\beta}_{(0)} = -\hat{f}_{22}(0)^{-1} \hat{f}_{2*}(0) \hat{f}_{vv}^{-1}(0) \gamma / \gamma' \hat{f}_{vv}^{-1}(0) \gamma = \hat{f}_{22}(0)^{-1} \hat{f}_{2*}(0) \hat{f}_{vv}^{-1}(0) e / e' \hat{f}_{vv}^{-1}(0) e$$

which relies only on spectral estimates at the origin. The information that is neglected in the formation of $\tilde{\beta}_{(0)}$ (in relation to $\tilde{\beta}$) turns out to be unimportant at least asymptotically as we shall show in Section 3.

In formulae (12) and (15) above, $\tilde{\beta}$ and $\tilde{\beta}_{(0)}$ have been constructed from the smoothed periodogram spectral estimates (11), (13) and (14). We observe that other conventional choices of spectral estimate may be employed in these formulae without affecting the asymptotic theory obtained below.

3. ASYMPTOTIC THEORY

Our main result is the following:

THEOREM 3.1

$$(a) \quad T(\tilde{\beta} - \beta) \Rightarrow \left(\int_0^1 B_2 B_2' \right)^{-1} \left(\int_0^1 B_2 dB_{1,2} \right)$$

$$(b) \quad T(\tilde{\beta}_{(0)} - \beta) \Rightarrow \left(\int_0^1 B_2 B_2' \right)^{-1} \left(\int_0^1 B_2 dB_{1,2} \right)$$

where

$$\begin{bmatrix} B_{1.2} \\ B_2 \end{bmatrix} \begin{matrix} 1 \\ m \end{matrix} = \text{BM} \left(\begin{bmatrix} \omega_{11.2} & 0 \\ 0 & \Omega_{22} \end{bmatrix} \right)$$

and

$$\omega_{11.2} = \omega_{11} - \omega_{21}' \Omega_{22}^{-1} \omega_{21} .$$

The common limit distribution in (a) and (b) is given explicitly in mixed normal form by the integral

$$(16) \quad \int_{g>0} N(0, g\omega_{11.2}\Omega_{22}^{-1}) dP(g)$$

where

$$(17) \quad g = e_1' \left(\int_0^1 W_2 W_2' \right)^{-1} e_1 ,$$

e_1 is the first unit m -vector and $W_2 = \text{BM}(I_m)$.

REMARK (a). We see from the above result that $\tilde{\beta}$ and $\tilde{\beta}_{(0)}$ are asymptotically equivalent. Information in the component spectral estimates at the origin is all that is relevant in the limit distribution and this is all that is used in the construction of $\tilde{\beta}_{(0)}$. In empirical work, of course, $\tilde{\beta}$ and $\tilde{\beta}_{(0)}$ will differ. However, since much of the spectral power is concentrated in an immediate neighborhood of the origin for most aggregate economic time series it seems likely that this difference between the estimates will not be great in many practical applications.

REMARK (b). The representation (16) shows that the limit distribution is a continuous mixture of normals. The mixing variate is the scalar (17).

If we partition the m -vector standard Brownian motion W_2 as

$$W_2' = \begin{bmatrix} 1 & m-1 \\ W_{21} & W_{22}' \end{bmatrix}$$

then we can also write (17) in the form

$$g = \left\{ \int_0^1 W_{21}^2 - \int_0^1 W_{21} W_{22}' \left(\int_0^1 W_{22} W_{22}' \right)^{-1} \int_0^1 W_{22} W_{21} \right\}^{-1}.$$

REMARK (c). The limit distribution (16) is the same as that of the full maximum likelihood estimator of β in (5) when an explicit parametric model is assumed for the data generating mechanism of the innovation vector v_t . When v_t is generated by an ARMA process this estimator is obtained by constructing the full (Gaussian) likelihood by a method such as the innovations algorithm (see, for example, Brockwell and Davis (1986)). The properties of this maximum likelihood estimator of β are explored in Phillips (1988). The above result shows that spectral regression offers a simple alternative to maximum likelihood that has several advantages:

- (i) The method leads to explicit easily calculated formulae;
- (ii) it offers the additional generality of stationary (rather than ARMA) errors in (5);
- (iii) it avoids the methodological problems that are involved in the specification of short-run dynamics through what is, in effect, a nonparametric treatment of the errors.

REMARK (d). As remarked earlier in Section 2, the choice of an efficient estimator is critical to the above result. Suppose, for example, a general weight function $\Phi(\cdot)$ were employed in (9). This would lead to the following estimator in place of $\tilde{\beta}$:

$$\tilde{\beta}_{\Phi} = - \left[\frac{1}{2M} \sum_{j=-M+1}^M \gamma' \Phi(\omega_j) \gamma \hat{f}_{22}(\omega_j) \right]^{-1} \left[\frac{1}{2M} \sum_{j=-M+1}^M \hat{f}_{2*}(\omega_j) \Phi(\omega_j) \gamma \right] .$$

The asymptotics for this estimator are given by:

THEOREM 3.2

$$(18) \quad T(\tilde{\beta}_{\Phi} - \beta) \Rightarrow \left[\int_0^1 B_2 B_2' \right]^{-1} \left(\int_0^1 B_2 dB' \Phi(0) e + \sum_{g=-\infty}^{\infty} \Delta_2(g+1) F_g e \right) / e' \Phi(0) e$$

where $\Phi(\lambda)$ has the following Fourier series representation:

$$(19) \quad \Phi(\lambda) = (1/2\pi) \sum_{g=-\infty}^{\infty} F_g e^{ig\lambda}$$

and where

$$\Delta_2(g) = \sum_{j=0}^{\infty} E(u_{20} v'_{j+g}) .$$

The limit distribution (18) is no longer mixed normal. The distribution is, in fact, miscentered by a second order bias that arises from two sources: the term $\sum_g \Delta_2(g+1) F_g e$; and the fact that the Brownian motion $B_2(r)$ is in general correlated with the Brownian motion $B'(r)\Phi(0)e$. Second and perhaps more importantly, the limit distribution (18) involves nuisance parameters which inhibit statistical inference. These nuisance parameters involve both the bias effects and the covariance matrix of the Brownian motion $B(r)$. They cannot be easily eliminated and their

presence in the limit distribution renders (18) effectively impotent for inferential purposes.

REMARK (e). The limit results given in Theorem 3.1 belong to the LAMN theory of Jeganathan (1980, 1982), LeCam (1986) and Davies (1986). As pointed out earlier, the criterion function (9) is asymptotically proportional to the exponent of the Gaussian likelihood of the model (5). This Gaussian likelihood belongs to the LAMN family of Jeganathan (1980) (see Phillips (1988) for details). The estimators $\bar{\beta}$ and $\bar{\beta}_{(0)}$ may therefore be regarded as spectral versions of maximum likelihood. As such they have all of the advantages of the latter, viz

- (i) they are asymptotically median unbiased and symmetrically distributed;
- (ii) the nuisance parameters that appear in the limit distribution (16) involve only scale effects and are readily eliminated to facilitate inference;
- (iii) an optimal theory of inference applies (from LeCam (1986));
- (iv) hypothesis testing may be conducted using conventional asymptotic chi-squared criteria.

REMARK (f). To pursue point (iv) above suppose we wish to test the following hypotheses about the cointegration space

$$H_0 : h(\beta) = 0 , \quad H_1 : h(\beta) \neq 0$$

where $h(\)$ is a twice continuously differentiable q -vector function of restrictions on β . We assume that $H = \partial h(\beta)/\partial \beta'$ has rank $q < m$.

To test H_0 against H_1 we may employ the Wald statistic in its

usual form. Thus for the estimator $\bar{\beta}$ we set up

$$M_1 = h(\bar{\beta})' [\bar{H} V_T \bar{H}']^{-1} h(\bar{\beta})$$

where $\bar{H} = H(\bar{\beta})$ and

$$V_T = \frac{1}{T} \left[\frac{1}{2M} \sum_{j=-M+1}^M \gamma' \hat{f}_{vv}^{-1}(\omega_j) \gamma \hat{f}_{22}(\omega_j) \right]^{-1}.$$

Here V_T is the conventional estimate of the asymptotic variance matrix of $\bar{\beta}$ from spectral regression theory (see Hannan (1970) page 442).

Similarly for $\bar{\beta}_{(0)}$ we construct

$$M_2 = h(\bar{\beta}_{(0)})' [\bar{H}_0 V_{T0} \bar{H}_0']^{-1} h(\bar{\beta}_{(0)})$$

where $\bar{H}_0 = H(\bar{\beta}_{(0)})$ and

$$V_{T0} = \frac{1}{T} [\gamma' \hat{f}_{vv}^{-1}(0) \gamma \hat{f}_{22}(0)]^{-1}.$$

We have

THEOREM 3.3

$$M_1, M_2 \Rightarrow \chi_q^2.$$

Thus, statistical tests of H_0 may be conducted in the usual fashion of asymptotic chi-squared tests. Interestingly no modification to the conventional formulae from spectral regression theory are required. This is because the variance matrix estimates V_T and V_{T0} , although random upon appropriate standardization in the limit, still provide the right

metric for measuring departures of $h(\bar{\beta})$ and $h(\bar{\beta}_0)$ from the null hypothesis.

REMARK (g). Single equation spectral regression methods do not have the same advantages as the systems estimators $\bar{\beta}$ and $\bar{\beta}_0$. To see this it is helpful to consider the following estimate which is the analogue of $\bar{\beta}$ for the first equation of (5)

$$\beta^* = \left[\frac{1}{2M} \sum_{j=-M+1}^M \hat{f}_{v_1 v_1}^{-1}(\omega_j) \hat{f}_{22}(\omega_j) \right]^{-1} \left[\frac{1}{2M} \sum_{j=-M+1}^M \hat{f}_{21}(\omega_j) \hat{f}_{v_1 v_1}^{-1}(\omega_j) \right]$$

where $\hat{f}_{21}(\lambda)$ is an estimate of the cross spectrum between y_{2t-1} and y_{1t} . The estimator β^* is the Hannan (1963) efficient estimator of β in the equation

$$(20) \quad y_{1t} = \beta' y_{2t-1} + v_{1t}.$$

After minor modifications to adjust for the lag in (20) this is just the standard spectral regression estimator of β in the cointegrating regression equation (1). The asymptotic theory for β^* is given by

THEOREM 3.4

$$(20) \quad T(\beta^* - \beta) \Rightarrow \left[\int_0^1 B_2 B_2' \right]^{-1} \left(\int_0^1 B_2 dB_1 + \delta \right)$$

where

$$\delta = \left(\sum_{g \rightarrow \infty} \Delta_{21}(g+1) d_g \right) / \left(\sum_{g \rightarrow \infty} d_g \right)$$

and

$$\Delta_{21}(g) = \sum_{j=0}^{\infty} E(u_{20} v_{1j+g})$$

The limit distribution (20) involves second bias effects and nuisance parameters arising from the presence of δ in the second factor of (20) and the correlation between the Brownian motions B_1 and B_2 . As in the case of (18) these problems severely inhibit the usefulness of the estimator β^* for inferential purposes.

Note that by decomposing B_1 as follows

$$B_1(r) = \omega'_{21} \Omega_{22}^{-1} B_2(r) + \omega_{11 \cdot 2}^{1/2} W_1(r)$$

where $W_1(r)$ is standard Brownian motion, i.e. BM(1), and W_1 is independent of B_2 we deduce an alternative representation of (20) in the form

$$(21) \quad \left(\int_0^1 B_2 B_2' \right)^{-1} \left(\int_0^1 B_2 dB_2' \Omega_{22}^{-1} \omega_{21} + \delta \right) + \omega_{11 \cdot 2}^{1/2} \left(\int_0^1 B_2 B_2' \right)^{-1} \int_0^1 B_2 dW_1 .$$

The first term of (21) involves the "unit root" distribution

$\left(\int_0^1 B_2 B_2' \right)^{-1} \left(\int_0^1 B_2 dB_2' \right)$ and the "bias effects" from the factor $\Omega_{22}^{-1} \omega_{21}$ and δ . The second term of (21) is mixed normal with the same distribution as (16).

The decomposition (21) highlights the differences between single equation and systems spectral regressions in the model (5). Single equation methods neglect the prior information of the m unit roots in (5) and ignore the joint dependence of y_{1t} and y_{2t} . As a result these methods implicitly involve the estimation of unit roots and this is responsible for the presence of the unit root distribution in the first

term of (21). In addition we see that neglect of the rest of the system in (5) imports a second order bias effect through the term δ . The magnitude of this term depends on the extent of the contemporaneous and serial correlation between u_{2t} and v_{1t} .

Finally, we observe that Theorem 3.4 gives the correct asymptotic theory for the (full band) spectral estimator used by Engle (1974) in his application of spectral regression to the aggregate consumption function with quarterly U.S. data on money income and consumption. Our results suggest that the estimates of the propensity to consume obtained by Engle in this study are likely to be biased and that conventional tests are invalidated by the asymptotic theory. It would be worthwhile to reanalyze this data set using the systems estimator $\bar{\beta}$ (and $\bar{\beta}_{(0)}$) and associated test statistics such as M_1 (and M_2).

4. CONCLUSION

This paper provides a frequency domain extension of the results in Phillips (1988) on the maximum likelihood estimation of cointegrated systems. Indeed full system spectral regression in an ECM is asymptotically equivalent to maximum likelihood and shares with it the advantages of belonging to the LAMN family. But spectral regression techniques seem to have more appeal in the context of cointegrated time series. This is because:

1. they involve only linear estimating equations and thereby avoid the nonlinear optimization methods that are typically called for in the application of maximum likelihood (for instance, when there are ARMA error processes);

2. the nonparametric treatment of regression errors that is inherently involved in spectral methods avoids the methodological difficulties that are encountered with the need to completely specify the data generating mechanism of the errors before maximum likelihood is applied;
3. the nonparametric approach brings with it additional generality concerning the error processes at what seems to be little or no extra cost;
4. even simpler methods are available like the systems band spectral regression estimator given by (15) and such estimators continue to enjoy the same asymptotic properties as the full system estimator (12);
5. standard systems spectral regressions may be used with no modifications being necessary to deal with the regressor endogeneity that is characteristic of cointegrated systems.

We emphasize that it is the systems spectral estimators given by (12) and (15) that have these advantages. Single equation or subsystem spectral regressions have quite different asymptotic properties as shown in Theorem 3.4. In particular, they suffer from bias and nuisance parameter dependencies that seriously inhibit their use for inference. Thus systems estimation brings with it considerably more than the usual efficiency gains we have come to expect from traditional asymptotic theory. In view of these apparent advantages of systems spectral estimators over direct maximum likelihood and single equation spectral methods it would seem worthwhile to investigate their performance in sampling experiments and in empirical work.

APPENDIX

Proof of Theorem 3.1. (a) From (12) we find that

$$(A1) \quad T(\tilde{\beta} - \beta) = \left[\frac{1}{2MT} \sum_{j=-M+1}^M e^{i\omega_j} \hat{f}_{VV}^{-1}(\omega_j) e^{-i\omega_j} \hat{f}_{22}(\omega_j) \right]^{-1} \left[\frac{1}{2M} \sum_{j=-M+1}^M \hat{f}_{2V}(\omega_j) \hat{f}_{VV}^{-1}(\omega_j) e^{i\omega_j} \right].$$

Our approach follows Hannan (1963) in general outline with the main differences arising from the treatment of the nonstationary elements.

It is convenient to work with spectral estimates in (A1) of the same general form, say

$$\hat{f}_{ab}(\lambda) = \frac{1}{2\pi} \sum_{n=-M}^M k\left(\frac{n}{M}\right) c_{ab}(n) e^{-in\lambda}$$

where

$$c_{ab}(n) = T^{-1} \sum_{t=1}^T a_t b_{t+n}, \quad 1 \leq t+n \leq T$$

and where the lag window $k(\cdot)$ is a bounded even function defined on $[-1, 1]$ with $k(0) = 1$. For example, when $k(n/M) = 1 - |n|/M$, $\hat{f}_{ab}(\lambda)$ is the Bartlett estimator (e.g. Hannan (1970), p. 278). We may also replace $k(n/M)$ by $k(n/M)(1 - |n|/T)$ in the above formula without affecting the arguments that follow. Other spectral estimates may also be employed but the above formula helps to simplify derivations and avoid repetition.

As in Hannan (1963) we have

$$\max_{\lambda} \left\| \hat{f}_{VV}(\lambda) - f_{VV}(\lambda) \right\| \xrightarrow{p} 0$$

as $T \rightarrow \infty$ and then the limit behavior of (A1) is equivalent to that of the same expression but with $f_{vv}(\omega_j)$ replacing $\hat{f}_{vv}(\omega_j)$ in both factors on the right hand side. We take each of these in turn.

Using the Fourier series

$$f_{vv}^{-1}(\lambda) = \frac{1}{2\pi} \sum_{g=-\infty}^{\infty} D_g e^{ig\lambda}$$

we have

$$\begin{aligned} & \frac{1}{2MT} \sum_{j=-M+1}^M e' f_{vv}^{-1}(\omega_j) e \hat{f}_{22}(\omega_j) \\ &= \frac{1}{2\pi T} \sum_{g=-\infty}^{\infty} e' D_g e \frac{1}{2M} \sum_{j=-M+1}^M e^{ig\pi j/M} \hat{f}_{22}(\pi j/M) \\ (A2) \quad &= \left(\frac{1}{2\pi}\right)^2 \frac{1}{T} \sum_{g=-\infty}^{\infty} e' D_g e c_{22}(g) k\left(\frac{g}{M}\right) \end{aligned}$$

where

$$g + 2\ell M = g, \quad -M+1 \leq g \leq M$$

for some integer ℓ and where

$$c_{22}(n) = T^{-1} \sum_1^T y_{2t} y_{2t+n}, \quad 1 \leq t+n \leq T.$$

The next step is to determine the asymptotic behavior of (A2). We start by defining the random elements

$$X_T(r) = T^{-1/2} P_{[Tr]}, \quad Z_{Ti}(r) = T^{-1} \sum_1^{[Tr]} P_{t \ v'_{t+i}}.$$

In view of (6) and (7) we have the weak convergence

$$X_T(r) \Rightarrow X_{\infty}(r) = B(r)$$

$$Z_{T1}(r) \Rightarrow Z_{\infty 1}(r) = \int_0^r B dB + r\Delta(1)$$

where

$$\Delta(1) = \sum_{j=0}^{\infty} E(v_0 v'_{j+1}) .$$

Using the Skorohod construction we now employ a new probability space with random elements $((X_T^*, Z_{T1}^*))$, $(X_{\infty}^*, Z_{\infty 1}^*)$ for which

$$(A3) \quad X_T^* \xrightarrow[\text{a.s.}]{} X_{\infty}^* , \quad Z_{T1}^* \xrightarrow[\text{a.s.}]{} Z_{\infty 1}^*$$

and where

$$X_T = X_T^* , \quad Z_{T1} = Z_{T1}^*$$

(with " = " as usual representing equivalence in distribution). It will be convenient to use a superscript "2" on these random elements to signify subelements of matrices and vectors that correspond to the component u_{2t} of v_t . Thus we write

$$X_{\infty}^{(2)} = B_2 , \quad Z_{\infty 1}^{(2)} = \int_0^r B_2 dB' + r\Delta_2(1)$$

$$Z_{\infty 1}^{(22)} = \int_0^r B_2 dB'_2 + r\Delta_{22}(1)$$

and so on. With this notation we have for $0 \leq n \leq M$ and up to a term of $O_p(T^{-1})$

$$\begin{aligned}
T^{-1}c_{22}(n) &= T^{-2} \sum_{1 \leq t \leq T} y_{2t} y'_{2t+n} \\
&= \int_0^1 X_T^{(2)} X_T^{(2)'} + T^{-1} \{ Z_{T1}^{(22)}(1) + \dots + Z_{Tn}^{(22)}(1) \} \\
&= \int_0^1 X_T^{*(2)} X_T^{*(2)'} + T^{-1} \{ Z_{T1}^{*(22)}(1) + \dots + Z_{Tn}^{*(22)}(1) \} \quad (:= T^{-1}c_{22}^*(n), \text{ say}) \\
&\xrightarrow{\text{a.s.}} \int_0^1 B_2^* B_2^{*'} .
\end{aligned}$$

In view of (A3) the final convergence takes place almost surely and uniformly in $|n| \leq M$ as $T \rightarrow \infty$. The same result also applies when $-M \leq n \leq 0$.

Since $k(g/M) \rightarrow 1$ for all fixed g as T (and hence M) $\rightarrow \infty$ we deduce that

$$\left(\frac{1}{2\pi} \right)^2 \frac{1}{T} \sum_{g \rightarrow \infty} e' D_g e c_{22}^*(g) k\left(\frac{g}{M}\right) \xrightarrow{\text{a.s.}} \left(\frac{1}{2\pi} \right)^2 \left(\sum_{g \rightarrow \infty} e' D_g e \right) \int_0^1 B_2^* B_2^{*'} .$$

This, of course, implies that

$$\left(\frac{1}{2\pi} \right)^2 \frac{1}{T} \sum_{g \rightarrow \infty} e' D_g e c_{22}^*(g) k\left(\frac{g}{M}\right) \Rightarrow \left(\frac{1}{2\pi} \right)^2 \left(\sum_{g \rightarrow \infty} e' D_g e \right) \int_0^1 B_2^* B_2^{*'} .$$

However,

$$c_{22}(g) = c_{22}^*(g) \quad \text{for all } g$$

and

$$B_2(r) = B_2^*(r)$$

so that by a simple modification of the Skorohod, Dudley, Wichura theorem (e.g. Shorack and Wellner (1987), p. 47) we deduce that

$$(A4) \quad \left(\frac{1}{2\pi}\right)^2 \frac{1}{T} \sum_{g \rightarrow \infty} e' D_g e c_{22}(g) k\left(\frac{g}{M}\right) \Rightarrow \left(\frac{1}{2\pi}\right)^2 \left(\sum_{g \rightarrow \infty} e' D_g e\right) \int_0^1 B_2 B_2' .$$

Next observe that

$$(1/2\pi) \sum_{g \rightarrow \infty} e' D_g e = e' f_{vv}^{-1}(0) e$$

so that the right hand side of (A4) is simply

$$(A5) \quad e' \Omega^{-1} e \int_0^1 B_2 B_2' = (1/\omega_{11,2}) \int_0^1 B_2 B_2' .$$

It remains to consider the second factor in the right hand element of

(A1). Replacing \hat{f}_{vv} by f_{vv} for the reason given earlier we have

$$\begin{aligned} & \frac{1}{2M} \sum_{j \rightarrow M+1}^M \hat{f}_{2v}(\omega_j) f_{vv}^{-1}(\omega_j) e \\ &= \frac{1}{2\pi} \sum_{g \rightarrow \infty} \left[\frac{1}{2M} \sum_{j \rightarrow M+1}^M \hat{f}_{2v}(\omega_j) e^{ig\omega_j} \right] D_g e \\ &= \left(\frac{1}{2\pi}\right)^2 \sum_{g \rightarrow \infty} c_{2v}(g) D_g e k(g/M) . \end{aligned}$$

Now

$$\begin{aligned} c_{2v}(n) &= T^{-1} \sum_1^T y_{2t-1} v_{t+n} , \quad 1 \leq t+n \leq T \\ &= Z_{Tn}^{(2)}(1) \\ &= Z_{Tn}^{*(2)}(1) \quad (:= c_{2v}^*(n) , \text{ say}) \\ &\xrightarrow{\text{a.s.}} Z_{\infty n}^{*(2)}(1) = \int_0^1 B_2^* dB_2^{*'} + \Delta_2(n+1) . \end{aligned}$$

We deduce that

$$\left(\frac{1}{2\pi}\right)^2 \sum_{g \rightarrow \infty} c_{2v}^*(g) D_g e^k \left(\frac{g}{M}\right) \xrightarrow{\text{a.s.}} \left(\frac{1}{2\pi}\right)^2 \int_0^1 B_2^* dB^* (\sum_{g \rightarrow \infty} D_g e) + \left(\frac{1}{2\pi}\right)^2 \sum_{g \rightarrow \infty} \Delta_2(g+1) D_g e$$

Using the Skorohod-Dudley-Wichura theorem as before we obtain

$$(A6) \quad \left(\frac{1}{2\pi}\right)^2 \sum_{g \rightarrow \infty} c_{2v}^*(g) D_g e^k \left(\frac{g}{M}\right) \Rightarrow \left(\frac{1}{2\pi}\right)^2 \int_0^1 B_2 dB' (\sum_{g \rightarrow \infty} D_g e) + \left(\frac{1}{2\pi}\right)^2 \sum_{g \rightarrow \infty} \Delta_2(g+1) D_g e$$

Note that

$$(A7) \quad e' (1/2\pi)^2 (\sum_g D_g) B(r) - e' \Omega^{-1} B(r) := B_e(r) = BM(1/\omega_{11.2})$$

Next we define

$$\underline{v}_j = \sum_{g \rightarrow \infty} v'_{g+j+1} D_g e$$

from which we deduce

$$(A8) \quad \sum_g \Delta_2(g+1) D_g e = \sum_{j=0}^{\infty} E(u_{20-\underline{j}}) = 0$$

since

$$E(u_{20-\underline{j}}) = \int_{-\pi}^{\pi} e^{ij\lambda} f_{2\underline{v}}(\lambda) d\lambda = 0 \quad \text{for all } j.$$

The latter follows in view of the fact that

$$\begin{aligned} f_{2\underline{v}}(\lambda) &= [0 \quad I] f_{\underline{v}\underline{v}}(\lambda) (\sum_{g \rightarrow \infty} D_g e^{ig\lambda}) e \\ &= 2\pi [0 \quad I] f_{\underline{v}\underline{v}}(\lambda) f_{\underline{v}\underline{v}}^{-1}(\lambda) e \\ &= 0 \end{aligned}$$

for all $\lambda \in (-\pi, \pi]$.

From (A6)-(A8) we obtain

$$(A9) \quad (1/2\pi)^2 \sum_{g=-\infty}^{\infty} c_{2v}(g) D_g \text{ek}(g/M) \Rightarrow \int_0^1 B_2 dB_e .$$

Combining (A9) with (A4), (A5) and (A1) we deduce that

$$T(\tilde{\beta} - \beta) \Rightarrow \left[(1/\omega_{11.2}) \int_0^1 B_2 B_2' \right]^{-1} \left[\int_0^1 B_2 dB_e \right] .$$

Since

$$\omega_{11.2} B_e(r) = B_{1.2}(r) = BM(\omega_{11.2})$$

the stated result (a) follows immediately.

The proof of part (b) follows similar lines. For the reasons advanced earlier $\hat{f}_{vv}(0)$ may be replaced by $f_{vv}(0)$ in the formula for $\tilde{\beta}_{(0)}$ giving

$$(A10) \quad T(\tilde{\beta}_{(0)} - \beta) - \left\{ \frac{1}{TM} \hat{f}_{22}(0) \right\}^{-1} \left\{ \frac{1}{M} \hat{f}_{2v}(0) f_{vv}^{-1}(0) e \right\} / e' f_{vv}^{-1}(0) e .$$

Using the Skorohod construction given in the proof of part (a) we have

$$\begin{aligned} \frac{1}{MT} \hat{f}_{22}(0) &= \frac{1}{2\pi M} \sum_{n=-M}^M k \left(\frac{n}{M} \right) \frac{1}{T} c_{22}^*(n) \\ &\xrightarrow{\text{a.s.}} \frac{1}{2\pi} \int_0^1 B_2^* B_2' \end{aligned}$$

as $T \rightarrow \infty$ from which we deduce that

$$(A11) \quad \frac{1}{MT} \hat{f}_{22}(0) \Rightarrow \frac{1}{2\pi} \int_0^1 B_2 B_2' .$$

Similarly we find

$$(A12) \quad \frac{1}{M} \hat{f}_{2v}^*(0) = \frac{1}{2\pi M} \sum_{k=-M}^M \binom{n}{M} c_{2v}^*(n) \\ \xrightarrow{\text{a.s.}} \frac{1}{2\pi} \int_0^1 B_2^* dB^{*'} + \Delta_2$$

where

$$\Delta_2 = \sum_{j=-\infty}^{\infty} E(u_{20} v_j')$$

To see this note that for fixed n we have

$$c_{2v}^*(n) \xrightarrow{\text{a.s.}} \int_0^1 B_2^* dB^{*'} + \Delta_2(n+1)$$

where

$$\Delta_2(n) = \sum_{j=0}^{\infty} E(u_{20} v_{j+n}') = \sum_{\ell=n}^{\infty} E(u_{20} v_{\ell}')$$

Since $\Delta_2(n) \rightarrow 0$ as $n \rightarrow \infty$ and $\Delta_2(n) \rightarrow \Delta_2$ as $n \rightarrow -\infty$ we find that Cesaro sum

$$\frac{1}{M} \sum_{k=-M}^M \Delta_2(n) \rightarrow \Delta_2, \quad M \rightarrow \infty$$

giving (A12). Hence

$$(A13) \quad \frac{1}{M} \hat{f}_{2v}^*(0) \Rightarrow \frac{1}{2\pi} \int_0^1 B_2^* dB^{*'} + \Delta_2$$

Next we observe that

$$(A14) \quad f_{vv}^{-1}(0)e = 2\pi\Omega^{-1}e$$

and

$$\begin{aligned}
\Delta_2 \Omega^{-1} e &= [0, I_m] (\sum_{j=-\infty}^{\infty} E(v_0 v_j')) \Omega^{-1} e \\
&= [0, I_m] \Omega \Omega^{-1} e \\
\text{(A15)} \quad &= [0 \quad I_m] e = 0 .
\end{aligned}$$

It then follows from (A10), (A11), (A13)-(A15) that

$$\begin{aligned}
T(\bar{\beta}_{(0)} - \beta) &= \left[\int_0^1 B_2 B_2' \right]^{-1} \int_0^1 B_2 dB' \Omega^{-1} e / e' \Omega^{-1} e \\
&= \left[\int_0^1 B_2 B_2' \right]^{-1} \int_0^1 B_2 dB_{1.2}
\end{aligned}$$

as required.

Proof of Theorem 3.2. This follows in the same way as the proof of part (a) of Theorem 3.1. We simply use (19) in place of the Fourier series for $f_{vv}^{-1}(\lambda)$.

Proof of Theorem 3.3. This is the same as the proof of Theorem 4.1 of Phillips (1988).

Proof of Theorem 3.4. This follows the same lines as the proof of part (a) of Theorem 3.1 above. We use the Fourier series

$$f_{v_1 v_1}^{-1}(\lambda) = (1/2\pi) \sum_{g=-\infty}^{\infty} d_g e^{ig\lambda}$$

and then

$$\begin{aligned} & \frac{1}{2MT} \sum_{j=-M+1}^M \hat{f}_{v_1 v_1}^{-1}(\lambda) \hat{f}_{22}(\omega_j) \\ & \Rightarrow \left(\frac{1}{2\pi}\right)^2 \left(\sum_{g=-\infty}^{\infty} d_g\right) \int_0^1 B_2 B_2' - (1/\omega_{11}) \int_0^1 B_2 B_2' \end{aligned}$$

since

$$(1/2\pi)^2 \sum_g d_g - (1/2\pi) \hat{f}_{v_1 v_1}^{-1}(0) = 1/\omega_{11} .$$

Next

$$\begin{aligned} & \frac{1}{2M} \sum_{j=-M+1}^M \hat{f}_{2v_1}(\omega_j) \hat{f}_{v_1 v_1}^{-1}(\omega_j) \\ & \Rightarrow \left(\frac{1}{2\pi}\right)^2 \int_0^1 B_2 dB_1(\sum_g d_g) + \left(\frac{1}{2\pi}\right)^2 \sum_g \Delta_{21}^{(g+1)} d_g \\ & = (1/\omega_{11}) \int_0^1 B_2 dB_1 + \delta_1 . \end{aligned}$$

The result stated now follows with $\delta = \omega_{11} \delta_1$.

REFERENCES

- Brillinger, D. R. (1974). *Time Series: Data Analysis and Theory*. New York: Holt, Rinehart and Winston.
- Brockwell, P. J. and R. A. Davis (1987). *Time Series: Theory and Methods*. New York: Springer.
- Davies, R. B. (1986). "Asymptotic inference when the amount of information is random," in L. M. LeCam and R. A. Olshen (eds.), *Proceedings of the Berkeley Conference in Honor of Jerzy Neyman and Jack Kiefer*, Vol. II. Wadsworth Inc.
- Engle, R. F. (1974). "Band spectrum regression," *International Economic Review*, 15, 1-11.
- _____ and C. W. J. Granger (1987). "Cointegration and Error Correction: Representation, Estimation and Testing," *Econometrica*, 55(87), 251-276.
- Hannan, E. J. (1963). "Regression for time series," in M. Rosenblatt (ed.), *Time Series Analysis*. New York: Wiley.
- _____ (1970). *Multiple Time Series*. New York: Wiley.
- _____ (1971). "Nonlinear time series regression," *Journal of Applied Probability*, 8, 767-80.
- _____ and P. M. Robinson (1973). "Lagged regression with unknown lags," *Journal of the Royal Statistical Society, Series B*, 35, 252-267.
- Journal of Economic Dynamics and Control* (1988), forthcoming.
- Jeganathan, P. (1980). "An extension of a result of L. LeCam concerning asymptotic normality," *Sankhya Series A*, 42, 146-160.
- _____ (1982). "On the asymptotic theory of estimation when the limit of the log-likelihood ratios is mixed normal," *Sankhya Series A*, 44, 173-212.
- LeCam, L. (1986). *Asymptotic Methods in Statistical Decision Theory*. New York: Springer.
- Oxford Bulletin of Economics and Statistics* (1986).
- Phillips, P. C. B. (1987). "Multiple regression with integrated processes," *AMS/IMS/SIAM Conference on Stochastic Processes Proceedings* (forthcoming).

- _____ (1988). "Optimal inference in cointegrated systems," Cowles Foundation Discussion Paper No. xxx, Yale University.
- _____ and S. N. Durlauf (1986). "Multiple time series with integrated variables," *Review of Economic Studies*, 53, 473-496.
- Robinson, P. M. (1972). "Nonlinear vector time series regression," *Journal of Applied Probability*, 9, xxx.
- Sharack, G. R. and J. A. Wellner (1986). *Empirical Processes with Applications to Statistics*. New York: Wiley.
- Stock, J. H. (1987). "Asymptotic properties of least squares estimators of cointegrating vectors," *Econometrica*, 55, 1035-1056.
- Whittle, P. (1951). *Hypothesis Testing in Time Series Analysis*. Uppsala, Sweden: Almqvist and Wicksell.