

DEVELOPMENTS IN THE STUDY OF COINTEGRATED ECONOMIC VARIABLES

*C. W. J. Granger**

I. INTRODUCTION

At the least sophisticated level of economic theory lies the belief that certain pairs of economic variables should not diverge from each other by too great an extent, at least in the long-run. Thus, such variables may drift apart in the short-run or according to seasonal factors, but if they continue to be too far apart in the long-run, then economic forces, such as a market mechanism or government intervention, will begin to bring them together again. Examples of such variables are interest rates on assets of different maturities, prices of a commodity in different parts of the country, income and expenditure by local government and the value of sales and production costs of an industry. Other possible examples would be prices and wages, imports and exports, market prices of substitute commodities, money supply and prices and spot and future prices of a commodity. In some cases an economic theory involving equilibrium concepts might suggest close relations in the long-run, possibly with the addition of yet further variables. However, in each case the correctness of the beliefs about long-term relatedness is an empirical question. The idea underlying cointegration allows specification of models that capture part of such beliefs, at least for a particular type of variable that is frequently found to occur in macroeconomics. Since a concept such as the long-run is a dynamic one, the natural area for these ideas is that of time-series theory and analysis. It is thus necessary to start by introducing some relevant time series models.

Consider a single series x_t , measured at equal intervals of time. Time series theory starts by considering the generating mechanism for the series. This mechanism should be able to generate all of the statistical properties of the series, or at very least the conditional mean, variance and temporal autocorrelations, that is the 'linear properties' of the series, conditional on past data. Some series appear to be 'stationary', which essentially implies that the linear properties exist and are time-invariant. Here we are concerned with the weaker but more technical

* I would like to acknowledge the excellent hospitality that I enjoyed at Nuffield College and the Institute of Economics and Statistics, Oxford whilst this paper was prepared.

requirement that the series has a spectrum which is finite but non-zero at all frequencies. Such a series will be called $I(0)$, denoting 'integrated of order zero'. Some series need to be differenced to achieve these properties and these will be called integrated of order one, denoted $x_t \sim I(1)$. More generally, if a series needs differencing d times to become $I(0)$, it is called integrated of order d , denoted $x_t \sim I(d)$. Let Δ^b denote application of the difference operator b times, if $x_t \sim I(d)$ then the b th difference series $\Delta^b x_t$ is $I(d-b)$. Sometimes a series needs to be integrated (summed) to become $I(0)$, for example the difference of an $I(0)$ series is $I(-1)$ and its integral is again $I(0)$. Most of this paper will concentrate on the practically important cases when $d=0$ or 1. The simplest example of an $I(0)$ series is a white noise ϵ_t , so that $\rho_k = \text{corr}(\epsilon_t, \epsilon_{t-k}) = 0$ for all $k \neq 0$. Another example is a stationary $AR(1)$ series, x_t , generated by

$$x_t = \alpha x_{t-1} + \epsilon_t \quad (1.1)$$

where $|\alpha| < 1$ and ϵ_t is white noise with zero mean. The simplest example of an $I(1)$ series is a random walk, where x_t is generated by

$$x_t = x_{t-1} + \epsilon_t \quad (1.2)$$

as would theoretically occur for a speculative price generated by an informationally efficient market. Here, the first differenced series is white noise. The most general $I(1)$ series replaces ϵ_t in equation (1.2) by any $I(0)$ series not necessarily having zero mean. Many macro economic series appear to be $I(1)$, as suggested by the 'typical spectral shape' (see Granger (1966)), by analysis of Box-Jenkins (1970) modelling techniques or by direct testing, as in Nelson and Plosser (1982). Throughout the paper all error processes, such as those in (1.1), (1.2) are assumed to have finite first and second moments.

There are many substantial differences between $I(0)$ and $I(1)$ series. An $I(0)$ series has a mean and there is a tendency for the series to return to the mean, so that it tends to fluctuate around the mean, crossing that value frequently and with rare extensive excursions. Autocorrelations decline rapidly as lag increases and the process gives low weights to events in the medium to distant past, and thus effectively has a finite memory. An $I(1)$ process without drift will be relatively smooth, will wander widely and will only rarely return to an earlier value. In fact, for a random walk, for a fixed arbitrary value the expected time until the process again passes through this value is infinite. This does not mean that returns do not occur, but that the distribution of the time to return is very long-tailed. Autocorrelations $\{\rho_k\}$ are all near one in magnitude even for large k ; an innovation to the process affects all later values and so the process has indefinitely long memory. To see this, note that the pure random walk $I(1)$ solves to give

$$x_t = \epsilon_t + \epsilon_{t-1} + \epsilon_{t-2} + \dots + \epsilon_1 \quad (1.3)$$

assuming the process starts at time $t = 0$, with $x_0 = 0$. Note that the variance of x_t is $t\sigma_\epsilon^2$ and becomes indefinitely large as t increases and $\rho_k = 1 - |k|/t$.

If x_t is a random walk with 'drift' (1.2) becomes

$$x_t = x_{t-1} + m + \epsilon_t$$

where ϵ_t is zero-mean white noise. The solution is now

$$x_t = mt + \sum_{j=0}^{t-1} \epsilon_{t-j} \quad (1.4)$$

so that x_t consists of a linear trend plus a drift-free $I(1)$ process (random walk) being the process in (1.3). The only more general univariate process considered in this section is

$$x_t = m(t) + x'_t$$

where x'_t is a drift-free random walk, such as generated by (1.3), and $m(t)$ is some deterministic function of time, being the 'trend in mean' of x_t .

II. COINTEGRATION

Consider initially a pair of series x_t, y_t , each of which is $I(1)$ and having no drift or trend in mean. It is generally true that any linear combination of these series is also $I(1)$. However, if there exists a constant A , such that

$$z_t = x_t - Ay_t \quad (2.1)$$

is $I(0)$, then x_t, y_t will be said to be cointegrated, with A called the cointegrating parameter. If it exists, A will be unique in the situation now being considered. As z_t has such different temporal properties from those of either of its components it follows that the x_t and y_t must have a very special relationship. Both x_t and y_t have dominating low-frequency or 'long wave' components, and yet z_t does not. Thus, x_t and Ay_t must have low-frequency components which virtually cancel out to produce z_t . A good analogy is two series each of which contain a prominent seasonal component. Generally, any linear combination of these series will also contain a seasonal, but if the seasonals are identical in shape there could exist a linear combination which has no seasonal.

The relationship

$$x_t = Ay_t \quad (2.2)$$

might be considered a long-run or 'equilibrium' relationship, perhaps as suggested by some economic theory, and z_t given by (2.1) thus measures the extent to which the system x_t, y_t is out of equilibrium,

and can thus be called the 'equilibrium error'. The term 'equilibrium' is used in many ways by economists. Here the term is not used to imply anything about the behaviour of economic agents but rather describes the tendency of an economic system to move towards a particular region of the possible outcome space. If x_t and y_t are $I(1)$ but 'move together in the long-run', it is necessary that z_t be $I(0)$ as otherwise the two series will drift apart without bound. Thus, for a pair of $I(1)$ series, cointegration is a necessary condition for the ideas discussed in the first section of this paper to hold. In some circumstances, an even stronger condition may be required, such as putting complete bounds on z_t , which will guarantee that it is $I(0)$, but such cases are not considered here.

The extension to series having trends in their means is straightforward. Consider

$$\begin{aligned}x_t &= m_x(t) + x'_t \\y_t &= m_y(t) + y'_t\end{aligned}\tag{2.3}$$

where x'_t, y'_t are both $I(1)$ but without trends in mean, and let

$$\begin{aligned}z_t &= x_t - Ay_t \\&= m_x(t) - Am_y(t) + x'_t - Ay'_t.\end{aligned}$$

For z_t to be $I(0)$, and x_t, y_t not to drift too far apart, it is necessary both that z_t have no trend in mean, so that

$$m_x(t) = Am_y(t)\tag{2.4}$$

for all t , and that x'_t, y'_t be cointegrated with the *same* value of A as the cointegrating parameter. It is seen that if the two trends in mean are different functions of time, such as an exponential and a cubic, then (2.4) cannot hold.

One thing that should be noted is that a model of the form

$$x_t = \beta y_t + e_t$$

where x_t is $I(0)$ and y_t is $I(1)$, makes no sense as the independent and dependent variables have such vastly different temporal properties. Theoretically the only plausible value for β in this regression is $\beta = 0$.

If x_t, y_t are both $I(1)$ without trends in mean and are cointegrated it has been proved in Granger (1983) and Granger and Engle (1985) that there always exists a generating mechanism having what is called the 'error-correcting' form:

$$\begin{aligned}\Delta x_t &= -\rho_1 z_{t-1} + \text{lagged}(\Delta x_t, \Delta y_t) + d(B)\epsilon_{1t} \\ \Delta y_t &= -\rho_2 z_{t-1} + \text{lagged}(\Delta x_t, \Delta y_t) + d(B)\epsilon_{2t}\end{aligned}\tag{2.5}$$

where

$$z_t = x_t - Ay_t,$$

$d(B)$ is a finite polynomial in the lag operator B (so that $B^k x_t = x_{t-k}$) and is the same in each equation, and ϵ_{1t} , ϵ_{2t} are joint white noise, possibly contemporaneously correlated and with $|\rho_1| + |\rho_2| \neq 0$. Not only must cointegrated variables obey such a model but the reverse is also true; data generated by an error-correction model such as (2.5) must be cointegrated. The reason for this is easily seen, as if x_t , y_t are $I(1)$ their changes will be $I(0)$ and so every term in the equations (2.5) is $I(0)$ provided z_t is also $I(0)$ meaning that x_t , y_t are cointegrated. If z_t is not $I(0)$, i.e., if x_t , y_t are not cointegrated, then the z_t term does not belong in these equations given that the dependent variables are $I(0)$ and hence at least one of ρ_1 , ρ_2 does not vanish.

These models were introduced into economics by Sargan (1964) and Phillips (1957) and have generated a lot of interest following the work of Davidson, Hendry, Srba and Yeo (1978), Hendry and von Ungern Sternberg (1980), Curry (1981), Dawson (1981) and Salmon (1982) amongst others. The models are seen to incorporate equilibrium relationships, perhaps suggested by an economic theory of the long-run, with the type of dynamic model favoured by time-series econometricians. The equilibrium relationships are allowed to enter the model but are not forced to do so. The title 'error-correcting' for equations such as (2.5) is a little optimistic. The absolute value of z_t is the distance that the system is away from equilibrium. Equation (2.5) indicates that the amount and direction of change in x_t and y_t take into account the size and sign of the previous equilibrium error, z_{t-1} . The series z_t does not, of course, certainly reduce in size from one time period to another but is a stationary series and thus is inclined to move towards its mean. A constant should be included in the equilibrium equation (2.2) and in (2.1) if needed, to make the mean of z_t zero.

There are a number of theoretical implications of cointegratedness that are easily derived from the results so far presented:

- (i) If x_t , y_t are cointegrated, so will be x_t and $by_{t-k} + w_t$, for any k where $w_t \sim I(0)$, with a possible change in cointegrating parameter. Formally, if x_t is $I(1)$ then x_t and x_{t-k} will be cointegrated for any k , but this is not an interesting property as it is true for any $I(1)$ process and so does not suggest a special relationship, unlike cointegration of a pair of $I(1)$ series. It follows that if x_t , y_t are cointegrated but are only observed with measurement error, then the two observed series will also be cointegrated if all measurement errors are $I(0)$.
- (ii) If x_t is $I(1)$ and $f_{n,h}(J_n)$ is the optimal forecast of x_{n+h} , based on the information set J_n available at time n , then x_{t+h} , $f_{t,h}(J_t)$ are cointegrated if J_n is a proper information set, that is if it includes x_{n-j} , $j \geq 0$. If J_n is not a proper information set, x_{t+h} and its optimum forecast are only cointegrated if x_t is cointegrated with variables in J_t .

- (iii) If x_{n+h} , y_{n+h} are cointegrated series with parameter A and are optimally forecast using the information set $J_n: x_{n-j}, y_{n-j}$, $j \geq 0$, then the h -step forecasts $f_{n,h}^x, f_{n,h}^y$ will obey

$$f_{n,h}^x = A f_{n,h}^y$$

as $h \rightarrow \infty$ (proved by S. Yoo (1986)). Thus, long-term optimum forecasts of x_t , y_t will be tied-together by the equilibrium relationships. Forecasts formed without cointegration terms such as univariate forecasts will not necessarily have this property.

- (iv) If T_t is an $I(1)$ target variable and x_t is an $I(1)$ controllable variable, then T_t, x_t will be cointegrated if optimum control is applied. (See Nickell (1985).)
- (v) If x_t, y_t are $I(1)$ and cointegrated, there must be Granger causality in at least one direction, as one variable can help forecast the other. This follows directly from the error-correlation model and the condition that $|\rho_1| + |\rho_2| \neq 0$, as z_{t-1} must occur in at least one equation and thus knowledge of z_t must improve forecastability of at least one of x_t, y_t . Here causality is with respect to the information set J_t defined in (iii).
- (vi) If x_t, y_t are a pair of prices from a jointly efficient, speculative market, they cannot be cointegrated. This follows directly from (v) as if the two prices were cointegrated, one can be used to help forecast the other and this would contradict the efficient market assumption. Thus, for example, gold and silver prices, if generated by an efficient market, cannot move closely together in the long-run. Tests of this idea have been conducted by Granger and Escribano (1986).

III. TESTING FOR COINTEGRATION

This topic has been discussed at some length by Granger and Engle (1985) and so only an outline of their conclusions is presented here. It is necessary to start with a test for whether a series x_t is $I(0)$ and a useful test has been provided by Dickey and Fuller (1981). The following regression is formed

$$\Delta x_t = \beta x_{t-1} + \sum_{j=1}^p \gamma_j \Delta x_{t-j} + e_t$$

where p is selected to be large enough to ensure that the residual e_t is empirical white noise. The test statistic is the ratio of $\hat{\beta}$ to its calculated standard error obtained from an ordinary least squares (OLS) regression. The null hypothesis is $H_0: x_t \sim I(1)$. This is rejected if $\hat{\beta}$ is negative

and significantly different from zero. However, the test-statistic does *not* have a t -distribution but tables of significance levels have been provided by Dickey and Fuller (1979).

To test for cointegration between a pair of series, that are expected to be $I(1)$, one method is to first form the 'cointegration regression'

$$x_t = c + \alpha y_t + a_t \quad (3.1)$$

and then to test if the residual a_t appears to be $I(0)$ or not. It might be noted that when x_t and y_t are cointegrated, this regression when estimated using, say, OLS should give an excellent estimate of the true cointegrating coefficient A , in large samples. Note that a_t will have a finite (or small) variance only if $\alpha = A$, otherwise a_t will be $I(1)$ and thus have theoretically a very large variance in a large sample. Stock (1984) has shown that when series are cointegrated, OLS estimates of A are highly efficient with variances $O(T^{-2})$ compared to more usual situations where the variances are $O(T^{-1})$, T being the sample size. Stock also shows that the estimates are consistent with an $O(T^{-1})$ bias. However, some recent Monte Carlo simulations by Banerjee *et al.* (1986) suggest that these bias terms can be very substantial in some cases. Two simple tests of the null hypothesis

$$H_0: x_t, y_t \text{ not cointegrated}$$

are based either on a Durbin-Watson statistic (D/W) for (3.1), but testing if D/W is significantly greater than zero, (see Sargan and Bhargava (1983) who provide critical values) or using the previously mentioned Dickey-Fuller test for \hat{a}_t . The latter test was found by Granger and Engle (1985) to have more stable critical values from a small simulation study and with $T = 100$ observations approximate significance levels for the pseudo t -statistic testing $\beta = 0$ are, 10 per cent ~ 2.88 , 5 per cent ~ 3.17 , 1 per cent ~ 3.75 . A great deal more experience with these tests, and more extensive simulation studies, are required before confidence in the quality of these, or alternative, testing procedures is assured. Some estimates of power for this test were found to be quite satisfactory for a sample size of 100.

Applying this test, some examples of the outcomes of empirical analysis are (mostly from Granger and Engle, 1985)

apparently cointegrated

US national income and consumption

US non-durables, production and sales

US short and long-term interest rates

UK W, P, H, U, T, (Hall, - this issue)

UK Velocity and short-term interest rates (Hendry and Ericsson, 1983)

apparently not cointegrated
 US wages and prices
 US durables, production and sales
 US money and prices.

Of course, some of the examples where cointegration was not found strongly suggest that further variables should be included in the investigation, such as the addition of productivity to wages and prices. This extension is considered next.

IV. GENERALISATION: MANY VARIABLES AND GENERAL COINTEGRATION

Let x_t be a vector of N component time series, each without trend in mean and each $I(d)$, $d > 0$. For the moment, it is assumed that the d -differenced vector series is a zero mean, purely non-deterministic stationary process, so that there is a Wold representation

$$(1 - B)^d x_t = C(B)\epsilon_t \quad (4.1)$$

where this is taken to mean that both sides have the same spectral matrix and ϵ_t is an $N \times 1$ zero-mean white noise vector with

$$\begin{aligned} E[\epsilon_t \epsilon_s'] &= 0 & t \neq s \\ &= G & t = s \end{aligned}$$

so that only contemporaneous correlations can occur. Equation (4.1) is normalized by taking $C(0) = I_N$, the unit matrix.

Then x_t will be said to be cointegrated $CI(d, b)$ if there exists a vector α such that

$$z_t = \alpha' x_t$$

is $I(d - b)$, $b > 0$.

The case considered in earlier sections has $N = 2$, $d = b = 1$. Moving to general values for N , d , b adds a large number of possible inter-relationships and models. In particular it is clear that α need no longer be unique, as there can be several 'equilibrium' relationships linking $N > 2$ variables. If there are r vectors α , each of which produces z 's integrated of order less than d , then r is called the 'order of cointegration' and it is easily seen that $r \leq N - 1$.

For the practically important case $d = b = 1$, it is shown in Granger (1983) and in Granger and Engle (1985) that

- (i) $C(1)$ is of rank $N - r$
- (ii) there exists a vector autoregressive (VAR) representation

$$A(B)x_t = d(B)\epsilon_t$$

where $A(1)$ is of rank r with $A(0) = I_N$ and $d(B)$ is a scalar stable lag polynomial. If a finite order VAR model exists, it takes this form but with $d(B) = 1$.

(iii) there exist $N \times r$ matrices α, γ of rank r such that

$$\begin{aligned}\alpha' C(1) &= 0 \\ C(1)\gamma &= 0 \\ A(1) &= \gamma\alpha'\end{aligned}$$

(iv) there exists an error-correction representation with

$$z_t = \alpha' x_t$$

an $r \times 1$ stationary vector, of the form

$$A^*(B)(1 - B)x_t = -\gamma z_{t-1} + d(B)\epsilon_t \quad (4.2)$$

where $A^*(0) = I_N$, $A^*(1)$ is of full rank and $|A^*(w)| = 0$ has all its roots outside the unit circle. It should be noted that the first term on the right hand side can be written as (given (iii) and (v))

$$\gamma z_{t-1} = A(1)x_{t-1}$$

and so, for all terms in (4.2) to be $I(0)$ it is necessary that $A(1)$ does not have a row consisting of just one non-zero term. A resulting condition on α is mentioned below.

Commenting on these results, (i) concerning the rank of $C(1)$ is a necessary and sufficient condition for cointegration and all other results are derived from it. In (ii) concerning VAR, $A(B)$ is the adjoint matrix of $C(B)$ and $d(B)$ is proportional to the determinant of $C(B)$ after dividing out unit roots. It follows from (ii) that if a VAR model is estimated for cointegrated variables, efficiency will be lost unless $A(1)$ is restricted to being of rank r .

In (iii) it should be noted that the matrices γ, α , are not uniquely defined by the set of equations shown. If θ is an $r \times r$ matrix of full rank, then γ can be replaced by $\gamma\theta$ and α' by $\theta^{-1}\alpha'$ and the equations will still hold. This lack of uniqueness leads to some interpretational problems in the error-correction model (4.2), which are similar to the identification problems of classical simultaneous equations models. To illustrate the problem, suppose that $N = 3$ and $r = 2$ and that α_1, α_2 are a pair of cointegrating vectors, giving

$$\begin{aligned}z_t(\alpha_1) &= \alpha_{11}x_{1t} + \alpha_{12}x_{2t} + \alpha_{13}x_{3t} \\ z_t(\alpha_2) &= \alpha_{21}x_{1t} + \alpha_{22}x_{2t} + \alpha_{23}x_{3t}\end{aligned}$$

as a pair of $I(0)$ variables corresponding to equilibrium relationships $\alpha_1' x_t = 0, \alpha_2' x_t = 0$. However, generally any combination of a pair of

$I(0)$ variables will also be $I(0)$ and so

$$z_t(\lambda) = (1 - \lambda)z_t(\alpha_1) + \lambda z_t(\alpha_2)$$

will also be $I(0)$ [it is assumed that for no λ will $z_t(\lambda)$ consist of just one component of x_t : this is a constraint on the matrix α preventing $z_t(\lambda) = x_t$, for example, which would make $z_t \sim I(1)$]. Thus, the equilibrium relations are not uniquely identified, and the error-correction models cannot be strictly interpreted as 'correcting' for deviations from a particular pair of equilibrium relationships. The only invariant relationship is the line in the (x_1, x_2, x_3) space defined by

$$z_t(\alpha_1) = 0, z_t(\alpha_2) = 0$$

This same line is given by

$$z_t(\lambda_1) = 0, z_t(\lambda_2) = 0$$

for any $\lambda_1 \neq \lambda_2$ and will be called the 'equilibrium sub-space'. The error-correction model might thus be interpreted as Δx_t being influenced by the distance the system is from the equilibrium sub-space. For general N, r , the equilibrium sub-space will be a hyper-plane of dimension $N - r$.

It is unclear if the identification question can be solved in ways similar to those used with simultaneous equations, that is by adding sufficient zeros to $A(1)$ or by appeals to 'exogeneity'.

For the $N = 3, r = 2$ case, λ 's can be chosen to give

$$z_t = \alpha_1 x_{1t} + \alpha_2 x_{2t}$$

and

$$z_t = \alpha_3 x_{1t} + \alpha_4 x_{3t}$$

and these seem to provide a natural way for testing for cointegration. For more general N and r , the number of possible combinations becomes extensive and testing will be more difficult, particularly when r is an unknown, as will be usual in practice.

Turning briefly to the most general case, with any N, d, b and r , the error-correction model becomes

$$A^*(B)(1 - B)^d x_t = -\gamma \{ [1 - (1 - B)^b] (1 - B)^{d-b} z_{t-1} + d(B) \epsilon_t \} \quad (4.3)$$

where $d(B)$ is a scalar polynomial in B .

It should be noted that $\{1 - (1 - B)^b\}$, if expanded in powers of B , has no term in B^0 and so only lagged z_t occur on the right hand side. Again, every term in (4.3) is $I(0)$ when cointegration is present. It is possible to define fractional differencing, as in Granger and Joyeux (1980), and equation (4.3) still holds in this case, although its practical importance has yet to be established.

In the general case (with integer N, b, d, r) Yoo (1986) has considered alternative ways of defining the z_t 's possibly using lagged x_t components, for a given $C(B)$ matrix but with some added assumptions about its form. Johanssen (1985) has also found some mathematically exact and attractive results for the general case, which do not rely on the assumption that all components of x_t are integrated of the same order. He points out, for example, that if x_{1t} is $I(1)$ and x_{2t} is $I(0)$, then x_{1t} and $\bar{x}_{2t} = \sum_{j=0}^t x_{2,t-j}$ could be cointegrated, thus expanding the class of variables that might be tested.

The work of Yoo and Johanssen suggests a more general definition of cointegration. Let $\alpha(B)$ be an $N \times 1$ vector of functions of the lag operator B , such that each component, such as $\alpha_j(B)$ has the property that $\alpha_j(1) \neq 0$. Then if x_t is a vector of $I(d)$ series such that

$$z_t = \alpha'(B)x_t$$

is $I(d-b)$, x_t may be called cointegrated. If a cointegrating vector α occurs, as defined in earlier sections there will be many $\alpha(B)$ that also cointegrate, and so uniqueness is lost but extra flexibility is gained. Consideration of these possibilities does allow for a generalisation that is potentially very important in economics. Suppose that $N = 2$, so that x_t has just two components, and let α be a cointegrating vector, with $\alpha' = (1, A)$. In this case α will be unique, if it does not depend on B , so that $r = 1$. [Generally, one would expect $r < N$]. However, there may exist another cointegrating vector of quite a different form,

$$\alpha'(B) = \left(1 - \frac{A'}{\Delta}, \frac{AA'}{\Delta} \right)$$

$\alpha' = (1, -A')$ and $\Delta = 1 - B$. An example of this possibility is where $x_t = (x_t, y_t)$, x_t, y_t are cointegrated with vector α , giving equilibrium error:

$$z_t = x_t - Ay_t$$

and $x_t, Sz_t = \sum_{j=0}^t z_{t-j}$ are cointegrated, so that $x_t - A'Sz_t$ is $I(0)$. This would correspond to a cointegrating vector of the form

$$\alpha(B) = (1 - SA', SAA')$$

where $S = 1/\Delta$ and $\Delta = 1 - B$.

For example, x_t, y_t could be sales and production of some industry, z_t = change in inventory, Sz_t inventory and x_t, y_t could be cointegrated as well as x_t, Sz_t . Another example might be x_t = income, y_t = expenditure, z_t = savings, Sz_t = wealth. Such series might be called 'multi-cointegrated'.

Throughout this section, if the series involved have deterministic trends in mean, these need to be estimated and removed before the concepts discussed can be applied. One method of removing trends of general shape is discussed in Granger (1985).

V. FURTHER GENERALIZATIONS

The processes considered so far have been linear and with time-invariant parameters. Clearly general models, and possibly more realistic ones, are achieved by removing these restrictions.

As institutions, technology and society changes, so may any equilibrium relationships. In the bivariate case, the cointegrating parameter may be slowly changing with time, for instance. To proceed with analysis of this idea, it is necessary to define time-varying parameter (TVP) $I(0)$ and $I(1)$ processes. Using concepts introduced by Priestley (1981), it is possible to define a time-varying spectrum $f_t(w)$ for a process such as one generated by an ARMA model with TVP. For example, consider

$$x_t = \beta(t)x_{t-1} + \epsilon_t$$

where $\beta(t)$ is a deterministic function of time, obeying the restriction that $|\beta(t)| < 1$ all t . If $f_t(w)$ is bounded above and also is positive for all t, w , the process may be called TVP $I(0)$. If the change of x_t is TVP $I(0)$, then x_t can be called TVP $I(1)$.

For a vector process x_t that is TVP $I(d)$ and has no deterministic components Cramer (1961) has shown that there exists a generalised Wold representation

$$(1 - B)^d x_t = C_t(B)\epsilon_t \quad (5.1)$$

where

$$\begin{aligned} E[\epsilon_t] &= 0 \\ E[\epsilon_t \epsilon_s'] &= 0 \\ E[\epsilon_t \epsilon_t'] &= \Omega_t \\ C_t(0) &= I_N \end{aligned}$$

and if

$$C_t(B) = \sum C_{jt} B^j$$

it will be assumed that

$$\sum_j C_{jt} \Omega_t C_{jt}' < \infty$$

so that the variance of $(1 - B)^d x_t$ is finite.

Assume now that $C_t(1)$ has rank $N - 1$ for all t , so that the cointegration rank is 1, then there will exist $N \times 1$ vectors $\alpha(t)$, $\gamma(t)$ such that

$$\alpha'(t)C_t(1) = 0$$

$$C_t(1)\gamma(t) = 0.$$

The TVP equilibrium error process will then be

$$z_t = \alpha'(t)x_t. \quad (5.2)$$

The corresponding error-correction models will be as (4.2) but with $A^*(B)$, γ , $d(B)$ all functions of time. A testing procedure would involve estimating the equilibrium regression (5.2) using some TVP techniques, such as a Kalman filter procedure, probably assuming that the components of $\alpha(t)$ are stochastic but slowly changing.

It might be thought that allowing $\alpha(t)$ to change with time can always produce an $I(0)z_t$. For example, suppose that $N = 2$ and consider

$$z_t = x_t - A(t)y_t$$

Taking $A(t) = x_t/y_t$ clearly gives $z_t = 0$, which is an uninteresting $I(0)$ situation. However, it is also clear that taking, $\bar{A}(t) = x_t/y_t + \delta$ will produce a z_t that is $I(1)$ in general. Interpretation of any TVP cointegration test will have to consider this possible difficulty.

Turning to the possibility of non-linear cointegration, it might be noted that in the basic error-correlation model (2.5) or (4.2) z_{t-1} terms appear linearly so that changes in dependent variables are related to z_{t-1} , whatever its size. In the actual economy, a more realistic behaviour is to ignore small equilibrium errors but to react substantially to large ones, suggesting a non-linear relationship. An error-correction model that captures this idea is, in the bivariate case,

$$\begin{aligned} \Delta x_t &= f_1(z_{t-1}) + \text{lagged}(\Delta x_t, \Delta y_t) + \epsilon_{1t} \\ \Delta y_t &= f_2(z_{t-1}) + \text{lagged}(\Delta x_t, \Delta y_t) + \epsilon_{2t} \end{aligned} \quad (5.3)$$

where

$$z_t = x_t - Ay_t.$$

It is generally true that if z_t is $I(0)$ with constant variance, then $f(z_t)$ will also be $I(0)$. Similarly, if z_t is $I(1)$ then generally $f(z_t)$ is also $I(1)$, provided $f(z)$ has a linear component for large z , i.e. $f(z) \rightarrow \sum_{j=0}^{\infty} a_j z^j$ with $a_0 \neq 0$. A rigorous treatment of these results is provided by Escriano (1986). As generally z_t and $f(z_t)$ will be integrated of the same order, if a test suggests that a pair of series are cointegrated, then a non-linear error-correction model of form (5.3) is a possibility. Of course, most of the other results of previous sections do not hold as they are based on the linear Wold representation. Equation (5.3) can be estimated by one of the many currently available non-linear, non-para-

metric estimation techniques such as that employed in Engle, Granger, Rice and Weiss (1986).

Error correction models essentially consider process whose components drift widely but the joint process has a generalised preference towards a certain part of the process space. In the cases so far considered this preferred sub-space is a hyper-plane but more general preferred sub-spaces could be considered although with considerably increased mathematical difficulty.

VI. CONCLUSION

This paper has attempted to expand the discussion about differencing macro-economic series when model building by emphasizing the use of a further factor, the 'equilibrium error', that arises from the concept of cointegration. This factor allows the introduction of the impact of long-run or 'equilibrium' economic theories into the models used by the time-series analysts to explain the short-run dynamics of economic data. The resulting error-correction models should produce better short-run forecasts and will certainly produce long-run forecasts that hold together in economically meaningful ways.

If long-run economic theories are to have useful impact on econometric models they must be helpful in model specification and yet not distract from the short-run aspects of the model. Historically, many econometric models were based on equilibrium relationships suggested by a theory, such as

$$x_t = Ay_t + e_t \quad (6.1)$$

without any consideration of the levels of integratedness of the observed variables x_t , y_t or of the residual series e_t . If x_t is $I(0)$ but y_t is $I(1)$, for example, the value of A in the resulting regression is forced to be near zero. If e_t is $I(1)$, standard estimation techniques are not appropriate. A test for cointegration can thus be thought of as a pre-test to avoid 'spurious regression' situations. Even if x_t and y_t are cointegrated an equation such as (6.1) can only provide a start for the modelling process, as e_t may be explainable by lagged changes in x_t and y_t , eventually resulting in an error-correction model of the form (2.5). However, there must be two such equations, which again makes the equation (2.5) a natural form. Ignoring the process of properly modelling the e_t can lead to forecasts from (6.1) that can be beaten by simple time-series models, at least in the short-term.

Whilst the paper has not attempted to link error-correction models with optimizing economic theory, through control variables for example, there is doubtless much useful work to be done in this area.

Testing for cointegration in general situations is still in an early stage of development. Whether or not cointegration occurs is an empirical

question but the beliefs of economists do appear to support its existence and the usefulness of the concept appears to be rapidly gaining acceptance.

University of California, San Diego

Date of Receipt of Final Manuscript: April 1986

REFERENCES

- Box, G. E. P. and Jenkins, G. M. (1970). *Time Series Analysis Forecasting and Control*, San Francisco, Holden Day.
- Cramer, H. (1961). 'On Some Classes of Non-Stationary Processes', *Proceedings 4th Berkeley Symposium on Math, Stats and Probability*, pp. 157-78, University of California Press.
- Currie, D. (1981). 'Some Long-Run Features of Dynamic Time-Series Models', *The Economic Journal*, Vol. 363, pp. 704-15.
- Davidson, J. E. H., Hendry, D. F., Srba, F. and Yeo, S. (1978). 'Econometric Modelling of the Aggregate Time-Series Relationship Between Consumer's Expenditure and Income in the United Kingdom', *The Economic Journal*, Vol. 88, pp. 661-92.
- Dawson, A. (1981). 'Sargan's Wage Equation: A Theoretical and Empirical Reconstruction', *Applied Economics*, Vol. 13, pp. 351-63.
- Dickey, D. A. and Fuller, W. A. (1979). 'Distributions of the Estimators for Autoregressive Time Series with a Unit Root', *Journal of the American Statistical Association*, Vol. 74, pp. 427-31.
- Dickey, D. A. and Fuller, W. A. (1981). 'The Likelihood Ratio Statistics for Autoregressive Time Series with a Unit Root', *Econometrica*, Vol. 49, pp. 1057-72.
- Engle, R. F., Granger, C. W. J., Rice, J. and Weiss, A. (1986). 'Non-Parametric Estimation of the Relationship Between Weather and Electricity Demand', *Journal of the American Statistical Association*, (forthcoming).
- Escribano, A. (1986). Ph.D. thesis, Economics Department, University of California, San Diego.
- Granger, C. W. J. (1966). 'The Typical Spectral Shape of an Economic Variable', *Econometrica*, Vol. 34, pp. 150-61.
- Granger, C. W. J. (1983). 'Co-Integrated Variables and Error-Correcting Models', UCSD Discussion Paper, pp. 83-13a.
- Granger, C. W. J. and Engle, R. F. (1985). 'Dynamic Specification with Equilibrium Constraints: Cointegration and Error-Correction', (forthcoming, *Econometrica*).
- Granger, C. W. J. and Escribano, A. (1986). 'Limitation on the Long-Run Relationship Between Prices from an Efficient Market', UCSD Discussion Paper.
- Granger, C. W. J. and Joyeux, R. (1980). 'An Introduction to Long-Memory Time Series and Fractional Differencing', *Journal of Time Series Analysis*, Vol. 1, pp. 15-29.
- Hendry, D. F. and Ericsson, N. R. (1983). 'Assertion without Empirical Basis: An Econometric Appraisal of "Monetary Trends in ... the United Kingdom" by Milton Friedman and Anna Schwartz', *Bank of England Academic Panel Paper No. 22*

- Hendry, D. F. and von Ungern-Sternberg, T. (1981). 'Liquidity and Inflation Effects on Consumer's Expenditure', in Deaton, A. S. (ed.), *Essays in the Theory and Measurement of Consumer's Behaviour*, Cambridge University Press.
- Johanssen, S. (1985). 'The Mathematical Structure of Error-Correction Models', Discussion Paper, Maths Department, University of Copenhagen.
- Nelson, C. R. and Plosser, C. I. (1982). 'Trends and Random Walks in Macroeconomic Time Series', *Journal of Monetary Economics*, Vol. 10, pp. 139-62.
- Nickell, S. (1985). 'Error-Correction, Partial Adjustment and all that: an Expository Note', *BULLETIN*, Vol. 47, pp. 119-29.
- Phillips, A. W. (1957). 'Stabilization Policy and the Time Forms of Lagged Responses', *Economic Journal*, Vol. 67, pp. 265-77.
- Priestley, M. B. (1981). *Spectral Analysis of Time Series*, Academic Press, New York.
- Salmon, M. (1982). 'Error Correction Mechanisms', *The Economic Journal*, Vol. 92, pp. 615-29.
- Sargan, J. D. (1964). 'Wages and Prices in the United Kingdom: A Study in Economic Methodology', in Hart, P., Mills, G. and Whittaker, J. N. (eds.), *Econometric Analysis for National Economic Planning*, Butterworths, London.
- Sargan, J. D. and Bhargava, A. (1983). 'Testing Residuals from Least Squares Regression for being generated by the Gaussian Random Walk', *Econometrica*, Vol. 51, pp. 153-74.
- Stock, J. H. (1984). 'Asymptotic Properties of a Least Squares Estimator of Co-Integrating Vectors', manuscript Harvard University.
- Yoo, S. (1986). Ph.D. thesis, Economics Department, University of California, San Diego.

Copyright of Oxford Bulletin of Economics & Statistics is the property of Blackwell Publishing Limited. The copyright in an individual article may be maintained by the author in certain cases. Content may not be copied or emailed to multiple sites or posted to a listserv without the copyright holder's express written permission. However, users may print, download, or email articles for individual use.