

**The Limiting Distribution of the Cointegration
Test Statistic in $VAR(n)$ Models**

by

Phoebus Dhrymes, Columbia University

October 1995

1994-95 Discussion Paper Series No. 748

The Limiting Distribution of the Cointegration Test Statistic in $VAR(n)$ Models

PHOEBUS J. DHRYMES*
Columbia University

October 9, 1995

Abstract

This paper obtains the limiting distribution of the trace test for cointegration in the context of the $VAR(n)$ model dealt with in Johansen (1988), (1991). The limiting distribution in question turns out to be that of a linear combination of mutually independent chi-squared variables all with the same degree of freedom parameter. The coefficients of the linear combination are characteristic roots of a certain positive definite matrix which may be estimated consistently.

Key Words: Cointegration, Cointegration test, characteristic roots, $VAR(n)$, Wishart distribution.

1 Introduction and Summary

Let $\{X_t : t \in \mathcal{N}\}$ be a stochastic sequence defined on some probability space $(\Omega, \mathcal{A}, \mathcal{P})$. If the sequence is taken to be $I(1)$, in the sense that $\{(I-L)X_t : t \in \mathcal{N}\}$ is **strictly stationary**, the question often arises as to whether the sequence in question is cointegrated. The latter means, in this context, that there exists a $q \times r$ matrix B of rank $r \leq q$ such that $X_t B$ is (strictly) stationary. In this paper we follow the derivation and arguments in Johansen (1988), (1991) and obtain the result that the trace test for cointegration, i.e. the test that the smallest $q - r$ roots of a certain determinantal equation are null, whose test statistic

* I wish to thank Ioannis Karatzas for a number of conversations, and Morris L. Eaton for helpful comments.

is the trace of a certain function of the original matrix is asymptotically distributed as the trace of a **Wishart** matrix. We show that this trace is distributed as a linear combination of chi-squared variables with $q - r$ degrees of freedom. Finally relying on earlier work by Box (1954), and Siddiqui (1965) we show the precise form of the distribution needed in order to carry out significance test. Although the form of this distribution **cannot be tabulated** owing to the fact that it depends on certain characteristic roots of an unknown matrix, still the form is quite simple and easily programmed to be used in conjunction with any estimation software dealing with such issues.

2 Notation and Problem Formulation

Consider the standard VAR

$$X_t \Pi(L) = \sum_{j=0}^n X_{t-j} \Pi_j = \epsilon_t, \quad t \geq 1, \quad \Pi_0 = I_q, \quad X_{-t} = 0, \quad t \geq 0, \quad (1)$$

where X_t is a q -element **row** vector, the error process being a multivariate white noise process with mean zero and covariance matrix $\Sigma > 0$, denoted by $MWN(\Sigma)$.

Remark 1. Although, as in the case of Johansen (1988) (1991), we shall formally assume the MWN process to be normal, this particular assumption merely motivates the form of the likelihood function. In fact, normality is quite irrelevant to most of our discussion, and the “likelihood function” may be thought of as the “objective function” of the problem much as in the so-called quasi-maximum likelihood literature of simultaneous equations. In this interpretation, the ϵ -process is $MWN(\Sigma)$ with unspecified distribution.

“Dividing” $\Pi(L)$ by $(I - L)$, where L is the usual lag operator we find, after some rearrangement,

$$(I - L)X_t = -X_{t-1} \Pi(1) + x_t \Pi^* + \epsilon_t, \quad (2)$$

where

$$x_t = (\Delta X_{t-1}, \Delta X_{t-2}, \dots, \Delta X_{t-n+1}), \quad t = 1, 2, \dots, T,$$

$$\Pi^* = (\Pi_1', \dots, \Pi_{n-1}')', \quad \Pi_j^* = \sum_{i=j+1}^n \Pi_i, \quad \Pi(1) = \sum_{j=0}^n \Pi_j.$$

It is assumed that

$$|\Pi(\phi)| = 0, \quad \text{where } \Pi(\phi) = \sum_{j=0}^n \Pi_j' \phi^j, \quad (3)$$

has r_0 **unit roots**, and the remaining roots are less than one in absolute value, so that the X -process is cointegrated of rank $r = q - r_0$. It is an easy consequence of the preceding that

$$(I - L)X'_t = \eta'_t, \quad \eta'_t = \sum_{j=0}^{\infty} A_j \epsilon'_{t-j}, \quad \|A_j\|^2 \sim c^j, \quad c \in (0, 1). \quad (4)$$

If the process is cointegrated of rank r then $\Pi(1)$ is of (reduced) rank $r < q$. Hence, by the rank factorization theorem, see Dhrymes (1984) p. 23, there exist matrices Γ , B both of dimension $q \times r$ and rank r such that $\Pi(1) = B\Gamma'$. We must now estimate the parameters of the model given T observations on $\{X_t : t = 1, 2, 3, \dots, T\}$; in order to give this problem the familiar look of systems of general linear models we define

$$\begin{aligned} y_t &= (I - L)X_t + X_{t-1}\Pi(1) = \Delta X_t + X_{t-1}B\Gamma'; \\ Y &= (y_t), \quad X = (x_t), \quad U = (\epsilon_t), \quad t = 1, 2, 3, \dots, T, \\ y &= \text{vec}(Y), \quad \pi^* = \text{vec}(\Pi^*), \quad u = \text{vec}(U), \quad \gamma = \text{vec}(\Gamma'). \end{aligned} \quad (5)$$

Thus, we may write

$$y_t = x_t\Pi^* + \epsilon_t, \quad Y = X\Pi^* + U, \quad \text{or} \quad y = (I_q \otimes X)\pi^* + u, \quad (6)$$

for a single observation and the entire sample, respectively, and define the estimation problem as: maximize the likelihood function with respect to the unknown parameters **subject to the condition** that B is $q \times r$ of **known rank** r . In view of the preceding, the loglikelihood (LF) of the T observations is given by

$$F_0 = -\frac{qT}{2}\ln(2\pi) - \frac{T}{2}\ln|\Sigma| - \frac{1}{2}\text{tr}(Y - X\Pi^*)\Sigma^{-1}(Y - X\Pi^*)'. \quad (7)$$

From Dhrymes (1984) p. 106, we have

$$\text{tr}(Y - X\Pi^*)\Sigma^{-1}(Y - X\Pi^*)' = [y - (I_q \otimes X)\pi^*]'(\Sigma^{-1} \otimes I_T)[y - (I_q \otimes X)\pi^*],$$

and solving the first order condition $(\partial F_0 / \partial \pi^*) = 0$, we find

$$\hat{\pi}^* = [I_q \otimes (X'X)^{-1}X']y. \quad (8)$$

Inserting this in Eq. (7), we obtain the concentrated LF

$$F_1 = -\frac{qT}{2}\ln(2\pi) - \frac{T}{2}\ln|\Sigma| - \frac{1}{2}y'(\Sigma^{-1} \otimes N)y, \quad N = I_T - X(X'X)^{-1}X'. \quad (9)$$

Next, we note that $y = \Delta p - (I_q \otimes P_{-1}B)\gamma$, where

$$\begin{aligned} P &= (X_t), \quad P_{-i} = (X_{t-i}), \quad t = 1, 2, 3, \dots, T, \quad i = 1, 2, 3, \dots, n-1 \\ p &= \text{vec}(P), \quad p_{-i} = \text{vec}(P_{-i}), \quad \gamma = \text{vec}(\Gamma') \end{aligned} \quad (10)$$

and the concentrated LF may be rendered as

$$\begin{aligned} F_1 &= -\frac{qT}{2}\ln(2\pi) - \frac{T}{2}\ln|\Sigma| \\ &\quad - \frac{1}{2}[\Delta p - (I_q \otimes P_{-1}B)\gamma]'(\Sigma^{-1} \otimes N)[\Delta p - (I_q \otimes P_{-1}B)\gamma]; \end{aligned} \quad (11)$$

solving the first order conditions $(\partial F_1/\partial \gamma) = 0$, we find

$$\hat{\gamma} = [I_q \otimes (B'P_{-1}'NP_{-1}B)^{-1}B'P_{-1}'N]\Delta p. \quad (12)$$

Noting that

$$\begin{aligned} \phi &= (I_q \otimes N)[\Delta p - (I_q \otimes P_{-1}B)\hat{\gamma}] = (I_q \otimes N^*)(I_q \otimes N)\Delta p, \\ N^* &= I_T - NP_{-1}B(B'P_{-1}'NP_{-1}B)^{-1}B'P_{-1}'N, \end{aligned} \quad (13)$$

we may write the (once again) concentrated LF as

$$\begin{aligned} F_2 &= -\frac{qT}{2}\ln(2\pi) - \frac{T}{2}\ln|\Sigma| \\ &\quad - \frac{1}{2}[(I_q \otimes N^*N)\Delta p]'(\Sigma^{-1} \otimes I_T)[(I_q \otimes N^*N)\Delta p]. \end{aligned} \quad (14)$$

Using the results in Dhrymes (1984) p. 106 and “rematricizing” the last expression in the LF, we may rewrite the latter as

$$F_2 = -\frac{qT}{2}\ln(2\pi) - \frac{T}{2}\ln|\Sigma| - \frac{T}{2}\text{tr}\Sigma^{-1}S, \quad S = \frac{1}{T}(N\Delta P)'N^*(N\Delta P). \quad (15)$$

Maximizing F_2 with respect to the elements of Σ^{-1} , we obtain

$$\frac{\partial F_2}{\partial \text{vec}(\Sigma^{-1})} = \frac{T}{2}\text{vec}(\Sigma)' - \frac{T}{2}\text{vec}(S) = 0, \quad \text{or} \quad \hat{\Sigma} = S. \quad (16)$$

Inserting this in Eq. (14), we find the ultimately concentrated LF

$$F_3 = -\frac{qT}{2}[\ln(2\pi) + 1] - \frac{T}{2}\ln|S|, \quad (17)$$

which is now to be **maximized with respect to B** .

To continue, it is convenient to define

$$W = N\Delta P = N(P - P_{-1}), \quad V = NP_{-1}, \quad (18)$$

and to omit the term $(1/T)$ from the definition of S ; in this notation the maximization of F_3 in Eq. (17) is **equivalent to the minimization of**

$$\begin{aligned} D(B) &= |S| = |W'W - W'VB(B'V'VB)^{-1}B'V'W| \quad (19) \\ &= |W'W| |I_q - J|, \quad J = W'VB(B'V'VB)^{-1}B'V'W(W'W)^{-1}. \end{aligned}$$

From Dhrymes (1989) p. 39, we have that

$$|I_q - J| = |B'V'VB|^{-1} |B'[V'V - V'W(W'W)^{-1}W'V]B|. \quad (20)$$

Since the function $D(B)$ is **homogeneous of degree zero in B** , we need to impose a normalization, else the problem is not well defined, or there is an infinitude of solutions. A convenient normalization, under the circumstances, is $B'V'VB = I_r$, so that the quantity to be minimized, with respect to B , is

$$D(B) = |W'W| |B'[V'V - V'W(W'W)^{-1}W'V]B|. \quad (21)$$

To do so we proceed in a roundabout fashion, abstracting for the moment from the probabilistic aspects of the problem. Consider the characteristic roots of $V'W(W'W)^{-1}W'V$ in the metric of $V'V$, i.e consider the equation

$$|\lambda V'V - V'W(W'W)^{-1}W'V| = 0. \quad (22)$$

Since the matrix $V'V - V'W(W'W)^{-1}W'V$ is **positive definite**, all such characteristic roots are (positive and) less than unity, see Dhrymes (1984) p.75. Let the roots $\{\lambda_j : j = 1, 2, 3, \dots, q\}$ be arranged in **decreasing magnitude**, let the corresponding characteristic vectors be $\{g_j : j = 1, 2, 3, \dots, q\}$, define

$$\Lambda = \text{diag}(\lambda_1, \lambda_2, \lambda_3, \dots, \lambda_q), \quad G = (g_{\cdot 1}, g_{\cdot 2}, g_{\cdot 3}, \dots, g_{\cdot q}), \quad (23)$$

and impose the normalization $G'V'VG = I_q$. Since

$$V'W(W'W)^{-1}W'VG = V'VGA, \quad (24)$$

we conclude that

$$D(G) = |W'W| |G'[V'V - V'W(W'W)^{-1}W'V]G| = |W'W| |I_q - \Lambda|. \quad (25)$$

Thus, the estimator we seek, \hat{B} , must be a subset of the columns of the matrix G , it must be of rank r , and must minimize $D(G)$. The solution is then obvious, **we must choose \hat{B} to be the submatrix $G_{(r)}$ containing the r characteristic vectors, corresponding to the r largest roots.** Thus, the maximized value of the LF is given by

$$F_4(\hat{B}; r) = \frac{qT}{2} [\ln(2\pi) + 1] - \frac{T}{2} \left| \frac{W'W}{T} \right| - \frac{T}{2} \sum_{j=1}^r \ln(1 - \hat{\lambda}_{(j)}), \quad (26)$$

where the notation $\hat{\lambda}_{(j)}, j = 1, 2, \dots, q$, indicates the **ordered roots**.

It is easily demonstrated that under the alternative of no cointegration,

$$H_1 : \text{rank}[\Pi(1)] = q,$$

the maximum of the LF is given by

$$F_4(\hat{B}; q) = \frac{qT}{2} [\ln(2\pi) + 1] - \frac{T}{2} \left| \frac{W'W}{T} \right| - \frac{T}{2} \sum_{j=1}^q \ln(1 - \hat{\lambda}_{(j)}). \quad (27)$$

so that the test of this hypothesis involves the question of whether the smallest $q - r$ roots are zero, and the log of the likelihood ratio test statistic is given by

$$\tau^* = \frac{T}{2} \sum_{j=r+1}^q \ln(1 - \hat{\lambda}_{(j)}). \quad (28)$$

Since the matrix whose characteristic roots are involved in the test is at worst positive semidefinite, an equivalent test is in terms of the alternative

$$H_1 : \sum_{j=r+1}^q \lambda_{(j)} = 0,$$

which will yield the test statistic

$$\hat{\tau} = T \sum_{j=r+1}^q \hat{\lambda}_{(j)}. \quad (29)$$

3 The Limiting Distribution of $\hat{\tau}$

To investigate this problem consider the matrix

$$S(\lambda) = \lambda \frac{V'V}{T} - \frac{V'W}{T} \left(\frac{W'W}{T} \right)^{-1} \frac{W'V}{T}.$$

Let G_2 be an **arbitrary** $q \times q - r$ matrix whose columns are linearly independent, such that (B, G_2) is a nonsingular matrix, and define

$$C_T = \left(B, \frac{1}{\sqrt{T}} G_2 \right), \quad S^*(\lambda) = C_T' S(\lambda) C_T. \quad (30)$$

It is easily verified that the characteristic roots of interest in our discussion are the roots of $|S^*(\lambda)| = 0$ and, moreover,

$$S^*(\lambda) = \begin{bmatrix} S_{11}^*(\lambda) & S_{12}^*(\lambda) \\ S_{21}^*(\lambda) & S_{22}^*(\lambda) \end{bmatrix}, \quad S_{21}^*(\lambda) = S_{12}^{*'}(\lambda) \quad \text{where} \quad (31)$$

$$S_{11}^*(\lambda) = \lambda \frac{B' P_{-1}' N P_{-1} B}{T} - \frac{B' V' W}{T} \left(\frac{W' W}{T} \right)^{-1} \frac{W' V B}{T},$$

$$S_{12}^*(\lambda) = \lambda \frac{B' V' V G_2}{T^{3/2}} - \frac{B' V' W}{T} \left(\frac{W' W}{T} \right)^{-1} \frac{W' V G_2}{T^{3/2}},$$

$$S_{22}^*(\lambda) = \lambda G_2' \frac{V' V}{T^2} G_2 - \frac{G_2' V' W}{T^{3/2}} \left(\frac{W' W}{T} \right)^{-1} \frac{W' V G_2}{T^{3/2}}.$$

It may be easily shown that $S_{12}^*(\lambda) \xrightarrow{d} 0$ and, asymptotically,

$$\frac{G_2' V' V G_2}{T^2} \sim \frac{G_2' P_{-1}' P_{-1} G_2}{T^2} \xrightarrow{d} G_2' P_0' \left(\int_0^1 B(s)' B(s) ds \right) P_0 G_2, \quad (32)$$

where P_0 is a triangular decomposition of the matrix Σ_0 , i.e.

$$P_0' P_0 = \Sigma_0 = \Sigma_1 + \Sigma_1^* + \Sigma_1^{*'} = E \eta_1' \eta_1 + \sum_{j=1}^{\infty} E \eta_1' \eta_{1+j} + \sum_{j=1}^{\infty} E \eta_{1+j}' \eta_1, \quad (33)$$

$B(s)$ being a SMBM¹ of dimension q . We also note

$$S_{11}^*(\lambda) = \lambda \frac{Z_{-1}' N Z_{-1}}{T} - \frac{Z_{-1}' N \Delta P}{T} \left(\frac{(\Delta P)' N \Delta P}{T} \right)^{-1} \frac{(\Delta P)' N Z_{-1}}{T},$$

where $Z = (z_t)$, and $z_t = X_t B$ is the **cointegral vector**. It follows, therefore, that

$$\frac{Z_{-1}' N Z_{-1}}{T} = \frac{Z_{-1}' Z_{-1}}{T} - \frac{Z_{-1}' X}{T} \left(\frac{X' X}{T} \right)^{-1} \frac{X' Z_{-1}}{T}$$

$$\xrightarrow{\text{a.c.}} M_{zz} - M_{zx} M_{xx}^{-1} M_{xz} = M_{zz}^*, \quad (34)$$

¹ In the notation and usage of this paper a standard multivariate Brownian Motion (SMBM) is a q -element row vector Brownian Motion such that $EB(s)' B(s) = sI_q$. A multivariate Brownian Motion (MBM) with covariance matrix $s\Phi$, is written as BP_0 , such that P_0 is the (unique) triangular matrix of the decomposition $\Phi = P_0' P_0$. For triangular decompositions of this type see Dhrymes (1984), pp. 68-69.

where M_{zz}^* denotes the (conditional) covariance matrix of the cointegral vector z_{t-1} . **given** x_t . In a similar fashion we may show that

$$\frac{Z'_{-1}N\Delta P}{T} \sim - \left[\frac{Z'_{-1}Z_{-1}}{T} - \frac{Z'_{-1}X}{T} \left(\frac{X'X}{T} \right)^{-1} \frac{X'Z_{-1}}{T} \right] \Gamma' \xrightarrow{\text{a.c.}} -M_{zz}^* \Gamma'$$

$$\frac{(\Delta P)'N\Delta P}{T} \xrightarrow{\text{a.c.}} \Sigma + \Gamma M_{zz}^* \Gamma' = M_{ww}, \quad \text{so that} \quad (35)$$

$$S_{11}^*(\lambda) \xrightarrow{\text{a.c.}} \lambda M_{zz}^* - M_{zz}^* \Gamma' M_{ww}^{-1} \Gamma M_{zz}^*,$$

$$S_{22}^*(\lambda) \xrightarrow{d} \lambda G'_2 P'_0 K P_0 G_2 = \lambda \bar{S}_{22}, \quad K = \int_0^1 B(s)' B(s) ds.$$

We therefore conclude that the characteristic roots of $|S(\lambda)| = 0$ converge, in distribution or a.c., to the characteristic roots of $|\bar{S}^*(\lambda)| = 0$, where

$$\bar{S}^*(\lambda) = \begin{bmatrix} \lambda M_{zz}^* - M_{zz}^* \Gamma' M_{ww}^{-1} \Gamma M_{zz}^* & 0 \\ 0 & \lambda \bar{S}_{22} \end{bmatrix}, \quad (36)$$

and it is evident that **there are at least** $q - r = r_0$ **zero roots**; to complete the argument we must show that the roots of

$$|\lambda M_{zz}^* - M_{zz}^* \Gamma' M_{ww}^{-1} \Gamma M_{zz}^*| = 0$$

are **nonnull**, so that the number of nonzero roots corresponds to the number of cointegrating vectors. Since the cointegral vector z_{t-1} and x_t are evidently **not linearly dependent**, $M_{zz}^* > 0$; moreover, since equally evidently $M_{zz}^* \Gamma' M_{ww}^{-1} \Gamma M_{zz}^* > 0$, we conclude from Proposition 62 in Dhrymes (1984) p. 74 that the roots in question obey $0 < \lambda_j < 1$, for $j = 1, 2, \dots, r$ and $\lambda_j = 0$ for $j = r + 1, r + 2, \dots, q$.

In view of the fact that the characteristic roots of $S(\lambda)$ are continuous functions of its elements, its ordered characteristic roots obey, asymptotically, the relations implied by Eq. (36), i.e. the largest r roots have positive a.c. limits, while the smallest $q - r$ roots have zero limits (in distribution).

Remark 2. Since we have employed the normalization $\hat{B}' V' V \hat{B} / T = I_r$, we are led to the conclusion that $M_{zz}^* = I_r$. Thus, the limiting form of the characteristic equation yielding the characteristic roots is given by

$$0 = |\lambda I_r - \Gamma' (\Sigma + \Gamma \Gamma')^{-1} \Gamma| |\lambda G'_2 P'_0 K P_0 G_2|, \quad (37)$$

which clearly shows that the r positive roots come from the first factor, while the $q - r$ zero roots come from the second factor. Finally, using

the partitioned inverse form, see Dhrymes (1984) p. 39, we conclude that the r largest characteristic roots of relevance to our discussion, $\hat{\lambda}_j$, $j = 1, 2, \dots, r$ converge to entities, λ_j , which are related to the roots of

$$0 = |(1 - \lambda)I_r - (I_r + \Gamma' \Sigma^{-1} \Gamma)^{-1}| = |\mu I_r - \Gamma' \Sigma^{-1} \Gamma|, \quad (38)$$

by $\mu_j = \lambda_j / (1 - \lambda_j)$. In view of the fact that $\Gamma' \Sigma^{-1} \Gamma > 0$, $\mu_j > 0$, for all j , and thus $0 < \lambda_j < 1$. Since $\Gamma' \Sigma^{-1} \Gamma$ is a **nonsingular matrix** of dimension equal to the **cointegration rank**, we see that the number of the characteristic roots of $S(\lambda)$ which, in the limit, are positive is precisely equal to the rank of cointegration.

We now turn to the question of the limiting distribution of the smallest $q - r$ characteristic roots; thus, we must determine the limiting distribution of $T\hat{\lambda}_{(j)} = \hat{\rho}_j$, for $j = r + 1, r + 2, \dots, q$. To this end, note first that

$$S_{22}^*(T^{-1}\rho) - S_{21}^*(T^{-1}\rho)[S_{21}^*(T^{-1}\rho)]^{-1}S_{12}^*(T^{-1}\rho) \xrightarrow{d} 0, \text{ but}$$

$$S^{**}(\rho) = TS_{22}^*(T^{-1}\rho) - TS_{21}^*(T^{-1}\rho)[S_{11}^*(T^{-1}\rho)]^{-1}S_{12}^*(T^{-1}\rho) \xrightarrow{d} \Upsilon(\rho),$$

where $\Upsilon(\rho)$ is a well defined matrix to be determined below, and moreover

$$\begin{aligned} S_{11}^*(T^{-1}\rho) &= \frac{\rho B' V' V B}{T^2} - \frac{B' V' W}{T} \left(\frac{W' W}{T} \right)^{-1} \frac{W' V B}{T} \\ &\xrightarrow{\text{a.c.}} -M_{zz}^* \Gamma' M_{ww}^{-1} \Gamma M_{zz}^* \\ TS_{22}^*(\rho) &= \rho \frac{G_2' V' V G_2}{T^2} - \frac{G_2' V' W}{T} \left(\frac{W' W}{T} \right)^{-1} \frac{W' V G_2}{T} \\ &\sim \rho G_2' P_0' K P_0 G_2 - \frac{G_2' V' W}{T} M_{ww}^{-1} \frac{W' V G_2}{T} \\ \sqrt{T} S_{12}^*(T^{-1}\rho) &\sim -M_{zz}^* \Gamma' M_{ww}^{-1} \frac{W' V G_2}{T}. \end{aligned} \quad (39)$$

Consequently,

$$\begin{aligned} S^{**}(\rho) &\sim \rho G_2' P_0' K P_0 G_2 - \left(\frac{G_2' V' W}{T} \right) F \left(\frac{W' V G_2}{T} \right), \\ F &= M_{ww}^{-1} - M_{ww}^{-1} \Gamma' (\Gamma' M_{ww}^{-1} \Gamma)^{-1} \Gamma' M_{ww}^{-1}. \end{aligned} \quad (40)$$

Since $V'W = P'_{-1}[NU - NP_{-1}B\Gamma']$ and $\Gamma'F = 0$, we need only be concerned about the limiting distribution of $P'_{-1}NU/T$, which obeys

$$\frac{P'_{-1}NU}{T} \xrightarrow{d} P'_0K_1T_1, \quad K_1 = \int_0^1 B(s)' dB_1(s), \quad T'_1T_1 = \Sigma,$$

where $B(s)$ and $B_{(1)}(s)$ are two mutually independent SMBM of dimension q . Finally, we obtain

$$S^{**}(\rho) \xrightarrow{d} \Upsilon(\rho) = \rho G'_2 P'_0 K P_0 G_2 - G'_2 P'_0 K_1 T_1 F T'_1 K'_1 P_0 G_2, \quad (41)$$

and thus the limiting distribution of the ordered roots $T\hat{\lambda}_j$, $j = r + 1, r + 2, \dots, q$ of $|S(\lambda)| = 0$ is given by the distribution of the ordered roots ρ_j , $j = 1, 2, \dots, q - r$, of

$$0 = \left| \rho G'_2 P'_0 \left(\int_0^1 B(s)' B(s) ds \right) P_0 G_2 - G'_2 P'_0 \left(\int_0^1 B(s)' dB_{(1)}(s) \right) T_1 F T'_1 \right. \\ \left. \left(\int_0^1 dB_{(1)}(s)' B(s) \right) P_0 G_2 \right|. \quad (42)$$

We have therefore proved

Proposition 1. In the context of the discussion above, let $\rho = T\lambda$ and consider the $q - r$ smallest roots of $|S(\lambda)| = 0$. The limiting distribution of such roots is the distribution of the roots of

$$|\Upsilon(\rho)| = |\rho G'_2 P'_0 K P_0 G_2 - G'_2 P'_0 K_1 T_1 F T'_1 K'_1 P_0 G_2| = 0.$$

Remark 3. A close examination of the conclusion of the proposition above indicates that this could not possibly be the **final** result since it would imply that the test for cointegration rank depends on **the arbitrary matrix** G_2 . Indeed, we may simplify the representation above, as well as the conclusion of the proposition. To this end, note that the characteristic roots of

$$|\mu M_{ww}^{-1} - M_{ww}^{-1} \Gamma (\Gamma' M_{ww}^{-1} \Gamma)^{-1} \Gamma' M_{ww}^{-1}| = 0,$$

consist of r unities and $q - r$ zeros. Consequently, using Proposition 63 in Dhrymes (1984) p. 75, we may write

$$F = F^* \left(I_q - \begin{bmatrix} 0 & 0 \\ 0 & I_r \end{bmatrix} \right) F'^* = F_1 F'_1, \quad F^* = (F_1, F_2), \quad \text{rank}(F) = q - r, \quad (43)$$

where F_1 is a $q \times q - r$ matrix of rank $q - r$.

Taking into account the discussion in the remark above we may simplify the representation of Proposition 1, as follows. In the first term the MBM is BP_0G_2 whose covariance matrix is given by $sG_2'\Sigma_0G_2 > 0$, which is of dimension $q - r$. Thus, define the SMBM $B_{(2)}$ such that

$$B_{(2)}(s)T_2 = B(s)P_0G_2, \quad G_2'\Sigma_0G_2 = T_2'T_2, \quad (44)$$

where T_2 is the triangular decomposition corresponding to the SMBM $B_{(2)}$, which is of dimension $q - r$. Similarly in the integral

$$G_2'P_0' \left(\int_0^1 B(s)' dB_{(1)}(s) \right) T_1F_1$$

we have integration with respect to the MBM $B_{(1)}(s)T_1F_1$, whose covariance matrix is $sF_1'\Sigma F_1 > 0$, of dimension $q - r$. Define

$$B_{(3)}(s)T_3 = B_{(1)}(s)T_1F_1, \quad F_1'\Sigma F_1 = T_3'T_3 \quad (45)$$

where T_3 is the triangular decomposition of $F_1'\Sigma F_1$, corresponding to the SMBM $B_{(3)}$, of dimension $q - r$. We further note that $B_{(2)}$ and $B_{(3)}$ are **mutually independent**. It follows, therefore, that the limit $\Upsilon(\rho)$ of Eq. (41) may be rendered as

$$\begin{aligned} \Upsilon(\rho) &= \rho T_2' \left(\int_0^1 B_{(2)}(s)' B_{(2)}(s) ds \right) T_2 - \left[T_2' \left(\int_0^1 B_{(2)}(s)' dB_{(3)}(s) \right) T_3 \right] \\ &\quad \times \left[T_2' \left(\int_0^1 B_{(2)}(s)' dB_{(3)}(s) \right) T_3 \right]' \end{aligned}$$

$$T_4'T_4 = \int_0^1 B_{(2)}(s)' B_{(2)}(s) ds.$$

Utilizing the results in Dhrymes (1995), Example 5 of Chapter 4, we obtain that

$$T_4'^{-1}T_2'^{-1}\Upsilon(\rho)T_2^{-1}T_4^{-1} = \rho I_{q-r} - XT_3T_3'X',$$

$$X = T_4'^{-1} \int_0^1 B_{(2)}(s)' dB_{(3)}(s) \quad (46)$$

and, moreover, that the rows of X are $N(0, I_{q-r})$ and **mutually independent**. Indeed, all elements of X are i.i.d. unit normal variables. Since, if both matrices are square, the characteristic roots of $|\lambda I - AB| = 0$ are identical to those of $|\lambda I - BA| = 0$ we have that the roots of interest in our discussion are given by

$$\begin{aligned} 0 &= |T_4^{-1}T_2'^{-1}\bar{S}^{**}(\rho)T_2^{-1}T_4'^{-1}| = |\rho I_{q-r} - YY'| \\ &= |\rho I_{q-r} - Y'Y|, \quad Y = XT_3. \end{aligned} \quad (47)$$

We note that the j^{th} row of Y is given by $y_j = x_j T_3$. Since the rows of X are i.i.d. $N(0, I_{q-r})$, it follows that

$$y'_j = T'_3 x'_j \sim N(0, T'_3 T_3), \quad (48)$$

and are mutually independent for $j = 1, 2, \dots, q - r$. Since

$$Y'Y = \sum_{j=1}^{q-r} y'_j y_j, \quad (49)$$

we conclude that

$$Y'Y \sim \mathbf{w}(\Phi, q - r), \quad \Phi = T'_3 T_3 = F'_1 \Sigma F_1, \quad (50)$$

where $\mathbf{w}(\Phi, q - r)$ indicates the **Wishart** distribution with parameters Φ and $q - r$, which evidently **does not depend** on the arbitrary matrix G_2 . For a discussion of the properties of the Wishart distribution see Eaton (1983), Chapter 8. We conclude, therefore, that the smallest $q - r$ (ordered) characteristic roots of $S(\lambda)$ converge in distribution to the ordered characteristic roots of a Wishart matrix with parameters Φ and $q - r$. This is a standard distribution but it does contain the nuisance parameter Φ .

We have therefore proved

Proposition 2. In the context of Proposition 1, the $q - r$ smallest (ordered) characteristic roots of $|S(\lambda)| = 0$ normalized by T , i.e. $T\hat{\lambda} = \hat{\rho}$, converge in distribution to the ordered characteristic roots of

$$0 = |\rho I_{q-r} - Y'Y| = 0, \quad Y = T_4^{-1} \left(\int_0^1 B_{(2)}(s)' dB_{(3)}(s) \right) T_3$$

$$T'_4 T_4 = \int_0^1 B_{(2)}(s)' B_{(2)}(s) ds, \quad T'_3 T_3 = F'_1 \Sigma F_1,$$

$$Y'Y \sim \mathbf{w}(T'_3 T_3, q - r).$$

Remark 4. The preceding discussion has established that the limiting distribution of the $q - r$ smallest roots of $|S(\lambda)| = 0$, i.e. the limiting distribution of $T\lambda_j$, $j = r+1, r+2, \dots, q$, under the null of cointegration, is the distribution of the roots of a Wishart matrix with parameters $\Phi = F'_1 \Sigma F_1$ and $q - r$. Since the test statistic for the test of the existence of cointegration is that the sum of the smallest $q - r$ roots is zero, we may utilize the result we have just established to formulate a test based on that limiting distribution. To accomplish this we need to determine

the distribution of the trace of a Wishart matrix with the characteristics above. To this end, we have

Lemma 1. Let

$$S = \sum_{j=1}^N X_j' X_j, \quad X_j \sim N(0, \Sigma), \quad j = 1, 2, \dots, N$$

the X_j being n -element **row** vectors and independent indentially distributed. Then

$$\text{tr} S \sim \sum_{i=1}^n r_i \zeta_i,$$

where the ζ_i are i.i.d. chi-squared variables with N degrees of freedom, and r_i are the characteristic roots of Σ .

Proof: Since $\Sigma > 0$ it is orthogonally similar to the diagonal matrix

$$R = \text{diag}(r_1, r_2, \dots, r_n), \quad \Sigma = QRQ',$$

where Q is the orthogonal matrix of the corresponding characteristic vectors. We note that

$$Q' S Q = \sum_{j=1}^N Q' X_j' X_j Q = \sum_{j=1}^N R^{1/2} Z_j' Z_j R^{1/2},$$

where $Z_j = R^{-(1/2)} Q' X_j \sim N(0, I_n)$, so that the Z_j are i.i.d. $N(0, I_n)$. Consequently,

$$\text{tr} S = \text{tr} Q' S Q = \sum_{j=1}^N Z_j' R Z_j = \sum_{i=1}^n r_i \left(\sum_{j=1}^N z_{ji}^2 \right).$$

Since the entities in parentheses in the rightmost member above are i.i.d. chi-squared variables each with N degrees of freedom and the r_i are the characteristic roots of Σ , the proof is completed.

q.e.d.

Remark 5. Unfortunately the chi-squared distribution does not have the reproductive properties of the normal distribution under linear combinations. Thus, $\text{tr} S$ is not necessarily a chi-squared variable, not even proportional to a chi-squared variable **unless** $r_j = r_1$, for all j . Thus, the distribution of the test statistic of Eq. (29) obeys

$$\hat{\tau} \stackrel{d}{\rightarrow} \tau \sim \sum_{j=1}^{q-r} r_j \zeta_j, \quad (51)$$

where $r_j, j = 1, 2, \dots, q - r$ are the characteristic roots of the positive definite matrix $\Phi = F_1' \Sigma F_1$ and the ζ_j are i.i.d. chi-squared variables with $q - r$ degrees of freedom.

To determine the nature of the limiting distribution above, consider the moment generating function of τ , say

$$h^*(s) = E e^{-s\tau} = \prod_{j=1}^{q-r} E e^{-sr_j \zeta_j} = \prod_{j=1}^{q-r} (1 + 2sr_j)^{-(q-r)/2}. \quad (52)$$

Since, if h is the density function of τ

$$h^*(s) = \int_0^\infty e^{-s\xi} h(\xi) d\xi, \quad (53)$$

it follows that h^* is simply the Laplace transform of the desired density function h . What is required in order to carry out the cointegration test is the function

$$H(x) = \Pr(\tau > x) = \int_x^\infty h(\xi) d\xi. \quad (54)$$

It is easily shown that the Laplace transform of H is given by

$$\begin{aligned} H^*(s) &= \int_0^\infty e^{-s\xi} H(\xi) d\xi \\ &= -\frac{1}{s} e^{-s\xi} H(\xi) \Big|_0^\infty - \int_0^\infty e^{-s\xi} h(\xi) d\xi = \frac{1}{s} [1 - h^*(s)]. \end{aligned} \quad (55)$$

The significance of the preceding discussion is evident from a result due to Box (1954a), (1954b), also given in somewhat more general form in Siddiqui (1965), as follows:

Theorem 1. If the Laplace transform of a density function h is representable as

$$h^*(s) = \sum_{j=1}^m k_j (1 + d_j s)^{-n_j} \quad (56)$$

then

$$H(x) = \sum_{j=1}^m k_j G(x; d_j, n_j), \quad (57)$$

$$G(x; d_j, n_j) = \frac{1}{\Gamma(n_j)} \int_x^\infty e^{-(\xi/d_j)} \left(\frac{\xi}{d_j}\right)^{n_j-1} d(\xi/d_j).$$

Remark 6. The usefulness of Theorem 1 is further enhanced by the fact that repeated integrations by part yields

$$G(x; d_i, n_i) = e^{-(x/d_i)} \sum_{j=0}^{n_i-1} \frac{(x/d_i)^j}{j!}. \quad (58)$$

Consider now the Laplace transform of the density of τ as exhibited in Eq. (52). We have

Proposition 3. In the context of Propositions 1 and 2, consider the test of the hypothesis

H_0 : the X – process is cointegrated of rank r

as against the alternative

H_1 : the X – process is not cointegrated ,

through the use of the test statistic

$$\hat{\tau} = T \sum_{j=r+1}^q \hat{\lambda}_{(j)}.$$

The critical region of size α for the test of the hypothesis above may be determined, for $q - r$ **even**, from the relation

$$\alpha = H(x_\alpha) = \sum_{j=1}^{q-r} \sum_{i=1}^{(q-r)/2} k_{ji} G(x_\alpha; d_j, i), \quad (59)$$

where

$$G(x; d_i, n_i) = e^{-(x/d_i)} \sum_{j=0}^{n_i-1} \frac{(x/d_i)^j}{j!}.$$

Proof: From Eq. (52) we have

$$h^*(s) = \prod_{j=1}^{q-r} (1 + d_j s)^{-(q-r)/2}, \quad d_j = 2r_j.$$

If $q - r$ is even, $m = (q - r)/2$ is an integer, so that each term has the partial fraction expansion

$$(1 + d_j s)^{-m} = \sum_{i=1}^m \frac{k_{ji}}{(1 + d_j s)^i}, \quad (60)$$

so that

$$h^*(s) = \sum_{j=1}^{q-r} \sum_{i=1}^m \frac{k_{ji}}{(1+d_j s)^i}. \quad (61)$$

In view of Theorem 1 we conclude that

$$H(x) = \sum_{j=1}^{q-r} \sum_{i=1}^n k_{ji} G(x; d_j, i). \quad (62)$$

q.e.d.

Remark 7. If $q-r$ is odd, $(q-r)/2$ is not an integer and Theorem 1 is not directly available. Nonetheless we may define bounding functions $H_1(x) < H(x) < H_2$ by the following procedure: let $q-r = 2m-1$ and define

$$\begin{aligned} \tau_1 &= \sum_{j=1}^{m-1} r_{2j+1}(\zeta_{2j-1} + \zeta_{2j}), \\ \tau_2 &= \sum_{j=1}^{m-1} r_{2j-1}(\zeta_{2j-1} + \zeta_{2j}) + r_{2m-1}(\zeta_{2m-1} + \zeta_{2m}^*), \end{aligned} \quad (63)$$

where z_{2m}^* is another chi-squared variable with $q-r$ degrees of freedom, independent of all others. Since the roots are ordered (in decreasing order) and since chi-squared variables are nonnegative, for any $\omega \in \Omega$

$$\tau_1(\omega) \leq \tau(\omega) \leq \tau_2(\omega).$$

If H_1, H, H_2 are the corresponding tail functions we may approximate H by H_θ , where $\theta \in (0, 1)$ and

$$H_\theta(x) = \theta H_1(x) + (1-\theta)H_2(x). \quad (64)$$

The approximation may be carried out by minimum distance methods, i.e. by defining a suitable metric and norm on an appropriate Hilbert space. The problem then would be

$$\min_{\theta \in (0,1)} \| H(x) - \theta H_1(x) - (1-\theta)H_2(x) \|^2.$$

Other approaches are also available, such as for example matching moments of H_θ to those of H .

We conclude our discussion with

Example 1. Suppose $q - r = 2$. Then we have

$$h^*(s) = \left(\frac{1}{1 + sd_1} \right) \left(\frac{1}{1 + sd_2} \right) = \frac{k_1}{1 + sd_1} + \frac{k_2}{1 + sd_2},$$

where $k_1 = d_1/(d_1 - d_2)$ and $k_2 = -d_2/(d_1 - d_2)$. Consequently, we obtain

$$H(x) = k_1 e^{-(x/d_1)} + k_2 e^{-(x/d_2)},$$

and the α -level of significance critical point may be easily found by solving the equation

$$\alpha = k_1 e^{-(x/d_1)} + k_2 e^{-(x/d_2)}$$

for x_α .

Example 2. Suppose in Example 1 $q - r = 3$. In this case $\tau = r_1 \zeta_1 + r_2 \zeta_2 + r_3 \zeta_3$. Since

$$h^*(s) = \prod_{j=1}^3 (1 + sd_j)^{-(3/2)}, \quad d_j = 2r_j,$$

h^* cannot be expanded by partial fractions. To overcome this define

$$\tau_1 = r_3(\zeta_1 + \zeta_2), \quad \tau_2 = r_1(\zeta_1 + \zeta_2) + r_3(\zeta_3 + \zeta_4^*),$$

where ζ_4^* is a an arbitrary chi-squared variable with 3 degrees of freedom, and independent of ζ_i , $i = 1, 2, 3$. It follows immediately that

$$h_1^*(s) = \left(\frac{1}{1 + sd_3} \right)^3, \quad h_2^*(s) = \left(\frac{1}{1 + sd_1} \right)^3 \left(\frac{1}{1 + sd_3} \right)^3,$$

and consequently we obtain the partial fraction expansion and the tail functions

$$h_1^*(s) = \sum_{i=1}^3 \frac{k_{i1}}{(1 + sd_3)^i},$$

$$h_2^*(s) = \sum_{i=1}^3 \frac{k_{i2}}{(1 + sd_1)^i} + \sum_{i=4}^6 \frac{k_{i2}}{(1 + sd_3)^{(i-3)}};$$

$$H_1(x) = e^{-(x/d_3)} \left[\sum_{i=1}^3 k_{i1} + (k_{21} + k_{31}) \frac{x}{d_3} + \frac{k_{31}}{2} \left(\frac{x}{d_3} \right)^2 \right],$$

$$H_2(x) = e^{-(x/d_1)} \left[\sum_{i=1}^3 k_{i2} + (k_{22} + k_{32}) \frac{x}{d_1} + \frac{k_{32}}{2} \left(\frac{x}{d_1} \right)^2 \right] \\ + e^{-(x/d_3)} \left[\sum_{i=4}^6 k_{i2} + (k_{52} + k_{62}) \frac{x}{d_3} + \frac{k_{62}}{2} \left(\frac{x}{d_3} \right)^2 \right].$$

The two tail functions determined above bracket the correct tail function; thus, if we are content to follow the procedure we follow in the Durbin-Watson test, we have a very simple procedure for determining the critical values of the trace cointegration test. If we are not so content, at least in the case of odd degrees of freedom we may devise a suitable approximation procedure.

REFERENCES

- Box, G. E. P. (1954a), "Some Theorems on Quadratic Forms Applied in the Study of Analysis of Variance Problems, I: Effect of Inequality of Variance in the one-way Classification", *Annals of Mathematical Statistics*, vol. 25, pp. 290-302.
- Box, G. E. P. (1954b), "Some Theorems on Quadratic Forms Applied in the Study of Analysis of Variance Problems, II: Effect of Inequality of Variance in the two-way Classification", *Annals of Mathematical Statistics*, vol. 25, pp. 484-498.
- Dhrymes, P. J. (1984), *Mathematics for Econometrics*, second edition, New York: Springer Verlag.
- Dhrymes, P. J. (1995), *Topics in Advanced Econometrics: vol. III, Topics in Time Series*, unpublished.
- Eaton, Morris L. (1983), *Multivariate Statistics: A Vector Space Approach*, New York: Wiley.
- Johansen, S. (1988), "Statistical Analysis of Cointegration Vectors", *Journal of Economic Dynamics and Control*, vol. 12, pp. 231-254.
- Johansen, S. (1991), "Estimation and Hypothesis Testing of Cointegration Vectors in Gaussian Vector Autoregressive Models", *Econometrica*, vol. 59, pp. 1551-1580.
- Phillips, P.C.B. and S. Ouliaris (1990), "Asymptotic Properties of Residual Based Tests for Cointegration", *Econometrica*, vol. 58, pp. 165-93.
- Siddiqui, M. M. (1965), "Approximations to the Distribution of Quadratic Forms", *Annals of Mathematical Statistics*, vol. 36, pp. 677-682.

1994-1995 Discussion Paper Series

Department of Economics
Columbia University
1022 International Affairs Bldg.
420 West 118th Street
New York, N.Y., 10027

The following papers are published in the 1994-95 Columbia University Discussion Paper series which runs from early November to October 31 (Academic Year). Domestic orders for discussion papers are available for purchase at \$8.00 (US) each and \$140.00 (US) for the series. Foreign orders cost \$10.00 (US) for individual paper and \$185.00 for the series. To order discussion papers, please send your check or money order payable to Department of Economics, Columbia University to the above address. Be sure to include the series number for the paper when you place an order.

- 708. Trade and Wages: Choosing among Alternative Explanations
Jagdish Bhagwati

- 709. Dynamics of Canadian Welfare Participation
Garrey F. Barret, Michael I. Cragg

- 710. Much Ado About Nothing? Capital Market Reaction to Changes in
Antitrust Precedent concerning Exclusive Territories.
Sherry A. Glied, Randall S. Kroszner

- 711. The Cost of Diabetes
Matthew Kahn

- 712. Evidence on Unobserved Polluter Abatement Effort
Matthew E. Kahn

- 713. The Premium for Skills: Evidence from Mexico
Michael Cragg

- 714. Measuring the Incentive to be Homeless
Michael Cragg, Mario Epelaum

- 715. The WTO: What Next?
Jagdish Bhagwati

- 716. Do Converters Facilitate the Transition to a New Incompatible Technology?
A Dynamic Analysis of Converters
Jay Phil Choi

- 716A. Shock Therapy and After: Prospects for Russian Reform
Padma Desai

- 717. Wealth Effects, Distribution and The Theory of Organization
-Andrew F. Newman and Patrick Legros

1994-95 Discussion Paper Series

718. Trade and the Environment: Does Environmental Diversity Detract from the Case for Free Trade?
-Jagdish Bhagwati and T.N. Srinivasan (Yale Univ)
719. US Trade Policy: Successes and Failures
-Jagdish Bhagwati
720. Distribution of the Disinflation of Prices in 1990-91 Compared with Previous Business Cycles
-Philip Cagan
721. Consequences of Discretion in the Formation of Commodities Policy
-John McLaren
722. The Provision of (Two-Way) Converters in the Transition Process to a New Incompatible Technology
-Jay Pil Choi
723. Globalization, Sovereignty and Democracy
-Jagdish Bhagwati
724. Preemptive R&D, Rent Dissipation and the "Leverage Theory"
-Jay Pil Choi
725. The WTO's Agenda: Environment and Labour Standards, Competition Policy and the Question of Regionalism
-Jagdish Bhagwati
726. US Trade Policy: The Infatuation with FTAs
-Jagdish Bhagwati
727. Democracy and Development: New Thinking on an Old Question
-Jagdish Bhagwati
728. The AIDS Epidemic and Economic Policy Analysis
-David E. Bloom, Ajay S. Mahal
729. Economics of the Generation and Management of Municipal Solid Waste
-David E. Bloom, David N. Beede
730. Does the AIDS Epidemic Really Threaten Economic Growth?
-David E. Bloom, Ajay S. Mahal
731. Big-City Governments
-Brendan O'Flaherty
732. International Public Opinion on the Environment
-David Bloom

1994-95 Discussion Paper Series

- 733. Is an Integrated Regional Labor Market Emerging in the East and Southeast Asia?
-David Bloom, Waseem Noor
- 734. Migration, Integration and Development
-Abhijit V. Banerjee, Andrew Newman
- 735. Infrastructure, Human Capital and International Trade
-Ronald Findlay
- 736. Short Ballots: Why Mayors Are in Charge of So Many Different Things
-Brendan O'Flaherty
- 737. Demand for Environmental Goods: Evidence from Voting Patterns on California Initiatives
-Matthew Kahn and John Matsusaka
- 738. Inflation and Stabilization in Poland 1990 - 1995
-S. Wellisz
- 739. Particulate Pollution Trends in the 1980's
-M. Kahn
- 740. Why has Wage Dispersion Grown in Mexico? Is it the Incidence of Reforms or the Growing Demand for Skills?
-M.I. Cragg and M. Epelbaum
- 741. Russia's Transition Toward the World Economy:
-P. Desai
- 742. Poland: Transition and Integration in the World Economy
-S. Wellisz
- 743. Team Effects on Compensation: An Application to Salary Determination in the National Hockey League
-T. Idson and L.H. Kahane
- 744. Economic Development, Growth of Human Capital, and the Dynamics of the Wage Structure
-J. Mincer
- 745. Technology Transfer with Moral Hazard
-J.P. Choi (Oct.)
- 746. Risk Preference Smoothness and the APT
-L.F. Herk, Columbia University, Institute on East Central Europe
- 747. A Characterization of Cointegration
P. Dhrymes (Oct.)
- 748. The Limiting Distribution of the Cointegration
P. Dhrymes (Oct.)