The Economic Journal, **101** (*March* 1991), 239–251
Printed in Great Britain

COINTEGRATION IN RECURSIVE SYSTEMS*

James Davidson and Stephen Hall

The theory of cointegration, developed in several recent papers such as Engle and Granger (1987), Engle and Yoo (1987), Johansen (1988*a*, *b*), Hendry (1986) *inter alia*, has been almost exclusively concerned with closed (complete) dynamic models. The exposition in Johansen (1988*b*) makes this point particularly clear, when a VAR system is specified which must contain all the relevant variables for the dynamic process under discussion. This contrasts with the most usual modelling framework in which a set of variables are modelled conditionally on another set, often taken to be strongly exogenous in the sense of Engle *et al.* (1983). This conditioning of the data generation process is central to any practical exercise in structural form modelling, otherwise the size of the data set under consideration quickly become intractable. In this paper we consider how the concept of cointegration may be extended to conditional models.

Every dynamic econometric equation which contains a levels solutions for an integrated variable must by definition contain a cointegrating vector; yet it is well-known that the cointegration of a set of variables implies nothing about the direction of causality within the set. Such equations seem to assume the long-run status of reduced forms, and the link with economic theory can become obscured. We attempt to clarify the role of economic theory by introducing the notion of *target relationships* which we assume to embody the main features of the underlying economic behaviour. These target relationships interact with the dynamic system to produce cointegrating vectors which may be thought of as linear combinations of the target relations, although the targets may not all be individually attainable. In this way it is possible for economic theory to play an explicit role in the modelling of cointegrated variables.

A simple question which is often neglected is: *why* do we observe integrated time series? In the context of a linear dynamic system, variables are integrated either because they are driven by other integrated variables, or because the dynamic processes generating them contain autoregressive roots of unity; in other words, unit roots may be found in either the marginal or conditional sub-systems, or of course both. In the former case we are able to draw a parallel with the conventional stability analysis of dynamic systems, although additional restrictions are involved in the case of higher orders of integration. In the latter case, the existence of unit roots is shown to be identified with a rank deficiency in the target relationships. There are two types of rank condition, which we call *insufficient targets* and *inconsistent targets* respectively, which appear to have distinctive behavioural interpretations.

* James Davidson acknowledges the support of the ESRC Macroeconomic Forecasting Consortium.

Much of the methodology for our analysis is derived from Davidson (1990), which examines the case of complete VAR processes. We begin, in Section I, by outlining the structure of a conditional sub-system of equations which determine the variables of interest. In Sections II and III we discuss the nature of cointegration when the conditional sub-system is stable and unstable, respectively. Section IV discusses the case of wage-price inflation as a simple illustration these concepts, and Section V concludes.

I. THE ERROR CORRECTION SYSTEM

Let $\mathbf{y}_t$ $(n \times 1)$ be a vector of endogenous variables observed at time t, and $\mathbf{z}_t$ $(m \times 1)$ a vector of variables which are exogenous in the sense of not being Granger-caused by $\mathbf{y}_t$.[1] A standard framework for modelling the joint generation process of $\mathbf{y}_t$ given $\mathbf{z}_t$ is the autoregressive distributed lag (ADL) system,

$$\mathbf{B}(\mathrm{L})\,\mathbf{y}_t + \boldsymbol{\Gamma}(\mathrm{L})\,\mathbf{z}_t = \mathbf{u}_t, \qquad (1)$$

where $\mathbf{B}(z) = \sum_{i=0}^{p} \mathbf{B}_i z^i$ and $\boldsymbol{\Gamma}(\mathbf{z}) = \sum_{i=0}^{p} \boldsymbol{\Gamma}_i z^i$, $z \in C$, are matrix polynomials of finite maximum order p, of dimension $n \times n$ and $n \times m$ respectively, and $\mathbf{u}_t$ $(n \times 1)$ is a vector of unobserved disturbances. (See e.g. Wallis (1977) for a survey of the properties of these models.) Strictly speaking, we should refer to (1) as a conditional sub-system of the complete system determining both $\mathbf{y}_t$ and $\mathbf{z}_t$. We will assume $\mathbf{z}_t$ is generated by a closed VAR,

$$\boldsymbol{\Theta}(\mathrm{L})\,\mathbf{z}_t = \mathbf{v}_t, \qquad (2)$$

(the marginal subsystem) such that there is no dynamic feedback from $\mathbf{y}$ to $\mathbf{z}$. In fact, no part of the analysis will depend upon either the contemporaneous or serial independence of $\mathbf{u}_t$ and $\mathbf{v}_t$, although the statistical treatment would obviously be simpler in the independent case.

As is well known, the system can be solved for $\mathbf{y}_t$ in terms of current and lagged $\mathbf{z}_t$ and $\mathbf{u}_t$ if and only if the roots of $|\mathbf{B}(z)| = 0$ all lie outside the unit circle, in which case the inverse $\mathbf{B}(z)^{-1}$ is an r-summable polynomial for all finite r. The system is then said to be *stable*, and it has the so-called final form representation

$$\mathbf{y}_t = \boldsymbol{\Pi}(\mathrm{L})\,\mathbf{z}_t + \mathbf{B}(\mathrm{L})^{-1}\,\mathbf{u}_t, \qquad (3)$$

where $\boldsymbol{\Pi}(z) = -\mathbf{B}(z)^{-1}\boldsymbol{\Gamma}(z)$ is likewise summable to all finite orders. If $\mathbf{z}_t$ has been fixed at $\mathbf{z}$ and $\mathbf{u}_t$ at $\mathbf{0}$ for an infinite number of periods, then $\mathbf{y}_t = \mathbf{y} = \boldsymbol{\Pi}\mathbf{z}$, where the matrix $\boldsymbol{\Pi} = \boldsymbol{\Pi}(1)$ is the sum of the lag weights, or matrix of total multipliers.[2] If likewise $\mathbf{B} = \mathbf{B}(1)$ and $\boldsymbol{\Gamma} = \boldsymbol{\Gamma}(1)$, it is easy to show that $\boldsymbol{\Pi} = -\mathbf{B}^{-1}\boldsymbol{\Gamma}$.

[1] See Granger (1969). The essential property required is the block-triangularity of the VAR generating the variables, which will often coincide with the concept of strong exogeneity in the sense of Engle *et al.* (1983); however, it will be seen in Section II below that cases arise where this assumption may be violated.

[2] To simplify the numerous partitioned matrix expressions which will arise, $\mathbf{P}$ will always be taken to denote $\mathbf{P}(1)$, for any matrix polynomial $\mathbf{P}(z)$ which has been previously defined.

Now suppose that an economic theory of long-run equilibrium suggests a set of static relations amongst the variables,

$$\bar{\mathbf{B}}\mathbf{y}+\bar{\mathbf{\Gamma}}\mathbf{z} = \mathbf{0}. \tag{4}$$

These will be called the *target relations*, since they are assumed to be formulated in terms of the goals of economic agents. If we observe the variables of the system at a point in time t, as $\mathbf{y}_t$ and $\mathbf{z}_t$, then the vector $\mathbf{e}_t = \bar{\mathbf{B}}\mathbf{y}_t+\bar{\mathbf{\Gamma}}\mathbf{z}_t$ defines the 'target error' at time t. If the target relations are equal in number to the endogenous variables and $|\bar{\mathbf{B}}| \neq 0$, the static system has a unique solution

$$\mathbf{y} = -\bar{\mathbf{B}}^{-1}\mathbf{\Gamma}z. \tag{5}$$

Obviously, if (1) is our model of the dynamic adjustment process which relates the paths of $\mathbf{y}_t$ and $\mathbf{z}_t$, we require

$$\mathbf{\Pi} = -\bar{\mathbf{B}}^{-1}\bar{\mathbf{\Gamma}} = -\mathbf{B}^{-1}\mathbf{\Gamma}. \tag{6}$$

But the matrices $\mathbf{B}$ and $\bar{\mathbf{B}}$, and $\mathbf{\Gamma}$ and $\bar{\mathbf{\Gamma}}$ are evidently different. The sums of the estimated coefficients of the ADL will not yield estimates of the parameters of interest in the target relations, but linear combinations of them. We can write

$$\mathbf{B} = \mathbf{C}\bar{\mathbf{B}}, \quad \mathbf{\Gamma} = \mathbf{C}\bar{\mathbf{\Gamma}}, \tag{7}$$

where $\mathbf{C}$ $(n \times n)$ is a matrix which is nonsingular if both $\mathbf{B}$ and $\bar{\mathbf{B}}$ are. The conventional analysis of the ADL model does not define this matrix, and focuses instead on the matrix of total multipliers $\mathbf{\Pi}$ in the long-run analysis. However, the identification of the target relations is a problem wholly distinct from the conventional identification of the ADL (as treated in Hsiao (1983) for example), although it is a simple matter of verifying the usual rank and order conditions on $[\bar{\mathbf{B}}:\bar{\mathbf{\Gamma}}]$.[3]

The elements of $\mathbf{C}$ are known as the error correction coefficients of the system, terminology which is better understood if we adopt an error correction parameterisation for the system. Using the equivalence

$$\mathbf{P}(z) = \sum_{j=0}^{p} \mathbf{P}_j z^j = \sum_{k=0}^{p} \mathbf{P}^{\{k\}}(1-z)^k, \tag{8}$$

we have

$$\mathbf{B}(\mathrm{L})\,\mathbf{y}_t+\mathbf{\Gamma}(\mathrm{L})\,\mathbf{z}_t = \mathbf{C}(\bar{\mathbf{B}}\mathbf{y}_t+\bar{\mathbf{\Gamma}}\mathbf{z}_t)+\sum_{j=1}^{p}[\mathbf{B}^{\{j\}}\Delta^j\mathbf{y}_t+\mathbf{\Gamma}^{\{j\}}\Delta^j\mathbf{z}_t].^4 \tag{9}$$

[3] To appreciate the distinction, compare Hatanaka's conditions (Hatanaka, 1975) for the identification of the ADL system, based on zero restrictions on $[\mathbf{B}(\mathrm{L}):\mathbf{\Gamma}(\mathrm{L})]$. These conditions are neither necessary nor sufficient for identification of $[\bar{\mathbf{B}}:\bar{\mathbf{\Gamma}}]$. (See also Granger (1986), Section IV, and Bewley (1979).)

[4] For purposes of estimation or prediction, the EC model can be further rearranged to eliminate the multiple occurrence of current $\mathbf{y}_t$ and $\mathbf{z}_t$, as

$$\mathbf{B}_0\Delta^p\mathbf{y}_t+\mathbf{\Gamma}_0\Delta^p\mathbf{z}_t+\sum_{j=1}^{p-1}(\mathbf{B}_j^*\Delta^j\mathbf{y}_{t-1}+\mathbf{\Gamma}_j^*\Delta^j\mathbf{z}_{t-1})+\mathbf{C}(\bar{\mathbf{B}}\mathbf{y}_{t-1}+\bar{\mathbf{\Gamma}}\mathbf{z}_{t-1}) = \mathbf{u}_t,$$

where

$$\mathbf{B}_j^* = \sum_{k=0}^{j}\mathbf{B}^{\{k\}} \text{ and } \mathbf{\Gamma}_j^* = \sum_{k=0}^{j}\mathbf{\Gamma}^{\{k\}}.$$

If the system is unstable because $|\mathbf{B}(z)| = 0$ possesses one or more roots of unity, the conventional analysis of final forms breaks down, and it used to be said that in this case the sub-system had no solution. It is more proper to say that there may exist a solution of reduced rank, such that there exist one or more linear combinations of $\mathbf{e}_t$ which are I(0). It is immediate that in this case $|\mathbf{B}| = 0$, which implies that either $\mathbf{C}$ or $\bar{\mathbf{B}}$, or both, have rank less than n. When there are unit roots in the autoregressive components, ADL models generate integrated variables. The case of a stable conditional sub-process in which the exogenous variables contain trends corresponds in this framework to the case where $|\mathbf{\Theta}(z)| = 0$ has unit roots, whereas in the unstable case $|\mathbf{B}(z)| = 0$ will also contain unit roots.

II. COINTEGRATION IN THE STABLE CASE

Stability implies that if $\mathbf{z}_t$ is a stationary process with fixed mean $\boldsymbol{\mu}_z$, say, then $\mathbf{y}_t$ is also stationary with mean $\boldsymbol{\mu}_y = \mathbf{\Pi}\boldsymbol{\mu}_z$. On the other hand, suppose $\mathbf{z}_t$ is integrated to order d. A scalar stochastic sequence x_t is said (e.g. by Engle and Granger (1987)) to be integrated to order d (I(d)) if d is the smallest positive integer such that $\Delta^d x_t$ is representable as a stationary, invertible ARMA process (termed I(0)) with zero mean. A vector sequence is said to be I(d) if at least one of its elements is I(d).[5] We shall assume that the disturbance vectors $\mathbf{u}_t$ and $\mathbf{v}_t$ are I(0) processes.

THEOREM 1. *Stability implies* (i) $\mathbf{y}_t$ *is* I(d') *for* $d' \leqslant d$, (ii) $\mathbf{e}_t$ *is* I(d'') *for* $d'' \leqslant d-1$. (See Appendix for proof.)

If n_d of the elements of the n-vector $\mathbf{\Pi}\mathbf{z}_t$ are I(d), we can accordingly say that the vector $(\mathbf{y}_t, \mathbf{z}_t)$ is cointegrated with cointegrating rank n_d,[6] with cointegrating vectors lying in the space spanned by the rows of $[\bar{\mathbf{B}}:\bar{\mathbf{\Gamma}}]$.

So for the case $d = 1$ the relatively novel concept of cointegration meshes naturally with the established concept of stability. It implies that the target errors $\mathbf{e}_t$ are I(d'') for $d'' \leqslant 0$,[7] stochastic processes with zero mean and finite variance. However, when $d > 1$ stability does not guarantee this last condition. In the case $\mathbf{e}_t \sim$ I(d'') for $d'' \leqslant 0$ we will say that $\mathbf{z}_t$ and $\mathbf{y}_t$ are *fully cointegrated*, and except when $d = 1$ this condition calls for some restrictions on the parameters of the model.

It is convenient to write the model in block-triangular VAR form, as

$$\begin{bmatrix} \mathbf{B}(L) & \mathbf{\Gamma}(L) \\ \mathbf{0} & \mathbf{\Theta}(L) \end{bmatrix} \begin{bmatrix} \mathbf{y}_t \\ \mathbf{z}_t \end{bmatrix} = \begin{bmatrix} \mathbf{u}_t \\ \mathbf{v}_t \end{bmatrix} \begin{matrix} n \\ m. \end{matrix} \tag{10}$$

[5] For the purposes of general dynamic modelling the Engle-Granger definition is somewhat restrictive; a more general definition of an integrated process might permit some heterogeneity in the definition of an I(0) process, and also allow deterministic components to be present. However, none of the theoretical results of the present paper hinge on the precise definition of I(0). The Engle-Granger definition can therefore be adopted with no loss of generality.

[6] In Granger's terminology, the vector sequence $\mathbf{x}_t$ ($n \times 1$) is cointegrated, or CI(d, b), if there exists a fixed n-vector $\boldsymbol{\alpha}$ such that $\boldsymbol{\alpha}'\mathbf{x}_t \sim$ I($d-b$), for $b \geqslant 1$. If there exist n_d linearly independent cointegrating vectors, $\mathbf{x}_t$ is said to have cointegrating rank n_d.

[7] Except in very special circumstances full cointegration would imply $d'' = 0$; but to specify conditions which would rule out the case of 'over-differenced' $\mathbf{e}_t$ would be distracting, and of marginal importance.

Since it is assumed that $\mathbf{z}_t$ is $I(d)$, $d > 0$, it follows that $|\mathbf{\Theta}(z)| = 0$ possesses at least one unit root, $\mathbf{\Theta}$ having rank $t < m$, where $|\mathbf{B}(z)| = 0$ has no unit roots and $\mathbf{B}$ has full rank.

THEOREM 2. *Let* $\mathbf{B}$ *have rank* n *and* $\mathbf{\Theta}$ *have rank* $t < m$ *in* (10), *and define*

$$\mathbf{J}(z) = \mathbf{B}(z)\,\mathbf{\Pi} + \mathbf{\Gamma}(z). \tag{11}$$

The vector $(\mathbf{y}_t', \mathbf{z}_t')'$ *is fully cointegrated if no element of* $\mathbf{J}(z)\mathbf{\Theta}(z)^{-1}$ *has a pole at* 1, *and only if no row of* $\mathbf{J}(z)\mathbf{\Theta}(z)^{-1}$ *has all elements with a pole at* 1. (See Appendix for proof.)

Writing $\mathbf{J}(z) = \Sigma_{j=0}^{p} \mathbf{J}^{\{j\}}(1-z)^j$ by analogy with (8), observe that $\mathbf{J}^{\{0\}} = 0$ by definition of $\mathbf{\Pi}$, and hence it is possible to write $\mathbf{J}(z) = (1-z)\tilde{\mathbf{J}}(z)$ where $\tilde{\mathbf{J}}(z) = \Sigma_{j=0}^{p-1} \mathbf{J}^{\{j+1\}}(1-z)^j$, which reproduces the result of Theorem 1; at least one pole of unity in $\mathbf{\Theta}(z)^{-1}$ is always cancelled by $\mathbf{J}(z)$. More generally, the significant feature of this result is that cointegration of $\mathbf{y}_t$ and $\mathbf{z}_t$ is a restriction across the parameters of both the conditional and marginal sub-processes. The parameter restrictions involved can be explored further using Theorems 3.2–3.5 of Davidson (1990).

One question immediately posed is answered by the following corollary, whose proof is just an extension of the argument for the case $j = 0$ which followed Theorem 2:

COROLLARY 1. *Necessary and sufficient conditions for* $\mathbf{y}_t$ *to be fully cointegrated with all* $I(d)$ *marginal sub-processes* $\mathbf{z}_t$ *are that* $\mathbf{J}^{(j)} = \mathbf{0}$ *for* $0 \leqslant j \leqslant \min(d-1, p)$.

These conditions, called *trend neutrality* conditions of the conditional sub-model, are necessary for the outputs of the conditional model to track the inputs when the latter are subject to arbitrary trends of a certain order. Noting how $\mathbf{J} = \mathbf{0}$ by definition, they can be interpreted as restrictions by which the dynamic components of the model 'mimic' the long-run components; the relationships between $\Delta^j\mathbf{y}_t$ and $\Delta^j\mathbf{z}_t$ for $1 < j \leqslant d-1$, have to match those between $\mathbf{y}_t$ and $\mathbf{z}_t$, so that the responses are the same at each frequency.

Models with these characteristics were investigated in a number of papers from the early 80s, including Currie (1981), Patterson and Ryding (1984), Davidson (1984), Kloek (1985) and Pagan (1985). These authors generally considered the case of deterministic $\mathbf{z}_t$ (polynomial trends), in which case (3) shows that $\mathbf{y}_t$ will be the sum of a deterministic function of time and an $I(0)$ stochastic process. The expected path of the system can be obtained by setting $\mathbf{u}_t$ to zero, and is called the 'growth steady state' of the system. Letting the relation $\mathbf{\Pi}(z) = \mathbf{\Pi} + \mathbf{\Pi}^*(z)(1-z)$ define $\mathbf{\Pi}^*(z)$ (as in the proof of Theorem 1), this has the form

$$\mathbf{y}_t = \mathbf{\Pi}\mathbf{z}_t + \mathbf{\Pi}^*(L)\,\Delta\mathbf{z}_t. \tag{12}$$

The dependence of the steady state on terms in $\Delta\mathbf{z}_t$ which have not been intended or justified theoretically in the design of a conditional model is a disturbing feature which modellers have sought to eliminate by imposing parameter restrictions. But such restrictions limit very severely the variety of permissible dynamic structures, and are usually rejected empirically. Just how severe they are can be seen by considering the case $d-1 \geqslant p$, where Corollary 1 gives $\mathbf{J}(z) = \mathbf{0}$, or

$$\mathbf{\Gamma}^{\{j\}} = -\mathbf{B}^{\{j\}}\mathbf{\Pi}, \quad j = 0, \dots, p \tag{13}$$

as necessary and sufficient. Using (13) in the error correction form (9), we obtain

$$\begin{aligned}\mathbf{B}(\mathrm{L})\,\mathbf{y}_t + \mathbf{\Gamma}(\mathrm{L})\,\mathbf{z}_t &= \mathbf{B}\mathbf{y}_t + \mathbf{\Gamma}\mathbf{z}_t + \sum_{j=1}^{p} [\mathbf{B}^{\{j\}}\Delta^j\,\mathbf{y}_t + \mathbf{\Gamma}^{\{j\}}\Delta^j\,\mathbf{z}_t] \\ &= \mathbf{B}(\mathbf{y}_t - \mathbf{\Pi}\mathbf{z}_t) + \sum_{j=1}^{p} \mathbf{B}^{\{j\}}\Delta^j(\mathbf{y}_t - \mathbf{\Pi}\mathbf{z}_t) \\ &= \mathbf{B}(\mathrm{L})(\mathbf{y}_t - \mathbf{\Pi}\mathbf{z}_t) = \mathbf{u}_t \end{aligned} \tag{14}$$

which solves as

$$\bar{\mathbf{B}}\mathbf{y}_t + \bar{\mathbf{\Gamma}}\mathbf{z}_t = \bar{\mathbf{B}}\mathbf{B}(\mathrm{L})^{-1}\,\mathbf{u}_t. \tag{15}$$

We can recognise this as the common factor model, the system without systematic dynamics but with a p^{th} order VAR disturbance. Whether or not such restrictions are plausible, it is worth noting that if $p = 1$ (15) is the only dynamic specification which is compatible with cointegration of any order. In a distributed lag model, high order lags of $\mathbf{y}_t$ or $\mathbf{z}_t$ are required for cointegration to be compatible with highly integrated exogenous processes without the model collapsing to (15).

The behavioural interpretation of trend neutrality is that the model dynamics are due to targeting errors in a stochastic or uncertain environment. Adjustment to smooth exogenous changes ('smooth' implying a low-order polynomial in time) is actually instantaneous. By contrast, the partial adjustment (positive mean lag) hypothesis, which typically has an interpretation in terms of the costs of change, implies that a moving target (even if moving predictably) is never attained.

However, the need for trend neutrality may often be the consequence of an improper assumption of exogeneity. The modeller observes cointegration between the input and output series of his estimated conditional model, and is tempted to impose neutrality restrictions to ensure the same property will hold in respect or arbitrarily chosen input paths in simulation exercises. Cointegration *in the data* will be due to the necessary condition of Theorem 2 being satisfied, but only the conditions of Corollary 1 can impose it *on the model* without sacrificing the assumption that $[\mathbf{B}(\mathbf{z}):\mathbf{\Gamma}(\mathbf{z})]$ and $\mathbf{\Theta}(\mathbf{z})$ are variation-free in the sense of Engle *et al.* (1983). This critique of exogeneity assumptions is familiar from the rational expectations literature, but clearly can be much more general, extending to any form of *ex post* error correction behaviour in an evolving environment.

III. COINTEGRATION IN THE UNSTABLE CASE

The assumption that the number of linearly independent target relations is equal to the number of dynamic adjustment equations, and that the two are linked by a nonsingular matrix $\mathbf{C}$ of error correction coefficients is implicitly 'conventional';[8] but there is really no reason why either more or fewer than n targets should not be specified. In a general equilibrium framework the targets might be thought of as characterising the goals of particular groups of agents. One type of target is a notional demand or supply schedule. Another might merely specify a desired growth path, or steady-state value for a variable. Different groups of agents are obviously capable of targeting different paths for the same variable, just as demand and supply schedules do not always intersect at feasible points. There is no reason to suppose that only compatible targets leading to a unique solution to the conditional system will be pursued. On the other hand, a deficiency of target relations is also possible, with certain variables having indeterminate long-run values, or no long-run constraints on their behaviour relative to other variables. Any situation where a random shock is expected not to be transitory, but to be built in to the subsequent path of the variable, might be modelled in a similar way as 'indeterminate'; demographic variables for example.

To free the number of target relations, we can assume that $[\bar{\mathbf{B}}:\bar{\mathbf{\Gamma}}]$ has $n+m$ columns but any number of rows, and that $\mathbf{C}$ has n rows but a column dimension to match that of $[\bar{\mathbf{B}}:\bar{\mathbf{\Gamma}}]$. Of course, a maximum of $n(n+m)$ parameters can be solved from estimates of $[\mathbf{B}:\mathbf{\Gamma}] = \mathbf{C}[\bar{\mathbf{B}}:\bar{\mathbf{\Gamma}}]$ so there will sometimes be an identification problem in estimation, although this is not our present concern. The question of interest is the dynamic properties of the model resulting from a particular configuration of targets and adjustment processes. When $\text{rank}(\mathbf{C}) = \text{rank}(\bar{\mathbf{B}}) = n$, and $\text{rank}([\bar{\mathbf{B}}:\bar{\mathbf{\Gamma}}]) \geqslant n$, the conditional model will be dynamically stable in the sense of Section II. There will be n 'effective' linearly independent targets, captured in the rows of $[\mathbf{B}:\mathbf{\Gamma}]$. If there are more than n linearly independent 'underlying' targets they cannot all be attained simultaneously, but n linear combinations of them will be attainable, the weights being supplied by the rows of $\mathbf{C}$.

On the other hand, suppose $\text{rank}(\mathbf{B}) < n$, implying instability of the conditional sub-system. If $\text{rank}(\mathbf{C}) < n$, this will typically imply that certain dynamic equations are insensitive to the targets, and determine only the rates of change of the variables concerned without reference to the level, either absolutely or relative to other variables. And if $\text{rank}([\bar{\mathbf{B}}:\bar{\mathbf{\Gamma}}]) < n$, the implication is that there are fewer targets than variables determined by the system. A third possibility is that $\text{rank}(\mathbf{C}) = n$ and $\text{rank}([\bar{\mathbf{B}}:\bar{\mathbf{\Gamma}}]) \geqslant n$, but $\text{rank}(\bar{\mathbf{B}}) = s < n$. This implies that although there are at least n targets, no more than s of them (or linear combinations of them) can be attained within the conditional subsystem. In this case the *conditional* targets are inconsistent, although they could be attainable if agents were able to gain some influence over $\mathbf{z}_t$.

Whichever of the above scenarios is in effect, there are just three distinct classes of model to be distinguished when we come to consider the directly estimable matrix $[\mathbf{B}:\mathbf{\Gamma}] = \mathbf{C}[\bar{\mathbf{B}}:\bar{\mathbf{\Gamma}}]$. The first is the stable case already considered, in which $\mathbf{B}$ has full rank n; the cases we are now concerned with are

[8] That is to say: it is not the error correction framework which is conventional, but the form of EC model implicit in the stability assumption.

those where $\mathbf{B}$ has rank $s < n$, such that $|\mathbf{B}(z)| = 0$ possesses unit roots, and the conditional process is integrating.[9] Then we either find $\text{rank}([\mathbf{B}:\mathbf{\Gamma}]) = \text{rank}(\mathbf{B}) = s$, or $\text{rank}([\mathbf{B}:\mathbf{\Gamma}]) > s$. In the first case we will say that instability comes about through an *insufficiency* of targets. On the other hand, when $\text{rank}([\mathbf{B}:\mathbf{\Gamma}]) = n$, the instability is attributed to the *inconsistency* of the targets. More generally, of course, we may find the mixed case where $s < \text{rank}([\mathbf{B}:\mathbf{\Gamma}]) < n$.

THEOREM 3. (*i*) *The maximum cointegrating rank of the complete system in* (10) *is* $s+t$. (*ii*) *The maximum cointegrating rank for the vector* $(\mathbf{y}'_t, \mathbf{z}'_t)'$ *is* s. (See Appendix for proof.)

What part (ii) of the theorem means is that if the number of linearly independent target relations in the conditional subsystem (rows of $[\bar{\mathbf{B}}:\bar{\mathbf{\Gamma}}]$) exceeds s, not all the targets are attainable, in the sense that the target error is an I(0) process. At most a set of s linear combinations of the targets will be fully cointegrating vectors.[10]

This typology of models can be related to the categorisation of integrating processes into *balanced* and *unbalanced* cases, which is introduced by Johansen (1988*b*) and discussed in Davidson (1990). Briefly, an integrating closed VAR process $\mathbf{A}(\mathrm{L})\,\mathbf{x}_t = \boldsymbol{\varepsilon}_t$ $((n+m)\times 1)$, in which $|\mathbf{A}(z)| = 0$ contains unit roots, is said to be balanced if it can be written in the form

$$\tilde{\mathbf{A}}(\mathrm{L})\,\mathbf{x}^*_t = \boldsymbol{\varepsilon}_t, \tag{16}$$

where $|\tilde{\mathbf{A}}(z)| = 0$ has all roots outside the unit circle, and

$$\mathbf{x}^*_t = \begin{bmatrix} \mathbf{x}_{1t} + \mathbf{A}_{11}^{-1}\mathbf{A}_{12}\,\mathbf{x}_{2t} \\ \mathbf{\Delta}_{22}(\mathrm{L})\,\mathbf{x}_{2t} \end{bmatrix}. \tag{17}$$

Here the $\mathbf{A}_{ij}$ correspond to the square partition of $\mathbf{A}$ such that $\text{rank}(\mathbf{A}_{11}) = \text{rank}(\mathbf{A})$, and $\mathbf{\Delta}_{22}(\mathrm{L})$ is a diagonal matrix having powers of Δ on the diagonal. Such systems were shown by Johansen (see Johansen (1988*a*), Davidson (1990)) to be fully cointegrating with cointegrating rank equal to $\text{rank}(\mathbf{A})$. Evidently, this is a property that models with insufficient targets may possess, but is incompatible with inconsistent targets models, which by definition have more target relations (rows of $\mathbf{A}$) than cointegrating vectors.

The next question is therefore to consider whether an insufficient targets subsystem actually is fully cointegrating. This may be answered with reference to the theorems in Davidson (1990). The arguments parallel the stable case rather closely, and will not be derived in detail. A condition that will always be necessary is that the closed VAR system with dynamic structure represented by $\mathbf{B}(z)$ should be fully cointegrating. In addition, there are conditions which call

[9] If the marginal process is also integrating we will have 'two-tier' integration, so that $\mathbf{z}_t$ is I(1) and $\mathbf{y}_t$ is I(2), for example: a block-recursive structure of unstable sub-models is just one way in which higher orders of integration may arise.

[10] The one exception here is when the rows of $\mathbf{\Gamma}$ are linearly dependent on the rows of $\mathbf{\Theta}$, so that the rank of $\begin{bmatrix} \mathbf{B} & \mathbf{\Gamma} \\ \mathbf{0} & \mathbf{\Theta} \end{bmatrix}$ is $s+t$, in spite of $\text{rank}([\mathbf{B}:\mathbf{\Gamma}]) > s$. The targets can then be attained because they impose conditions which the marginal process happens to obey. Strictly speaking, target inconsistency is a condition on the whole system (ruling out this case), not just on the conditional subsystem.

generally for dependence between the parameters of the conditional and marginal processes, and are best exemplified by the strong conditions under which the system is fully cointegrating with rank s under any exogenous process integrated up to order d. The requisite condition can be shown to be

$$\begin{bmatrix} \mathbf{B}_{11}^{\{j\}} \\ \mathbf{B}_{21}^{\{j\}} \end{bmatrix} \mathbf{B}_{11}^{-1}[\mathbf{\Gamma}_{11}\,\mathbf{\Gamma}_{12}] = \begin{bmatrix} \mathbf{\Gamma}_{11}^{\{j\}} & \mathbf{\Gamma}_{12}^{\{j\}} \\ \mathbf{\Gamma}_{21}^{\{j\}} & \mathbf{\Gamma}_{22}^{\{j\}} \end{bmatrix}, \quad 0 \leqslant j \leqslant \min(d-1, p) \tag{18}$$

which generalises the condition of Corollary 1. Observe that in that case the restriction held automatically for $j = 0$ by virtue of the stability condition. It can be verified that (18) holds for $j = 0$ because of the insufficient targets condition that $\mathbf{B}$ and $[\mathbf{B}:\mathbf{\Gamma}]$ have the same rank; i.e. that $\mathbf{B}_{21}\mathbf{B}_{11}^{-1}\mathbf{\Gamma}_{1i} = \mathbf{\Gamma}_{2i}$, $i = 1, 2$.

In the inconsistent targets case, full cointegration with maximum rank is known to fail. We will merely illustrate the contention that any actual cointegrating vectors are formed as linear combinations of the targets, with weights are derived from the higher-order dynamic components of the system. For simplicity, take the case where the closed model with dynamic structure $\mathbf{B}(z)$ is balanced. It is then possible to write the system as

$$\begin{bmatrix} \mathbf{B}(\mathrm{L}) & \mathbf{\Gamma}(\mathrm{L}) \\ \mathbf{0} & \mathbf{\Theta}(\mathrm{L}) \end{bmatrix} \begin{bmatrix} \mathbf{y}_t \\ \mathbf{z}_t \end{bmatrix} = \begin{bmatrix} \tilde{\mathbf{B}}(\mathrm{L}) & \mathbf{\Gamma}(\mathrm{L}) \\ \mathbf{0} & \mathbf{\Theta}(\mathrm{L}) \end{bmatrix} \begin{bmatrix} \mathbf{y}_t^* \\ \mathbf{z}_t \end{bmatrix} = \begin{bmatrix} \mathbf{u}_t \\ \mathbf{v}_t \end{bmatrix} \tag{19}$$

where $|\tilde{\mathbf{B}}(z)| = 0$ has all its roots outside the unit circle, and

$$\mathbf{y}_t^* = \begin{bmatrix} \mathbf{y}_{1t} + \mathbf{B}_{11}^{-1}\mathbf{B}_{12}\mathbf{y}_{2t} \\ \mathbf{\Delta}_{22}(\mathrm{L})\mathbf{y}_{2t} \end{bmatrix} \tag{20}$$

by exact analogy with (16). Thus, in the transformed system the conditional sub-model is actually stable, and the analysis of Section II can be applied. Following the approach of Theorem 2 we can write $\tilde{\mathbf{\Pi}} = -\tilde{\mathbf{B}}^{-1}\mathbf{\Gamma}$, where the matrix $\tilde{\mathbf{B}}$ has its first s columns those of $\mathbf{B}$, and its last $n-s$ columns composed of the corresponding elements of one or other of the matrices $\mathbf{B}^{\{j\}}$, $j = 1, \ldots, p$. We must therefore address the issue of whether

$$\mathbf{y}_t^* - \tilde{\mathbf{\Pi}}\mathbf{z}_t = \begin{bmatrix} \mathbf{y}_{1t} + \mathbf{B}_{11}^{-1}\mathbf{B}_{12}\mathbf{y}_{2t} \\ \mathbf{\Delta}_{22}(\mathrm{L})\,\mathbf{y}_{2t} \end{bmatrix} + \begin{bmatrix} \mathbf{B}_{11} & \tilde{\mathbf{B}}_{12} \\ \mathbf{B}_{21} & \tilde{\mathbf{B}}_{22} \end{bmatrix}^{-1} \begin{bmatrix} \mathbf{\Gamma}_{11} & \mathbf{\Gamma}_{12} \\ \mathbf{\Gamma}_{21} & \mathbf{\Gamma}_{22} \end{bmatrix} \begin{bmatrix} \mathbf{z}_{1t} \\ \mathbf{z}_{2t} \end{bmatrix} \tag{21}$$

is I(0), and the first s elements of the vector in particular, of course, since these do not involve differences. Any fully cointegrating vectors which may exist for $(\mathbf{y}_{1t}', \mathbf{y}_{2t}', \mathbf{z}_{1t}', \mathbf{z}_{2t}')'$ must be obtained as linear combinations of the rows of

$$\begin{aligned} &[\mathbf{I}_s:\mathbf{B}_{11}^{-1}\mathbf{B}_{12}:\mathbf{E}^{-1}(\mathbf{\Gamma}_{11} - \tilde{\mathbf{B}}_{12}\tilde{\mathbf{B}}_{22}^{-1}\mathbf{\Gamma}_{21}):\mathbf{E}^{-1}(\mathbf{\Gamma}_{12} - \tilde{\mathbf{B}}_{12}\tilde{\mathbf{B}}_{22}^{-1}\mathbf{\Gamma}_{22})] \\ &\quad = \mathbf{E}^{-1}([\mathbf{B}_{11}:\mathbf{B}_{12}:\mathbf{\Gamma}_{11}:\mathbf{\Gamma}_{12}] - \tilde{\mathbf{B}}_{12}\tilde{\mathbf{B}}_{22}^{-1}[\mathbf{B}_{21}:\mathbf{B}_{22}:\mathbf{\Gamma}_{21}:\mathbf{\Gamma}_{22}]) \\ &\quad = \mathbf{E}^{-1}[\mathbf{I}: -\tilde{\mathbf{B}}_{12}\tilde{\mathbf{B}}_{22}^{-1}]\,\mathbf{C}[\bar{\mathbf{B}}:\bar{\mathbf{\Gamma}}] \end{aligned} \tag{22}$$

where $\mathbf{E} = \mathbf{B}_{11} - \tilde{\mathbf{B}}_{12}\tilde{\mathbf{B}}_{22}^{-1}\mathbf{B}_{21}$, and the first equality follows from the fact that $\mathbf{B}_{22} = \mathbf{B}_{21}\mathbf{B}_{11}^{-1}\mathbf{B}_{12}$. In this example, the weights given to each target relation in the construction of the cointegrating vector depend not only on the error correction coefficients **C**, but on the elements of $\tilde{\mathbf{B}}$, and hence on the coefficients of $\Delta^j \mathbf{y}_t$ for $j > 0$. We can interpret this by saying that a target relation is relatively strongly represented in (22) if the rate of adjustment towards it is relatively rapid, and conversely.

Recalling the earlier suggestion that an inconsistent targets model could represent a bargaining process, we note that the dynamics of the model would in this case represent the relative influence of different groups of agents in their response to exogenous environmental changes, determining the extent to which the different parties tend to prevail in the long run. The cointegrating vectors such as (22) might be identified with a Nash bargaining solution to the contest between agents pursuing inconsistent targets.

IV. AN ILLUSTRATION: WAGES AND PRICES

Wages and prices are a highly suggestive case for several reasons. Inflation in the United Kingdom is an I(1) series, so that the (log) price and wage levels are I(2); these variables are cointegrated but not *fully* cointegrated, for the log of real wages is also I(1). (See Hall, 1986 and Davidson and Hall, 1990). Further, in an economy with a permissive monetary policy and a floating exchange rate (as characterised the United Kingdom over extended periods of the last 20 years) there can apparently be no theory of the equilibrium price level; in our terminology, no-one is pursuing a *nominal* target. Imagine a macroeconomic system in which the targets related only to real wages – or equivalently, only to prices expressed in wage units. In such a model, the columns of $\bar{\mathbf{B}}$ corresponding to the logs of wages and prices respectively would be equal and opposite; hence, $\bar{\mathbf{B}}$ would be singular.

We are reminded that 20 years ago, in an influential paper, Lipsey and Parkin (1970) estimated a model of UK inflation with the general form

$$\dot{p} = a_0 + a_1 \dot{w} + a_2 \dot{m} + a_3 \dot{q} \tag{23}$$

$$\dot{w} = b_0 + b_1 \dot{p} + b_2 U + b_3 \dot{n} \tag{24}$$

where $\dot{p}$, $\dot{w}$, $\dot{m}$, $\dot{n}$ and $\dot{q}$ are the proportionate rates of change of retail prices, wage rates, import prices lagged one quarter, the percentage of the labour force unionised, and output per head, respectively, while U is the unemployment rate. This paper was the subject of a critique by Kenneth Wallis (Wallis, 1971), who pointed out amongst other things that when the equations were estimated by FIML, the estimation collapsed as the estimated matrix $\begin{bmatrix} 1 & -a_1 \\ -b_1 & 1 \end{bmatrix}$ tended to singularity.

Lipsey and Parkin's model, with its primitive dynamics, is now something of an historical curiosity, and results like Wallis's have rarely been replicated since. The interesting point is that the evidence reported by Wallis is consistent with an interpretation of his application of FIML as an attempt to estimate the target matrix for an unstable conditional subsystem. The fact that rates of change rather than log-levels are being modelled is immaterial, since differencing a system which generates I(2) variables (logs of w and p) as functions of I(1) variables does not eliminate the singularity; it means only that $\dot{w}$ and $\dot{p}$ are I(1) variables generated by I(0) variables.[11] Elsewhere (Davidson and Hall, 1990) we have developed a dynamic model containing target relationships similar in spirit to Lipsey and Parkin's equations, although set up in levels and enlarged by linking domestic to import prices through the exchange rate. While this model does not answer the crucial questions we posed concerning the origins of nominal instability, the empirical evidence that it is only the real wage which matters is very strong and echoes Wallis's earlier finding, in data which scarcely overlap the original.

V. CONCLUSION

In this paper we have set out a framework for the analysis of linear, dynamic macroeconomic systems generating integrated variables. We have shown that the phenomena of integration and cointegration are associated with certain restrictions on the coefficients of these systems. It remains a considerable challenge to devise well-founded theories of economic behaviour which embody these restrictions; but since we observe nonstationarity in mean as a generic property of macroeconomic time series, it is a challenge which cannot ultimately be bucked, unless the linear paradigm is itself to be abandoned.

London School of Economics

Bank of England

APPENDIX

Proof of Theorem 1. Define $\mathbf{\Pi}^*(L)$ by the relation $\mathbf{\Pi}(z) = \mathbf{\Pi}(z) = \mathbf{\Pi}^*(z)(1-z)$ (compare Engle and Granger (1987) equation (3.2), or Lemma 2.1 of Davidson (1990)). Rewrite (3) as

$$\mathbf{y}_t = \mathbf{\Pi z}_t + \mathbf{\Pi}^*(L)\,\Delta\mathbf{z}_t + \mathbf{B}(L)^{-1}\mathbf{u}_t,$$

where the matrix polynomial $\mathbf{\Pi z}_t$ is at most $I(d)$, and $\mathbf{\Pi}^*(L)\Delta\mathbf{z}_t + \mathbf{B}(L)^{-1}\mathbf{u}_t$ is at most $I(d-1)$ since both $\mathbf{\Pi}^*(z)$ and $\mathbf{B}(z)^{-1}$ are summable matrices by assumption. This implies (i), and since it also implies that $\mathbf{y}_t - \mathbf{\Pi z}_t$ is at most $I(d-1)$, using (6) yields (ii). ■

Proof of Theorem 2.

$$\begin{bmatrix} \mathbf{B}(L) & \mathbf{\Gamma}(L) \\ 0 & \mathbf{\Theta}(L) \end{bmatrix} \begin{bmatrix} \mathbf{y}_t \\ \mathbf{z}_t \end{bmatrix} = \begin{bmatrix} \mathbf{B}(L) & \mathbf{\Gamma}(L) \\ 0 & \mathbf{\Theta}(L) \end{bmatrix} \begin{bmatrix} \mathbf{I}_n & \mathbf{\Pi} \\ 0 & \mathbf{I}_m \end{bmatrix} \begin{bmatrix} \mathbf{I}_n & -\mathbf{\Pi} \\ 0 & \mathbf{I}_m \end{bmatrix} \begin{bmatrix} \mathbf{y}_t \\ \mathbf{z}_t \end{bmatrix}$$

$$= \begin{bmatrix} \mathbf{B}(L) & \mathbf{J}(L) \\ 0 & \mathbf{\Theta}(L) \end{bmatrix} \begin{bmatrix} \mathbf{y}_t - \mathbf{\Pi z}_t \\ \mathbf{z}_t \end{bmatrix}$$

$$= \begin{bmatrix} \mathbf{B}(L) & \mathbf{J}(L)\mathbf{\Theta}(L)^{-1} \\ 0 & \mathbf{I} \end{bmatrix} \begin{bmatrix} \mathbf{y}_t - \mathbf{\Pi z}_t \\ \mathbf{v}_t \end{bmatrix} = \begin{bmatrix} \mathbf{u}_t \\ \mathbf{v}_t \end{bmatrix}.$$

[11] Note: either L-P's Phillips-style wage inflation equation containing U is mis-specified (as we would argue) or w must depend on the integral of U.

Hence

$$\mathbf{B}(L)\,\bar{\mathbf{B}}^{-1}\mathbf{e}_t = \mathbf{B}(L)\,(\mathbf{y}_t - \boldsymbol{\Pi}\mathbf{z}_t) = \mathbf{u}_t - \mathbf{J}(L)\,\boldsymbol{\Theta}(L)^{-1}\mathbf{v}_t.$$

Under the necessary and sufficient conditions of the theorem, $\mathbf{J}(L)\,\boldsymbol{\Theta}(L)^{-1}\mathbf{v}_t$ is I(0); a pole at 1 in $\mathbf{J}(L)\,\boldsymbol{\Theta}(L)^{-1}$ is necessary to integrate an element of $\mathbf{v}_t$, and hence the absence of a pole is sufficient for $\mathbf{J}(L)\,\boldsymbol{\Theta}(L)^{-1}\mathbf{v}_t \sim I(0)$; but zero elements of $\mathbf{v}_t$ (corresponding to identities) are not ruled out, and hence it is only the absence of a complete row of integrating elements in $\mathbf{J}(L)\,\boldsymbol{\Theta}(L)^{-1}$ which is necessary.

Since the sum of two I(0) vectors is at most I(0), $\mathbf{B}(L)\,\bar{\mathbf{B}}^{-1}\mathbf{e}_t$ is $I(d'')$ for $d'' \leqslant 0$. But $|\mathbf{B}(z)| = 0$ has no unit roots, and hence $\mathbf{e}_t$ is integrated to the same order as $\mathbf{B}(L)\,\mathbf{B}^{-1}\mathbf{e}_t$ (by, e.g. Theorem 2.1 (ii) of Davidson (1990)). ■

Proof of Theorem 3. Consider a threefold partition for the system in (10) such that

$$\mathbf{A} = \begin{bmatrix} \mathbf{B} & \boldsymbol{\Gamma} \\ 0 & \boldsymbol{\Theta} \end{bmatrix} = \begin{bmatrix} \mathbf{B}_{11} & \mathbf{B}_{12} & \boldsymbol{\Gamma}_{11} & \boldsymbol{\Gamma}_{12} \\ \mathbf{B}_{21} & \mathbf{B}_{22} & \boldsymbol{\Gamma}_{21} & \boldsymbol{\Gamma}_{22} \\ \multicolumn{2}{c}{0} & \multicolumn{2}{c}{\boldsymbol{\Theta}} \end{bmatrix} \begin{matrix} s \\ n-s, \\ m \end{matrix}$$

where $\mathbf{B}_{11}$ is square of rank s.[12] Define the nonsingular matrix

$$\mathbf{T} = \begin{bmatrix} \mathbf{I} & 0 & 0 \\ -\mathbf{B}_{21}\mathbf{B}_{11}^{-1} & \mathbf{I} & 0 \\ 0 & 0 & \mathbf{I} \end{bmatrix} \begin{matrix} s \\ n-s \\ m \end{matrix}$$

such that

$$\mathbf{TA} = \begin{bmatrix} \mathbf{B}_{11} & \mathbf{B}_{12} & \boldsymbol{\Gamma}_{11} & \boldsymbol{\Gamma}_{12} \\ 0 & 0 & \boldsymbol{\Gamma}_{21}^{*} & \boldsymbol{\Gamma}_{22}^{*} \\ \multicolumn{2}{c}{0} & \multicolumn{2}{c}{\boldsymbol{\Theta}} \end{bmatrix} \begin{matrix} s \\ n-s, \\ m \end{matrix}$$

where $\boldsymbol{\Gamma}_{21}^{*} = \boldsymbol{\Gamma}_{21} - \mathbf{B}_{21}\mathbf{B}_{11}^{-1}\boldsymbol{\Gamma}_{11}$, $\boldsymbol{\Gamma}_{22}^{*} = \boldsymbol{\Gamma}_{22} - \mathbf{B}_{21}\mathbf{B}_{11}^{-1}\boldsymbol{\Gamma}_{12}$. Since the cointegrating vectors of the complete system lie in the space spanned by the rows of $\mathbf{A}$, they also lie in the space spanned by the rows of $\mathbf{TA}$. And since $\mathbf{z}_t$ is determined by the recursive subsystem $\boldsymbol{\Theta}(L)\,\mathbf{z}_t = \mathbf{v}_t$, we know that the cointegrating vectors for $\mathbf{z}_t$ lie in the space spanned by the rows of $\boldsymbol{\Theta}$, and their maximum number is $t = \text{rank}(\boldsymbol{\Theta})$. In particular, the rows of $[\boldsymbol{\Gamma}_{21}^{*}\,\boldsymbol{\Gamma}_{22}^{*}]$ $((n-s)\times m)$ cannot be cointegrating vectors for $\mathbf{z}_t$ unless they are linearly dependent on the rows of $\boldsymbol{\Theta}$. There are at most s cointegrating vectors for $(\mathbf{y}_t', \mathbf{z}_t')'$ in the space spanned by $[\mathbf{B}_{11}\,\mathbf{B}_{12}\,\boldsymbol{\Gamma}_{11}\,\boldsymbol{\Gamma}_{12}]$, proving (ii), and since these rows are linearly independent of the rows of $\boldsymbol{\Theta}$, (i) also follows. ■

References

Bewley, R. A. (1979). 'The direct estimation of the equilibrium response in a linear dynamic model.' *Economics Letters* vol. 3, pp. 357–62.

Currie, David (1981). 'Some long-run features of dynamic time series.' Economic Journal vol. 91, pp. 704–15.

Davidson, J. E. H. (1984). 'Error correction systems', mimeo, London School of Economics.

—— (1990). 'The cointegration properties of VAR models.' Forthcoming. *Journal of Time Series Analysis.*

[12] For the case $s = 0$ the partition would be twofold in the obvious way.

Davidson, J. E. H. and Hall, S. G. (1990). 'Cointegration in recursive systems: the structure of wage and price determination in the United Kingdom.' Working paper presented to RES/AUTE Conference, Nottingham.

Engle, R. F., Hendry, D. F. and Richard, J.-F. (1983). 'Exogeneity.' *Econometrica* vol. 51, pp. 277–304.

—— and Granger, C. W. J. (1987). 'Co-integration and error correction: representation, estimation and testing.' *Econometrica* vol. 55, no. 2 (March) pp. 251–76.

—— and Yoo, Byung Sam (1987). 'Forecasting and testing in cointegrated systems.' *Journal of Econometrics* vol. 35, no. 1, pp. 143–60.

Granger, C. W. J. (1969). 'Investigating causal relations by econometric models and cross-spectral methods.' *Econometrica* vol. 52, pp. 424–38.

—— (1981). 'Some properties of time series data and their use in econometric model specification.' *Journal of Econometrics* vol. 16, pp. 121–30.

—— (1986). 'Developments in the study of cointegrated economic variables.' *Oxford Bulletin of Economics and Statistics* vol. 48, no. 3, pp. 213–28.

Hall, S. G. (1986). 'An application of the Granger and Engle two-step estimation procedure to United Kingdom aggregate wage data.' *Oxford Bulletin of Economics and Statistics* vol. 48, no. 3, pp. 229–39.

Hatanaka, M. (1975). 'On the global identification of the dynamic simultaneous equations model with stationary disturbances.' *International Economic Review* vol. 16, pp. 545–54.

Hendry, D. F. (1986). 'Econometric modelling with cointegrated variables: an overview.' *Oxford Bulletin of Economics and Statistics* vol. 48, no. 3, pp. 201–12.

Hsiao, Cheng (1983). 'Identification.' Chapter 4 of *Handbook of Econometrics*, vol. 1, (Zvi Griliches and Michael D. Intriligator, eds.) North-Holland.

Johansen, Søren (1988*a*). 'The mathematical structure of error correction models.' *Contemporary Mathematics* vol. 80 (American Mathematical Society), pp. 359–85.

—— (1988*b*). 'Statistical analysis of cointegration vectors.' *Journal of Economic Dynamics and Control* vol. 12, pp. 231–54.

Kloek, T. (1985). 'Dynamic adjustment when the target is nonstationary.' *International Economic Review* vol. 25, pp. 315–26.

Lipsey, R. G. and Parkin, J. M. (1970). 'Incomes policy: a re-appraisal.' *Economics* vol. 37 (May), pp. 115–38.

Pagan, A. (1985). 'Time series behaviour and dynamic specification.' *Oxford Bulletin of Economics and Statistics* vol. 47, no. 3 (August) pp. 199–212.

Patterson, K. (1983). 'Growth coefficients in error correction and autoregressive distributed lag models.' Bank of England Discussion Papers (Technical Series) No. 2.

—— and Ryding, J. (1984). 'Dynamic time series models with growth effects constrained to zero.' ECONOMIC JOURNAL vol. 94, pp. 137–43.

Phillips, P. C. B. and Durlauf, S. N. (1986). 'Multiple time series regression with integrated processes.' *Review of Economic Studies.* vol. 53(4) (August) pp. 473–96.

Salmon, Mark (1982). 'Error correction mechanisms.' ECONOMIC JOURNAL vol. 92, pp. 618–29.

Wallis, K. F. (1971). 'Wages, prices and incomes policies: some comments.' *Economica* vol. 38 (August), pp. 304–10.

—— (1977). 'Multiple time series and the final form of econometric models.' *Econometrica* vol. 45, pp. 1481–97.