Title
Self-Generating Variables in a Cointegrated VAR Framework

Authors
Granger, Clive W.J.
YOON, GAWON

Publication Date
2001-02-20

2001-04

UNIVERSITY OF CALIFORNIA, SAN DIEGO

DEPARTMENT OF ECONOMICS

SELF-GENERATING VARIABLES IN A COINTEGRATED VAR FRAMEWORK

BY

CLIVE W.J. GRANGER

AND

GAWON YOON

DISCUSSION PAPER 2001-04
FEBRUARY 2001

Self-generating variables in a cointegrated VAR framework

February 20, 2001

Clive W.J. Granger
Department of Economics, 0508
University of California San Diego
La Jolla, CA 92093-0508
USA
cgranger@ucsd.edu

Gawon Yoon*
Department of Economics
Pusan National University
Pusan, 609-735
S. Korea
gyoon@pusan.ac.kr

* Corresponding author.

Abstract

A variable is defined to be self-generating if it can be forecast efficiently from its own past only. Conditions are derived for certain linear combinations to be self-generating in error correction models. Interestingly, there are only two candidates for self-generation in an error correction model. They are cointegrating relationships and common stochastic trends defined by Gonzalo and Granger (1995). The usefulness of self-generation as a multivariate-modelling tool is investigated. A simple testing procedure is also presented. Some interesting economic hypothesis can be easily tested in the self-generation framework. For example, for forward exchange rate to have forecasting power for the future movements in spot rate, the latter should not be self-generating. Given that they are cointegrated, the spot exchange rate should not be a common stochastic trend, which can be easily tested. We also provide additional examples.

Keyword: Self-generation, cointegrated VAR, common stochastic trends, aggregation of time series, efficiency of forecasting

JEL classification: C22, C32

Contents

1. Introduction

Suppose initially that X_t is a vector of stationary series with six components to be modelled by a vector autoregression, or VAR. Thus, there will be six equations, each of which will have one variable, at time *t*, explained by itself and also each of the other variables at lags 1, 2, …, *p*. Suppose, for simplicity, *p* is the same for each variable and takes the modest value *p* = 4. Then each equation will contain 24 parameters, with matching t-values, giving 144 parameters plus t-values for the whole system, which is nearly 300 statistics to be reported apart from the covariance matrix of residuals and any further evaluation statistics. Not only is this a large estimation burden but, if presented on a printed page, is too much information for easy assimilation. In fact, two famous papers involving six variables, a four lag VAR, Sims (1980) and King *et al.* (1991), do not report any of the parameter values for the model, just a few summary statistics. On the other hand, VARs, or their modern reformulation as error-correction model, or ECM, are the major tools for investigating linear relationships between small groups of economic variables. It is the object of this paper to ask if simplified, more easily interpreted versions of the models are available on some occasions.

Write the original VAR as

$$X_t = A(B)X_t + e_t \tag{1}$$

$X_t = (x_{1t}, x_{2t}, \ldots, x_{nt})'$, so that X_t has n components with mean 0, for simplicity, and $A(B) = A_1B + A_2B^2 + \cdots + A_pB^p$, where A_j is an $n \times n$ matrix. Consideration will be given to instantaneous transformations of the original vector X_t into a new vector Y by

$$Y_t = QX_t,$$

where Q is an $n \times n$ matrix and will be assumed to have an inverse, so that $X_t = Q^{-1}Y_t$

and equation (1) now becomes

$$Y_t = G(B)Y_t + e_t$$

where $G(B) = QA(B)Q^{-1}$ and $e_t = Qe_t$. It will be asked if Q can be chosen so that $G(B)$ is somehow simpler or easier to interpret than $A(B)$. For example, the lags involved in $G(B)$ could be less or there could be blocks of zeros. In this paper, use is made of the following definition.

Definition: Self-generation

A component of X_t is said to be self-generating (in mean) if

$$E(x_{i,t+1} \mid x_{i,t-j},\ j \geq 0) = E(x_{i,t+1} \mid X_{t-j},\ j \geq 0),$$

so that $x_{jt}, j \neq i$, do not (Granger) cause $x_{i,t+1}$.

Throughout the paper, “causality” will be taken to mean “Granger causality”, as discussed in Granger (1969). It may be noted that a special case of self-generation is when $x_{i,t+1}$ is a martingale difference sequence and so is unforecastable from $X_{t-j}, j \geq 0$, so that $E(x_{i,t+1} \mid X_{t-j}, j \geq 0) = 0$. For ease of exposition this will be called the “white noise” case, although the above definition is more precise. The search for self-generation will be limited to the contemporaneous aggregation of data series. It is possible that while no contemporaneous aggregate is self-generating, a non-contemporaneous one is.[1]

From standard discussion on causality in VAR models, such as in Lütkepohl

[1] See equation (1) in Ericsson (1993) for an example.

(1991), self-generation will involve sections of zeros in $A(B)$. Kohn (1982) proves some useful results for self-generating processes. Let X_t have a moving average representation,

$$X_t = e_t + \sum_{j=1}^{\infty} C_j e_{t-j} \qquad (2)$$

where C_j is an $n \times n$ matrix, each j, and e_t is a white noise vector and define a single series $y_t = d' X_t$, where d is a vector of n components, then y_t is self-generating so that

$$E(y_{t+1} \mid y_{t-j}, j \geq 0) = E(y_{t+1} \mid X_{t-j}, j \geq 0)$$

if and only if

$$d' C_j = k_j d', \text{ all } j,$$

where k_j is a sequence of constants, in which case

$$y_t = e_t + \sum_{j=1}^{\infty} k_j e_{t-j}$$

where e_t is a univariate white noise. Write equation (2) in a VAR form, assuming that all X_t components are stationary,

$$X_t = \sum_{j=1}^{p} A_j X_{t-j} + e_t,$$

where e_t is a vector of white noise, then $y_t = d' X_t$ is self-generating if and only if

$$d' A_j = a_j d', \ j = 1, \ldots, p, \qquad (3)$$

where a_j is a sequence of constants, in which case

$$y_t = \sum_{j=1}^{p} a_j y_{t-j} + e_t.$$

Thus d is an eigenvector for each A_j', with a_j being the corresponding eigenvalue. Kohn also points out that if there is no one-step forecastability, then there will be no k-

step forecastability, $k > 1$.

A special case of interest is that of $p=1$, so that X_t obeys a first order (Markov) vector autoregressive model, then because of equation (3), there will always exist a vector d so that $y_t = d'X_t$ for every component that is self-generating. In this case the original X_t component can all be written as a linear combination of n self-generating series. It is generally the case that a linear combination of self-generating process is not self-generating. If x_{1t} and x_{2t} are both white noise, so unforecastable from X_{t-j}, $j>0$, then a linear combination of these variables will also be a white noise. Other than this very simple case, the conditions for the linear combinations to be self-generating seem to be very stringent: suppose that x_{1t} and x_{2t} have the following VAR(p) representation;

$$\begin{pmatrix} x_{1t} \\ x_{2t} \end{pmatrix} = \sum_{i=1}^{p} \begin{pmatrix} a_i & 0 \\ 0 & b_i \end{pmatrix} \begin{pmatrix} x_{1,t-i} \\ x_{2,t-i} \end{pmatrix} + \begin{pmatrix} e_{1t} \\ e_{2t} \end{pmatrix}.$$

The off-diagonal elements in the coefficients matrices should be zero because both x_{1t} and x_{2t} are assumed to be self-generating. Using the Kohn's result for a linear combination $y_t \equiv \alpha_1 x_{1t} + \alpha_2 x_{2t}$ to be self-generating, the condition becomes:

$$\begin{pmatrix} \alpha_1 & \alpha_2 \end{pmatrix} \begin{pmatrix} a_i & 0 \\ 0 & b_i \end{pmatrix} = m_i \begin{pmatrix} \alpha_1 & \alpha_2 \end{pmatrix},$$

where m_i is some scalar sequence. This implies that $a_i = b_i = m_i$ for $i = 1, \ldots, p,$ and this requires that all three processes have the same coefficients for all lags.

The above definition of self-generation can be easily extended to a vector x_t, which is a subset of X_t, being self-generating by requiring that

$$E(x_{t+1} \mid x_{t-j}, j \geq 0) = E(x_{t+1} \mid X_{t-j}, j \geq 0).$$

The definition is obviously of no interest if x_t and X_t are identical. Kohn points out that the definition will also apply to I(1) processes, which is easily expressed in the moving average form. Further if a variable x_{jt} is self-generating, it will follow as obvious that Δx_{jt} is also self-generating.

The remainder of the paper is organized as follows: in section 2, the self-generating variables will be discussed in a cointegrated VAR framework. In section 3, some remarks on self-generating variables, including a brief literature review, will be made. In section 4, the concept of partially commuting matrices will be introduced and a simplifying structure in error correction models will be discussed. In section 5, testing procedures for self-generation will be introduced. In section 6, several economic examples will be discussed in the self-generation framework. Section 7 concludes the paper. Appendix contains more detailed discussion on the condensed form of partially commuting matrices.

2. Cointegration and self-generating variables

Suppose now that X_t is a vector with n components, each of which is I(1), so that the differenced series are (second order) stationary. Further suppose that there are r linearly independent vectors a_j for $j = 1, \ldots, r$ such that $a_j' X_t$ is stationary; a_j is known as a cointegrating vector and the $n \times r$ matrix a with a_j in the j th column is the cointegrating matrix. It follows that X_t can be thought of having a (linear) data generating mechanism which is an error-correction model:

$$\Delta X_t = -ga' X_{t-1} + \sum_{j=1}^{p-1} A_j^* \Delta X_{t-j} + e_t \qquad (4)$$

where A_j^* are $n \times n$ matrices. Assume that the $n \times n$ matrix $(a \quad g_\perp)$ is nonsingular[2], where $g_\perp$ is an $n \times (n-r)$ matrix such that $g_\perp' g$ = 0. We set $m \equiv (n-r)$. Discussions of cointegration and error-correction models can be found in the collection of papers by Engle and Granger (1991). Clearly equation (4) can be rewritten as a VAR in levels, but with some constraints on the parameters.[3]

As we are looking for transformations of the form QX_t that produce interesting properties, there is obviously no problem with starting to search with a convenient transformation of this from. The transformation that will be used is

$$Y_t = MX_t$$

where

$$M = \begin{pmatrix} g_\perp' \\ a' \end{pmatrix}.$$

It is also convenient to define

$$W_t = g_\perp' X_t \text{ and}$$
$$z_t = a' X_t.$$

Assuming M has an inverse, X_t can be written as a linear combination of W_t and z_t. It should be noted that z_t is I(0) if X_t is cointegrated and that W_t is I(1) with $g_\perp \neq a$. W_t thus becomes a candidate for a common stochastic trend and has been proposed by

---

[2] The matrix is nonsingular when det($a'g$) $\neq 0$. It ensures that an $r \times r$ matrix $g_\perp' a_\perp$ has a full rank. See Gonzalo and Granger (1995).

[3] In this paper, we use error correction models exclusively to investigate the self-generation property. However, the discussion can be also made in the triangular representation of Phillips (1991). The relationship between an error correction model and triangular representation is studied in Cappuccio and Lubian (1996).

Gonzalo and Granger (1995) as a particularly convenient estimate of a common stochastic trend. It is easily estimated and has various other useful properties, some of which are mentioned below.[4] Note that $g_{\perp}$ is identified up to a nonsingular matrix multiplication to the left.

Multiplying equation (4) by M gives two block of equations: the first m equations explains W_t in terms of lagged z_t and W_t (after substituting X_t for z_t and W_t) and the final r equations explain z_t in terms of lagged z_t and W_t. Several self-generation questions can be asked.

Q1. Can the z_t block be self-generating, so that W_t does not appear in those equations?

Q2. Can the W_t vector be self-generating?

Q3. If the vector z_t is self-generating, can one member z_{jt} be self-generating in relation to the other z 's?

Q4. Similarly, if W_t is self-generating, can W_{jt} be self-generating with respect to the other W 's?

Q5. Can any other linear combinations of X 's be self-generating?

It is convenient to interpret these questions using the error-correction model, equation (4), after the transformation. Premultiplying equation (4) with $g_{\perp}^{'}$ and $a^{'}$, respectively, we will get the following two equations of common stochastic trends and cointegrating relationships:

---

[4] Gonzalo and Granger (1995) define common stochastic trends as a linear combination of data series and require them not to be Granger caused by the short-run deviation from equilibrium, z_t , at zero frequency. See Granger and Lin (1995) for more discussion.

$$\Delta W_t = \sum_{i=1}^{p-1} g_{\perp}^{'} A_i^{*} \Delta X_{t-i} + g_{\perp}^{'} e_t \qquad (5)$$

and

$$\Delta z_t = -a^{'} g\, z_{t-1} + \sum_{i=1}^{p-1} a^{'} A_i^{*} \Delta X_{t-i} + a^{'} e_t . \qquad (6)$$

Using the relation

$$X_t = a_{\perp}(g_{\perp}^{'} a_{\perp})^{-1} W_t + g\,(a^{'} g)^{-1} z_t ,$$

the two equations (5) and (6) can be collected as

$$\begin{pmatrix} W_t \\ z_t \end{pmatrix} = \begin{pmatrix} I_{n-r} & 0 \\ 0 & I_r - a^{'} g \end{pmatrix} \begin{pmatrix} W_{t-1} \\ z_{t-1} \end{pmatrix} + \sum_{i=1}^{p-1} \begin{pmatrix} g_{\perp}^{'} A_i^{*} a_{\perp} (g_{\perp}^{'} a_{\perp})^{-1} & g_{\perp}^{'} A_i^{*} g\,(a^{'} g)^{-1} \\ a^{'} A_i^{*} a_{\perp} (g_{\perp}^{'} a_{\perp})^{-1} & a^{'} A_i^{*} g\,(a^{'} g)^{-1} \end{pmatrix} \begin{pmatrix} \Delta W_{t-i} \\ \Delta z_{t-i} \end{pmatrix} + \begin{pmatrix} g_{\perp}^{'} \\ a^{'} \end{pmatrix} e_t . \qquad (7)$$

This representation will be used in the proof of theorem 3 below.

The five questions can now be considered in turn. The first is answered in the following theorem.

Theorem 1

Let $z_t = a^{'} X_t$ be the vector of cointegrating relationships. Then the vector z_t is self-generating if and only if

$$a^{'} A_i^{*} = m_i a^{'} \qquad (8)$$

for some $r \times r$ matrices m_i, $i = 1,\ldots, p-1$. If equation (8) holds, then z_t is a p th order autoregressive process, with all roots lying outside the unit circle.

Proof:

(a) Necessity: By the Wold decomposition, stationary ΔX_t and cointegrating relationship z_t processes can be represented as

$$\Delta X_t = C(B)e_t = (I + \sum_{i=1}^{\infty} C_i B^i)e_t \text{ and}$$

$$z_t = h_t + \sum_{i=1}^{\infty} f_i h_{t-i},$$

where f_i is an $r \times r$ matrix of coefficients and h_t is an $r \times 1$ vector of white noise. Also, cointegration restrictions imply that

$$z_t = a^{'} C^*(B)e_t = a^{'}(C_0^* + \sum_{i=1}^{\infty} C_i^* B^i)e_t = a^{'}(I_n + \sum_{i=1}^{\infty} C_i^* B^i)e_t, \qquad (9)$$

where $C(B) = C(1) + C^*(B)(1-B)$ and $C_j^* = -\sum_{i=j+1}^{\infty} C_i,\ j = 0,1,\ldots$. The last equality of equation (9) follows from the fact that $a^{'} C_0^* = a^{'}$, which is derived from the cointegration restrictions, $a^{'} C(1) = 0$. It is assumed that $C(B)$ is 1-summable and therefore $C^*(B)$ is absolutely summable. Self-generation implies that $h_t = a^{'} e_t$. So

$$z_t = h_t + \sum_{i=1}^{\infty} f_i h_{t-i} = a^{'} e_t + \sum_{i=1}^{\infty} f_i a^{'} e_{t-i}.$$

Since $z_t = a^{'}(I_n + \sum_{i=1}^{\infty} C_i^* B^i)e_t$, it follows that for $i = 1,2,\ldots$

$$f_i a^{'} = a^{'} C_i^*.$$

Using the fact that $\{I - A(B)\}C(B) = I - B,$ it can be shown that $a^{'} A_i^* = m_i a^{'}$ for some $r \times r$ matrices m_i, $i = 1,\ldots,p-1.$

(b) Sufficiency: Conversely suppose that $a^{'} A_i^* = m_i a^{'}$ or $a^{'} C_i^* = k_i a^{'}$ for some $r \times r$ matrices k_i, $i = 1,\ldots,p-1.$ Then

$$z_t = a^{'} e_t + \sum_{i=1}^{\infty} a^{'} C_i^* e_{t-i} = a^{'} e_t + \sum_{i=1}^{\infty} k_i a^{'} e_{t-i}.$$

We want to show that this is the Wold representation of z_t and $h_t = a^{'} e_t$. Suppose now that $1+\sum k_i z^i = 0$. Then,

$$0 = (1+\sum k_i z^i)a' = a'(I+\sum C_i^* z^i).$$

Therefore, $\det(I+\sum C_i^* z^i)=0$, implying that $|z|\geq 1$ from equation (9). Since the Wold representation is unique, it follows that $a'e_t = h_t$ and $k_i = f_i$. The remaining part of the theorem easily follows from the fact that z_t is stationary, assuming that $m_{p-1}\neq 0$.

¦

Under the hypothesis that z_t is self-generating, the newly transformed VAR process, equation (6), becomes

$$z_t = (I - a'g + m_1)z_{t-1} + (m_2 - m_1)z_{t-2} + \cdots + (m_{p-1} - m_{p-2})z_{t-p+1} - m_{p-1}z_{t-p} + a'e_t\,.$$

In answering the second question, a more general situation will first be considered in the following theorem:

Theorem 2

Let $M^* = \begin{pmatrix} S_m' \\ a' \end{pmatrix}$ where S_m is an $n\times m$ matrix and $y_t = M^* X_t$. Then $W_t^* = S_m' X_t$ can be self-generating if and only if

$$(a)\ \ S_m = g_\perp \text{ and}$$
$$\text{(b)}\ \ g_\perp' A_i^* = g_i g_\perp'$$

for some sequence of $m\times m$ matrices g_i.

Proof:

(a) follows informally by consideration of the error correction model. If $S_m \neq g_\perp$, then the m equations for the components of ΔW_t^* will contain z_{t-1} terms, so immediately are seen not to be self-generating. As there are only a finite $(p-1)$ number of lagged differenced terms, conditions on these differenced Δz_{t-j} terms cannot compensate for the presence of z_{t-1}. If $S_m = g_{\perp}$, then it will automatically be true that z_{t-1} will not enter the equations for ΔW_t, so this is a necessary condition for self-generation. Using the Kohn's results and a similar argument as in the Theorem 1, another condition (b) is determined. A formal proof is available but not given here. ¦

If W_t is self-generating, then the equation (5) becomes

$$\Delta W_t = \sum_{i=1}^{p-1} g_i \Delta W_{t-i} + g_{\perp}^{'} e_t \,. \qquad (10)$$

If z_t and ΔW_t are either self-generating, then particular linear combinations of them[5], such as $d_r^{'} z_t$ or $d_m^{'} \Delta W_t$, can be self-generating in relation to the other components of z_t or ΔW_t? d_r and d_m are $1 \times r$ and $1 \times m$ vectors, respectively. For instance, consider the self-generating ΔW_t process, equation (10); for $d_m^{'} \Delta W_t$ to be self-generating, the required conditions are

$$d_m^{'} g_i = a_i^* d_m^{'}$$

for constants a_i^* for $i = 1, \dots, p-1$. Combined with the initial conditions for W_t to be self-generating, the above requirement for $d_m^{'} \Delta W_t$ to be self-generation becomes

$$(g_{\perp} d_m)^{'} A_i^* = a_i^* (g_{\perp} d_m)^{'} \,. \qquad (11)$$

For $d_r^{'} z_t$ to be self-generating,

$$d_r^{'} (m_{i-1} - m_i) = a_i d_r^{'} \,, \; i = 2, \dots, p \qquad (12)$$

---

[5] Recall that a linear combination of self-generating variables is generally not self-generating as previously mentioned.

with $m_p \equiv 0$ and

$$d_r^{'}(-m_1 - I + a^{'}g) = a_1 d_r^{'} \qquad (13)$$

Some algebra simplifies equation (13) to

$$d_r^{'}(a^{'}g) = (m+1)d_r^{'}$$

where $m = \sum_{j=1}^{p} a_j$ so that $d_r^{'}$ is an eigenvector of $a^{'}g$ and $m+1$ is an eigenvalue, and with the further constraints equation (12) also operating for $i \geq 2$. Thus the number of possible self-generating combinations of z's is limited to consideration of the eigenvectors of $a^{'}g$.

The final question is answered in the following theorem.

Theorem 3

Suppose that q_t is a univariate series, $q_t = k^{'}X_t$, where k is an $n \times 1$ vector, expressed as

$$q_t = \begin{pmatrix} d_w^{'} & d_z^{'} \end{pmatrix} \begin{pmatrix} W_t \\ z_t \end{pmatrix},$$

then for q_t to be self-generating, either $d_z = 0$ or $d_w = 0$.

Proof.

Using the expression in equation (7) and Kohn's results in equation (8), a necessary condition for self-generation is that

$$\begin{pmatrix} d_w^{'} & d_z^{'} \end{pmatrix} \begin{pmatrix} I_{n-r} & 0 \\ 0 & I_r - a^{'}g \end{pmatrix} = n \begin{pmatrix} d_w^{'} & d_z^{'} \end{pmatrix},$$

where n is given by

$$q_t = l(B)q_{t-1} + z_t \text{ and}$$

$$n = l(0),$$

where z_t is an impulse of q_t. This gives the pair of sets of equations

$$(1-n)d_w^{'} = 0 \text{ and}$$

$$d_z^{'}(I - a^{'}g) = n d_z^{'}.$$

If $n \neq 1$, then $d_w = 0$, and the theorem is proved. If $n = 1$, the second set of equations give

$$d_z^{'}(a^{'}g) = 0$$

but as $a^{'}g$ has been assumed to be not singular for the theory of cointegration, it follows that $d_z = 0$, again proving the theorem.

It should be noted that if a different definition of the common stochastic trend is used, so that the new W_t includes z_t in its generation, this proof does not hold and possibly the theorem. Also in line with the questions posed at the start of this paper, only linear combinations of X_t are considered, which may consist of combinations of I(0) or I(1) variables. It follows that both z_t and W_t can not be white noise, so the rules for combining self-generating variables discussed in the first section do not apply. The possibility of finding linear combinations of z_t and ΔW_t that are self-generating is a different question as lagged X_t are also involved and will not be considered here, but in this case the various elements could be white noise, for example if W_t were random walks.

To summarize this section, both z_t and W_t can be self-generating, with conditions given in Theorem 1 and 2 and they are the only non-trivial linear combinations of X_t that can be self-generating. Linear combinations of z_t and W_t or

ΔW_t can also be self-generating according to extra conditions, equations (11) to (13).

3. Some remarks

We consider a special case in which both cointegrating relationships and common stochastic trends are self-generating and explore the relationship between self-generation and serial correlation common features. We also provide a brief review of previous literature on forecasting efficiency from full and aggregate information sets for both cointegrated and non-cointegrated series.

3.1 A special case

There is a special case worth considering; a first-order cointegrated system of

$$\Delta X_t = -ga' X_{t-1} + e_t.$$

It can be easily shown that both ΔW_t and z_t are self-generating; ΔW_t becomes white noise and z_t is stationary. Furthermore, if $a'g = I$, z_t is also white noise. Note that cointegration requires only that det($a'g$) $\neq$ 0. The following example used in Engle and Granger (1987) shows the special case. Assume that the two variables x_{1t} and x_{2t} are jointly generated as follows:

$$x_{1t} + bx_{2t} = u_{1t}, \;\; u_{1t} = u_{1t-1} + e_{1t}$$

$$x_{1t} + ax_{2t} = u_{2t}, u_{2t} = r u_{2t-1} + e_{2t}$$

where $e_{ú\, t}$ and $e_{ú\, t}$ are possibly correlated white noise disturbances and $|r|$ú ú . The ECM for the data generating process becomes,

$$\Delta x_{1t} = bd\,(x_{1t-1} + a\,x_{2t-1}) + h_{1t}$$

$$\Delta x_{2t} = -d\,(x_{1t-1} + a\,x_{2t-1}) + h_{2t}\,,$$

where $d = (1 - r)/(a - b)$ and h_{1t} and h_{2t} are linear combinations of white noise processes. Note that $a \neq b$. Clearly, $W_t\,(= x_{1t} + bx_{2t})$ is self-generating because it is random walk. It can be easily shown that cointegrating relationship, $z_t\,(= x_{1t} + ax_{2t})$, is also self-generating. If $r = 0$, z_t becomes white noise.

3.2 Comparison with serial correlation common features

We will only briefly mention the relationship between self-generation and serial correlation common features of Engle and Kozicki (1993). If a linear combination eliminates serial correlation in the original data series, so that the new variable becomes an innovation relative to the past information set, it is called a serial correlation common feature. Denote the serial correlation common feature by $\hat{a}$. From the error correction model in equation (4), for $\Delta \hat{a}' X_t$ to be serial correlation common features,

$$\hat{a}' A_j^* = 0,\ \ j = 0, \ldots, p-1 \text{ and } \hat{a}' g = 0,$$

so that $\Delta \hat{a}' X_t = \hat{a}' e_t$ is a white noise process. Serial correlation common features are an extreme case of self-generation in that they require the eigenvalues of A_j^* to be jointly null. It is also easy to infer conditions under which cointegrating relationships become a white noise process;

$$a' A_j^* = 0,\ j = 0, \ldots, p-1 \text{ and } a' g = I \,.$$

In the previous bivariate example in subsection 3.1, $(1, b)'$ is a cofeature vector. Additionally, if $r = 0$, $(1, a)'$ is also a cofeature vector, since z_t becomes white noise.

3.3 A brief literature review

For $y_t = d' X_t$, the set which is composed of the past values of X_t, $\{X_{t-j},\ j > 0\}$, is called a full or disaggregate information set, while the set of the past values of

y_t only, $\{y_{t-j}, \ j > 0\}$, is an aggregate information set. A great number of literature exists on the efficiency comparison of full and aggregate information sets. We will not try to be exhaustive. Many of the papers cited here contain more references. In theory at least, more information in general leads to improved forecasts. However, due to difficulties associated with model specification and estimation among others, no general consensus is reached among various empirical results. For instance, Fair and Shiller (1990) find that in forecasting real GNP growth, VAR models dominate autoregressive ones. In contrast, Bodo *et al.* (2000) find that the forecast of an aggregate ARIMA model beats that from disaggregate models in predicting industrial production of 11 countries in Euro area. Forecasting efficiency is compared in Rose (1977), Tiao and Guttman (1980), Wei and Abraham (1981), Lütkepohl (1984), among others. Kohn (1982) derives necessary and sufficient conditions for aggregate forecasts to be as efficient as disaggregate ones, as previously discussed in section 1 of this paper. He also provides a very simple testing procedure, which will be discussed below. In this paper, we extend his framework to the cointegrated vector autoregressive framework and provide additional results. Fliedner (1999) examines demand forecasts for firms maintaining thousands of items and discusses strategies that reduce forecast burden by considering families of items. He also contains a useful review of previous literature on the subject, especially from the hierarchical forecasting point of view.

If the variables of our interest are not self-generating, full information in general will produce better forecasts than an aggregate one does. However, it is not clear which forecasting method should be employed among, for instance, an unrestricted VAR, a VAR in first differences or an error correction model. Furthermore, if one method produces the better forecasts of X_t, it does not necessary follow that its forecasts are also better for its linear combinations, $a'X_t$ for example. Evidence from Monte Carlo simulations is of little help. For example, the simulation results in Clements and Hendry (1995) show that forecasts of X_t from a VAR in first differences are at par with those from an unrestricted VAR or an ECM, but that its forecasts of linear combinations of X_t are much worse. However, in an example regarding the demand for UK M1, a VAR in first differences performs very well in predicting a cointegrating relationship, contrary to their simulation results. Similar results are also reported in Hoffman and Rasche (1996)[6] and Clark (2000). Clearly, more works need to be done on this area.

The recent paper by Clark (2000) deserves special comment. He derives forecast equivalency conditions, as stated in theorem 2 above, among cointegrated variables and suggests a testing procedure. He also examines the effect of aggregation on forecasting with simulations in a cointegrated VAR framework. However, he mentions little on cointegrating relationship as a candidate for self-generation. As the examples in section 6 below show, there are many interesting economic hypotheses regarding cointegrating relationships as well as common stochastic trends. For instance, one would be interested in forecasting the spread between interest rates that have been

[6] They compare the forecasting ability of different approaches for the three cointegrating relationships, money demand relation, the Fisher relations and interest rate differential, among five variables of real money balances, inflation rates, commercial paper rates, real GDP, and the T-bill rate. The VAR in differences forecasts the money demand relation best, while ECMs forecast the remaining two cointegrating relationships best.

known to posses some forecasting power about future economic movements. Since the spreads are cointegrating relationships, it will be interesting to examine if they are self-generating. In this paper, we present a coherent framework to investigate self-generation properties of cointegrated variables, which turns out to be very useful in testing many interesting economic hypotheses.

4. Self-generation, partially commuting matrices, and error correction models

The conditions for self-generation require that the coefficient matrices in a VAR or an ECM have a common eigenvector. Interestingly, we can further simplify the error correction model in this case. For simplicity, we assume only two coefficient matrices A_1 and A_2 in the following discussion. First, a definition from Alpin *et al.* (2000).

Definition

Two complex $n \times n$ matrices A_1 and A_2 are said to be partially commuting if they have a common eigenvector.

If two matrices have a (nonzero) common eigenvector, then there exist nontrivial common invariant subspaces of A_1 and A_2, on which these matrices commute. The span of the common eigenvector is one of such subspaces. Shemesh (1980) gave a criterion for two matrices to have a common eigenvector. Let $[A_1, A_2]$ denote the commutator of A_1 and A_2 such that $[A_1, A_2] = A_1A_2 - A_2A_1$[7]. The symbol $M_n(\mathbb{C})$ stands for the space

[7] We list some properties of the commutator from Cohen-Tannoudji *et al.* (1977), p.168.

$[A_1, A_2] = -[A_2, A_1]$

of complex $n \times n$ matrices.

Theorem (Shemesh, 1980)

Let A_1, $A_2 \in M_n(\mathbb{C})$. Then A_1 and A_2 are partially commuting if and only if the subspace $\aleph$ defined by the formula

$$\aleph = \bigcap_{k,l=1}^{n-1} \ker[A_1^k, A_2^l]$$

is nontrivial.

The subspace $\aleph$ is invariant with respect to both A_1 and A_2, and they commute on $\aleph$. Such matrices are reducible by the same similarity to triangular form; there exists a nonsingular matrix T such that $T^{-1}A_1T$ and $T^{-1}A_2T$ are triangular and its first column is an eigenvector for A_1 (and for A_2 as well). Alpin *et al.* show that partially commuting matrices A_1 and A_2 can be brought by a simultaneous similarity transformation to block triangular matrices with special structure;

$$R = P^{-1}A_1P \text{ and } S = P^{-1}A_2P \qquad (14)$$

where

$[A_1, A_2 + A_3] = [A_1, A_2] + [A_1, A_3]$

$[A_1, A_2A_3] = [A_1, A_2]A_3 + A_2[A_1, A_3]$

$[A_1, [A_2, A_3]] + [A_2, [A_3, A_1]] + [A_3, [A_1, A_2]] = 0$

$[A_1, A_2]^H = [A_2^H, A_1^H]$, where A^H is the adjoint operator or Hermite adjoint of *A*.

$$R = \begin{pmatrix} R_{11} & R_{12} & \cdots & R_{1,k-1} & R_{1k} \\ & R_{22} & \cdots & R_{2,k-1} & R_{2k} \\ & & \cdots & \cdots & \cdots \\ & & & R_{k-1,k-1} & R_{k-1,k} \\ & & & & R_{kk} \end{pmatrix} \tag{15}$$

and

$$S = \begin{pmatrix} S_{11} & S_{12} & \cdots & S_{1,k-1} & S_{1k} \\ & S_{22} & \cdots & S_{2,k-1} & S_{2k} \\ & & \cdots & \cdots & \cdots \\ & & & S_{k-1,k-1} & S_{k-1,k} \\ & & & & S_{kk} \end{pmatrix}. \tag{16}$$

They call matrices (15) and (16) the first condensed form of the pair (A_1, A_2). The following theorem specifies the structure of the first condensed form.

Theorem (Theorem 3 in Alpin *et al.* (2000))

Let $A_1,\ A_2 \in M_n(\mathbb{C})$ be partially commuting matrices. Then for some positive integer k, A_1 and A_2 can be brought by a simultaneous similarity transformation in equation (14) to block upper triangular form (15) and (16). In this form, the diagonal blocks R_{11}, $R_{22}, \cdots, R_{k-1,k-1}$ commute with the corresponding blocks $S_{11}, S_{22}, \cdots, S_{k-1,k-1}$. The last diagonal blocks and either commute or do not partially commute. The integer k and the orders of diagonal blocks are uniquely defined by the matrices A_1 and A_2. Moreover, a similar matrix P that brings A_1 and A_2 to R and S, respectively, can be computed rationally.

With necessary modifications, the first condensed form of the pair (A_1, A_2) can be also defined as a pair of block lower triangular matrices instead. The number k, which is the index of the first condensed form of the pair (A_1, A_2) and the orders of diagonal blocks in equations (15) and (16) are uniquely defined by the matrices A_1 and A_2. It is also known that rank[A_1, A_2] ≥ k-1. As a corollary, it follows that if the rank[A_1, A_2] = 1, their condensed form is a pair of 2×2 block triangular matrices.

We can further refine the condensed form to produce as many zeros as possible in its subdiagonal part; our goal is for commuting $n\times n$ matrices A_1 and A_2 to find a simultaneous similarity transformation

$$F_1 = Q^{-1}A_1Q \text{ and } F_2 = Q^{-1}A_2Q$$

that produces as many zeros as possible in the subdiagonal part of both matrices F_1 and F_2. Alpin *et al.* use an approach based on the generalized Lanczos procedure given in Elsner and Ikramov (1997). We leave the discussion of the procedure in an appendix. Instead, we consider how the procedure can be used to simplify an error correction model. Suppose that X_t is cointegrated with the following ECM:

$$\Delta X_t = -\gamma\alpha' X_{t-1} + A_1^*\Delta X_{t-1} + A_2^*\Delta X_{t-2} + e_t \ .$$

If the cointegrating relationship, z_t, is self-generating so that $\alpha' A_1^* = a_1\alpha'$ and $\alpha' A_2^* = a_2\alpha'$ for constants a_1 and a_2,

$$z_t = (1-\alpha'\gamma + a_1)z_{t-1} + (a_2 - a_1)z_{t-2} - a_2 z_{t-3} + \alpha' e_t \ .$$

Furthermore, A_1^* and A_2^* are simultaneously triangularizable by a nonsingular matrix T such that

$$R_1 = T^{-1}A_1^*T \quad \text{and} \quad R_2 = T^{-1}A_2^*T$$

and the error correction model becomes

$$\Psi_t = R_0 z_{t-1} + R_1\Psi_{t-1} + R_2\Psi_{t-2} + h_t \, ,$$

where $\Psi_t = T^{-1}\Delta X_t$, $R_0 = -T^{-1}g$ and $h_t = T^{-1}e_t$. The subdiagonal parts of the upper triangular matrices R_1 and R_2 can be transformed to have as many zeros as possible following the procedure devised in Alpin *et al.* at least in principle. The same argument will be applied to self-generating common stochastic trends.

5. Testing for self-generation

At first, it seems natural to use the commutator of A_1 and A_2, $[A_1, A_2]$, for testing conditions for self-generation; if $[A_1, A_2]$ is nonsingular, A_1 and A_2 have no common eigenvector. However, the singularity of $[A_1, A_2]$ is not sufficient. Consider the following example from Shemesh (1987),

$$A_1 = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix} \text{ and } A_2 = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 1 & 0 & 0 \end{pmatrix},$$

where A_1 and A_2 have no common eigenvectors. Yet

$$[A_1, A_2] = \begin{pmatrix} 0 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & -1 & 0 \end{pmatrix}$$

is a nilpotent matrix. Or, one may follow the procedure developed in Johansen and Swensen (1999) for testing the restrictions implied by rational expectations in a cointegrated VAR framework. There is, however, a very simple testing procedure that can be easily implemented due to Kohn (1982). Suppose that we are interested in testing if a linear combination of x_t, $d' x_t$, is self-generating. Define a new matrix

$$D = \begin{pmatrix} d_1 & d_2 & \cdots & d_n \\ 0 & 1 & \cdots & 0 \\ \vdots & \vdots & & \\ 0 & 0 & \cdots & 1 \end{pmatrix}$$

such that $k_t = Dx_t$, where $d = (d_1 \ \ d_2 \ \ \cdots \ \ d_n)^{'}$ is an $n \times 1$ vector. If x_t is a stationary ARMA process, so is k_t. Assume for example that k_t has a VAR(p) representation;

$$k_t = \Phi_1 k_{t-1} + \Phi_2 k_{t-2} + \ldots + \Phi_p k_{t-p} + V_t ,$$

where Φ_i is an $n \times n$ coefficient matrix and V_t is white noise. For the newly defined k_t process, select $d = (1,0,\ldots 0)^{'}$, so that the conditions for $d^{'} x_t$ to be self-generating are

$$\Phi_{i,(1,k)} = 0 \text{ , for } i = 1,2,\ldots, p \text{ and } k = 2,\ldots, n ,$$

where $\Phi_{i,(1,k)} = 0$ denotes the first row and kth column element of Φ_i. The test for self-generation is, therefore, equivalent to examining if p lagged values of $(x_2, x_3, \ldots, x_n)$ do not Granger-cause $d^{'} x_t$.

Clark (2000) extends the above procedure to a univariate Δy_t, where y_t is a non-cointegrating linear combination of X_t; regress Δy_t with z_{t-1} and lagged values of all but one of the components of ΔX_t as well as lagged values of Δy_t and test if z_{t-1} and the lagged values of the included ΔX_t are significant. He shows that the testing procedure has an appropriate size and has good power. His testing procedure can be motivated and generalized by the following observation; define the number of common trends $m \equiv n - r$ and a selection matrix $S_r = (0_{r \times (n-r)} \, I_r)$, so that $S_r X_t = (x_{m+1,t} x_{m+2,t}, \ldots, x_{nt}) \equiv X_{-m,t}$. Assume that $y_t = d^{'} X_t$ is a $m \times 1$ vector. X_t can be written as a linear combination of y_t and $X_{-m,t}$, if the inverse of $\begin{pmatrix} d' \\ S_r \end{pmatrix}$ exists. From the error correction model, we get

$$\Delta d' X_t = -d' g a' X_{t-1} + \sum_{i=1}^{p-1} d' A_i^* \Delta X_{t-i} + d' e_t ,$$

so that Δy_t is a function of z_{t-1}, $(p-1)$ lagged values of Δy_t and $(p-1)$ lagged values of $\Delta X_{-m,t}$. The test of self-generation for Δy_t is the significance test of z_{t-1} and $(p-1)$ lagged values of $\Delta X_{-m,t}$. If they are significant, the null that Δy_t is self-generating is rejected. Clearly, we are assuming that $p > 1$.

We expand the procedure to the cointegrating relationship z_t. Define another selection matrix $S_{n-r} = (0_{(n-r)\times r} I_{(n-r)})$, so that $X_{-r,t} \equiv S_{n-r} X_t = (x_{r+1,t} x_{r+2,t}, \dots, x_{nt})$. For a $r \times 1$ vector $z_t = a' X_t$, if the inverse of $\begin{pmatrix} a' \\ S_{n-r} \end{pmatrix}$ exists, X_t can be written as a linear combination of z_t and $X_{-r,t}$. From the error correction model, we get

$$\Delta a' X_t = -a' g a' X_{t-1} + \sum_{i=1}^{p-1} a' A_i^* \Delta X_{t-i} + a' e_t ,$$

so that z_t can be written as a function of p lagged values of itself and $(p-1)$ lagged values of $\Delta X_{-r,t}$. Testing if z_t is self-generating is equivalent to the test of significance of $(p-1)$ lagged values of $\Delta X_{-r,t}$.

Ahn and Reinsel (1990) show that various estimators of “stationary parameters” g and A_i^*, $i = 1, \dots, p-1$ in the error correction model are asymptotically equivalent with limiting multivariate normal distributions that do not depend on the manner in which a is estimated. The common stochastic trend vector $g_{\perp}'$ has also asymptotic normal distribution, as shown in Gonzalo and Granger (1995) or in Ahn and Lee (2000). Therefore, the standard tests for significance will have a c^2 distribution, as also noted in Gonzalo and Granger (1995).

We conduct a limited number of computer simulations to investigate the size and power of the proposed tests. We assume that the cointegrating relationship is known; in fact, there are many economic models that imply cointegration with known cointegrating vectors. For instance, in the examples in section 6 below, all cointegrating relationships are known *a priori*. The lag order is also assumed to be known for simplicity, even though it can be easily relaxed. We test for self-generation at a 5% significance level for a sample of 125 and 250 observations, respectively, with 10,000 simulations. We set $X_0 = 0$ and discard initial 100 observations to lessen the influence of the initial condition.

First, the test of self-generation for common stochastic trends. We use the same data generating process [DGP] as in Clark (2000);

$$X_t = m + A_1 X_{t-1} + A_2 X_{t-2} + U_t,$$

where $X_t = (x_{1t} \quad x_{2t})'$ and U_t denotes bivariate standard normal random errors. We choose the following values for the DGP; $m = \begin{pmatrix} 0.5 \\ 1.5 \end{pmatrix}$ and $a = \begin{pmatrix} 1 \\ -1 \end{pmatrix}$. For parameterization 1-4, we set $A_2 = \begin{pmatrix} 0.6 & 0.2 \\ 0.0 & 0.4 \end{pmatrix}$ and $g = \begin{pmatrix} -0.3 \\ 0.3 \end{pmatrix}$, $\begin{pmatrix} -0.3 \\ 0.5 \end{pmatrix}$, $\begin{pmatrix} -0.1 \\ 0.5 \end{pmatrix}$, or $\begin{pmatrix} 0.2 \\ 0.5 \end{pmatrix}$. For parameterization 5-8, we set $A_2 = \begin{pmatrix} 0.6 & -0.3 \\ 0.2 & 0.4 \end{pmatrix}$ with the same values of g as used in parameterization 1-4. The unweighted sum $(x_1 + x_2)$ is a common stochastic trend when $A_2 = \begin{pmatrix} 0.6 & 0.2 \\ 0.0 & 0.4 \end{pmatrix}$ and $g = \begin{pmatrix} -0.3 \\ 0.3 \end{pmatrix}$. Therefore, the results for the parameterization 1 show the size of the test, while others show its power. The simulation results from likelihood ratio tests are summarized in table 1. Note that the results in the first column are corresponding to those in the table 2 in Clark (2000). The size of the test is appropriate and the power is very good. The power gets higher with more observations.

Second, testing of self-generation for cointegrating relationships. We use a slightly different DGP from above;

$$X_t = m + A_1 X_{t-1} + A_2 X_{t-2} + U_t,$$

where $m = \begin{pmatrix} 0.5 \\ 1.5 \end{pmatrix}$, $a = \begin{pmatrix} 1 \\ -1 \end{pmatrix}$, $g = \begin{pmatrix} -0.3 \\ g_2 \end{pmatrix}$, where $g_2 \in \{0.4,\ 0.3,\ 0.2,\ \ldots,\ \text{-}0.1\}$. Also, $A_2 = \begin{pmatrix} 0.6 & k \\ 0.0 & 0.4 \end{pmatrix}$, with $k \in$ {-0.5, -0.4, …, 0.4, 0.5}, which guarantees that X_t has a unit root. The only exception occurs when k = -0.5 and g_2 = 0.4, in which case a root lies outside unit circle. The cointegrating relationship is self-generating when $k =$-0.2, regardless less of the values of g. We still assume that the cointegrating relationship is known. Table 2 shows some of the simulation results from likelihood ratio tests. We test if the cointegrating relationship is self-generating at a 5% significance level with a sample of 125 and 250 observations, respectively. With k = -0.2, the size of the test is appropriate across different values of g_2. The results indicate that the test has a better power when the value of k is farther away from -0.2. Clearly, the power increases with more observations.

We can summarize the proposed testing procedure as follows. Run cointegration analysis using, for example, the maximum likelihood approach of Johansen (1991) and find cointegrating relationships and common stochastic trends as defined by Gonzalo and Granger (1995). Conduct hypothesis tests on the cointegrating relationships and common stochastic trends if necessary. Only cointegrating relationships and common stochastic trends can be candidates for self-generation. If the variables of interest are neither of them, conclude that they are not self-generating. If error correction models have no short-run dynamics, conclude that both cointegrating relationships and common stochastic trends are self-generating. Otherwise, run the self-generation test by examining the significance of "other variables" in the cointegrating relationships or common stochastic trends equations.

6. Examples

In this section, we present several examples to which the framework of self-generation can be fruitfully applied.

6.1 Spot and forward exchange rates

There is a large number of studies testing whether forward exchange rates or forward premium help forecasting the future movements of spot exchange rates. Clarida and Taylor (1997) call the problem one of the most persistent questions in international finance. Zivot (2000a), among others, contains previous references on the subject. The null hypothesis is

$$E[s_{t+1} \mid s_{t-j}, f_{t-j}, j \geq 0] = E[s_{t+1} \mid s_{t-j}, j \geq 0],$$

where s_t and f_t are the spot and forward exchange rates, respectively. Under the null, the spot rate is self-generating. Zivot (2000a) finds that the following error correction model is adequate for the monthly exchange rates on the pound, yen, and Canadian dollar relative to the US dollar;

$$\Delta f_t = m_f + g_f (f_{t-1} - a_s s_{t-1}) + e_{ft} ,$$

$$\Delta s_t = m_s + g_s (f_{t-1} - a_s s_{t-1}) + e_{st} ,$$

where a_s is a cointegrating coefficient, m_f and m_s are constants, and g_f and g_s are adjustment coefficients. Readers are referred to his paper for more details on the data and on empirical results. With no lagged values of Δf_t and Δs_t in the above error correction model, both cointegrating relationship and common stochastic trend are self-generating. We examine cointegrating relationship first, even though it is not the main hypothesis we want to test. Zivot shows that the exchange rates are cointegrated with $a_s = 1$ after restricting the intercept into the error correction term such that $m_f = -g_f m_c$ and $m_s = -g_s m_c$ for all three currencies. He also finds that $m_c = 0$. Therefore, the cointegrating relationship or forward premium ($f_t - s_t$) is self-generating. Note that the condition $|1+g_f - g_s|$ < 1 is required for stationarity and the forward premiums are highly autocorrelated, especially when g_f and g_s are close to each other.

Now consider the self-generation property for common stochastic trend. For the spot rate to be self-generating, it should be the common stochastic trend of the two exchange rates, which requires that $g_s = 0$. It follows that the changes in spot rate are white noise. However, formal tests strongly reject the hypothesis that $g_s = 0$ for all three currencies; see table 7 in Zivot (2000a). The spot exchange rate is not common stochastic trend of the exchange rates and therefore it cannot be self-generating; the information held in forward rates help forecasting the future movements in spot rates.

This conclusion agrees with that in Clarida and Taylor (1997), who use weekly data series for Germany, Japan and the UK.

6.2 Dividends/stock prices and consumption/income

These relationships are investigated in Cochrane (1994) and Gonzalo and Granger (1995). We will investigate the empirical results presented in Zivot (2000b) instead, mainly because the latter lends an easier interpretation of empirical results in terms of self-generation. The present value model of stock prices together with dividend smoothing model implies that dividends follow a random walk and price/dividend ratio is stationary. Zivot (2000b) finds that a simple error correction model with no short-run dynamics is adequate for dividends/stock prices; see table 2 of Zivot (2000b). The estimated error correction model indicates that the dividend price ratio ($d_t - p_t$) is self-generating, approximately $d_t - p_t \approx$ -.64+.80($d_{t-1} - p_{t-1}$) + h_t, where d_t and p_t are dividend and stock price, respectively, and h_t is white noise. The error correction model shows also that dividend is common stochastic trend and self-generating as well. Furthermore, Δd_t is approximately white noise.

For consumption and income, their difference is found to be stationary and consumption is their common stochastic trend, as the permanent income theory indicates. The empirical results in table 2 in Zivot (2000b) show, however, that neither their cointegrating relationship nor common stochastic trend is self-generating. For instance, the consumption equation is approximately $\Delta c_t = .08\Delta y_{t-1} + e_t$, even though the coefficient of determination, R^2, of the equation is pretty small. c_t and y_t denote consumption and income, respectively.[8]

[8] Zivot (2000b) contains additional results on spot and forward exchange rates, which are

We can add some additional remarks on investigating another implication of the permanent income hypothesis as tested in Campbell (1987). He showed that under the permanent income hypothesis restrictions, the revision from time (*t-1*) to t in the expected value of human wealth,

$$s_t - \Delta y_{lt} - (1+r)s_{t-1},$$

is unpredictable given lagged Δy_{lt} and s_t. Saving, $s_t \equiv y_{lt} + y_{kt} - c_t$, is the difference between (capital and labor) income and consumption[9] and r is an expected real interest rate. The condition implies that $s_t - \Delta y_{lt} - (1+r)s_{t-1}$ is a serial correlation common feature. Given that a serial correlation common feature is a special case of self-generation, we examine the condition for $s_t - \Delta y_{lt} - (1+r)s_{t-1}$ to be self-generating in the framework adapted in Campbell. Even if we make a strong assumption that both s_t and Δy_{lt} are self-generating, since s_t is a cointegrating relationship and y_{lt} is assumed to be a common stochastic trend, their linear combination will not be self-generating in general. In fact, Campbell found that s_t Granger caused Δy_{lt}, as shown in his table IV, and rejected the hypothesis that $s_t - \Delta y_{lt} - (1+r)s_{t-1}$ is unpredictable.

6.3 Real interest rate

There are large empirical studies on the time series behavior of real interest rate, see for example Mishkin (1995). Can we improve the forecasts of real interest rate by considering the information held in past nominal interest rates and inflation rates

---

consistent with those reported in Zivot (2000a) and in the previous subsection of this paper.

[9] Note that we use the same notation as in Campbell (1987) to be consistent with him. The notations in this subsection should not be confused with those in the previous subsection.

than just real interest rate's past history? Let R_t be the one-period nominal interest rate from time t to ($t+1$) and p_{t+1} the realized inflation during that period. Define $(R_t - p_t)$ as the proxy of real interest rate, since p_{t+1} is not observable at the beginning of time t.[10] Using quarterly U.S. data on three month Treasury bill rate and the inflation rate calculated from the CPI, Bidarkota (1998) finds that both R_t and p_t are nonstationary and cointegrated with $R_t - p_t$ as their cointegrating relationship. Results are reported for the following restricted error correction model;

$$\begin{pmatrix} \Delta R_t \\ \Delta p_t \end{pmatrix} = \begin{pmatrix} w_1 \\ w_2 \end{pmatrix} (R_{t-1} - p_{t-1}) - \begin{pmatrix} w_1 & 0 \\ w_2 & 0 \end{pmatrix} \begin{pmatrix} \Delta R_{t-1} \\ \Delta p_{t-1} \end{pmatrix} + \begin{pmatrix} z_{1t} \\ z_{2t} \end{pmatrix} + \begin{pmatrix} q_{11} & q_{12} \\ q_{21} & q_{22} \end{pmatrix} \begin{pmatrix} z_{1,t-1} \\ z_{2,t-1} \end{pmatrix},$$

where z_{it} is white noise, $i = 1, 2$. From the estimation results reported in his table 6, we can infer that the proxy of real interest rate is not self-generating when we acknowledge that w_1 is significantly different from zero and w_2 is insignificant. Therefore, improved forecasts of real interest rate will be possible by considering a bigger information set of nominal interest rate and inflation rate.[11] Bidarkota also finds that a bivariate error correction model produces better forecasts than a univariate unobserved component model of real interest rate. Additionally, we can infer that inflation rate is their common stochastic trend and self-generating.

[10] $(R_t - p_{t+1})$ is usually used as *ex-post* real interest rate, which is also employed in Bidarkota (1998). The following results on self-generation are not changing when this definition is used instead.

[11] The time series behavior of real interest rates is known to change, possibly due to monetary policy shifts or budget deficits. See for example Bekdache (1999).

6.4 Nominal GDP

Consider the forecast of nominal GDP growth from a model with real GDP, price, and M2. Can the nominal GDP growth be efficiently forecast by its own past only without using the information held in other variables? In a working paper version of Clark (2000), the full information set is found to produce improved forecasts of nominal GDP growth; Clark (1995). We can provide a simple argument that the nominal GDP growth cannot be self-generating using the self-generation framework. According to the P-star model of Hallman *et al.* (1991), there is a cointegrating relationship among (log transformed) real GDP, price, and M2, so that the M2 velocity is stationary, which in turn indicates a cointegrating vector of $(1\ 1\ -1)'$ among the three variables. Therefore, the nominal GDP is not a cointegrating relationship and cannot be self-generating. No hypothesis tests are reported in Clark (1995) on the cointegrating relationship, even though the estimated values seem to be close to $(1\ 1\ -1)'$. The nominal GDP can be a candidate for a self-generating process only on an extremely unlikely situation when real GDP and price are common stochastic trends respectively with the same stochastic structure. Empirical results for UK using M4 as presented in Hall and Milne (1994) show also that the velocity is stationary and the nominal GDP is not self-generating in the UK data. We acknowledge that in recent years the relationship among the variables considered in the example is changing due to financial innovations.

6.5 U.S. GNP, consumption, and investment

We investigate the self-generation property among the famous data set used in King, Plosser, Stock and Watson (1991, KPSW); per capita US quarterly GNP, consumption, and investment from 1947:1 ~ 1988:4. With a lag order of 6 as in KPSW, two cointegrating relationships are found. The hypothesis that the so-called great ratios are stationary is not rejected with a p-value of 0.15. Furthermore, the hypothesis that consumption is the only element of their common stochastic trend is not rejected with a p-value of 0.13, along with the restrictions on the cointegrating relationships. The finding that consumption is the only element of common stochastic trend of the three aggregate variables seems to quite robust. For instance, Penm *et al.* (1997) report similar findings, using quite different approach for a different sample period. Therefore, the maintained cointegrating and common stochastic vectors are, respectively,

$$a = \begin{pmatrix} 1 & -1 & 0 \\ 1 & 0 & -1 \end{pmatrix}' \text{ and } g_{\perp} = \begin{pmatrix} 0 & 1 & 0 \end{pmatrix}'.$$

We first test if the common stochastic trend is self-generating; the hypothesis that the lagged two cointegrating relationships and lagged GNP and investment are jointly zero is rejected with a p-value of 0.02. Therefore, consumption is not self-generating. For the hypothesis that the two cointegrating relationships are self-generating, we test if the lagged values of GNP growth are significant in the system of equations of the cointegrating relationships. The hypothesis is rejected with a p-value of 0.00 according to a likelihood ration test. The results are not changing if lagged values of consumption or investment is used instead of GNP. Therefore, neither common stochastic trend nor cointegrating relationships is self-generating among the three aggregates. The results are not changing when a lag order of 8 as in Proietti (1997) or a shorter one of 4 is used in an error correction model.

7. Conclusion

In this paper, a new class of model has been introduced, some properties of if determined, testing procedure considered, and various applications given. A variable is called self-generating if it can be efficiently forecast by its past only. The definition can be easily extended to vectors of variables. In a cointegrated VAR framework, cointegrating relationships and common stochastic trends are only two candidates for self-generation. The conditions for self-generation suggest a simplifying structure in multiple time series. Many interesting economic hypotheses can be easily examined by utilizing the properties of self-generation in a cointegrated VAR framework.

Table 1
Size and power of the self-generation test for a common stochastic trend

Parameterization	Sample size =125	Sample size = 250
1	0.051	0.051
2	0.653	0.922
3	0.979	0.998
4	0.999	0.999
5	0.996	1.000
6	0.999	0.998
7	0.985	0.998
8	1.000	1.000

The number shows the ratio in which the null hypothesis of self-generation is rejected at 5% significance level using likelihood ratio tests. The total number of simulations is 10,000. Parameterization 1 shows the size of the tests.

Table 2
Size and power of the self-generation test for a cointegrating relationship

$k \downarrow g_2 \rightarrow$	*0.4*	*0.3*	*0.1*	*0.0*	*-0.1*
0.5	0.995	0.993	0.999	0.996	0.989
	0.998	0.999	0.998	0.997	0.998
0.3	0.949	0.932	0.938	0.943	0.952
	0.994	0.996	0.992	0.994	0.990
0.1	0.512	0.509	0.537	0.528	0.545
	0.800	0.808	0.812	0.820	0.830
0.0	0.251	0.259	0.263	0.278	0.272
	0.447	0.457	0.468	0.481	0.486
-0.1	0.102	0.109	0.109	0.106	0.108
	0.149	0.152	0.159	0.156	0.162
-0.2	0.049	0.054	0.053	0.054	0.052
	0.053	0.052	0.051	0.050	0.050
-0.3	0.099	0.103	0.094	0.102	0.108
	0.149	0.155	0.150	0.156	0.154
-0.4	0.216	0.257	0.255	0.253	0.266
	0.406	0.447	0.470	0.463	0.458
-0.5	N/A.	0.476	0.487	0.488	0.499
		0.770	0.780	0.786	0.785

The number shows the ratio in which the null hypothesis of self-generation is rejected at 5% significance level using likelihood ratio tests. In each cell, the numbers are from a sample of 125 and 250 observations, respectively. The total number of simulations is 10,000. The result for k = -0.2 shows the size of the test. When k = -0.5 and g_2 = 0.4, a root lies outside unit circle, so no results are reported.

Appendix

In the appendix, we briefly present the procedure to have as many zeros as possible in the condensed form of two commuting matrices A_1 and A_2 as mentioned in section 4 of the text. The following discussion is from Alpin *et al.* (2000). Let $m_1,\dots,m_k$ be the orders of diagonal blocks in the condensed form of (A_1, A_2), which we denote by *R* and *S*, respectively, and consider a nonsingular block diagonal matrix

$$D = D_1 \oplus D_2 \oplus \cdots \oplus D_k, \qquad \text{(a1)}$$

where $D_i \in M_{m_i}(\mathbb{C})$, $i = 1,\dots,k$ and $\oplus$ denotes direct sum of matrices. Then the matrices

$$\tilde{R} = D^{-1}RD \text{ and } \tilde{S} = D^{-1}SD \qquad \text{(a2)}$$

are another condensed form of (A_1, A_2). We explore the possibility to gain more zeros in the subdiagonal part of matrices $\tilde{R}$ and $\tilde{S}$ by using a freedom in choosing transformations (a1) and (a2). The approach to solving this problem is based on the use of the generalized Lanczos procedure as presented in Elsner and Ikramov (1997). Let $u \in \mathbb{C}^n$ be a fixed nonzero vector and consider the vector sequence

$$u, A_1u, A_2u, A_1^2u, A_1A_2u, A_2^2u, A_1^3u, A_1^2A_2u, A_1A_2^2u, A_2^3u, A_1^4u, \dots \qquad \text{(a3)}$$

It is convenient to consider the sequence (a3) as consisting of segments of length 1, 2, 3,…, respectively. The (*k+1*)th segment is called the *k*th layer of (a3). It can be described as the totality of vectors of vectors of the form $u_k = W_k(A_1, A_2)u$, where $W_k(s,t)$ is any words of degree k in two commuting variables s and t; for k = 0, $W_0(A_1, A_2) = I_n$. The subspace

$$L_m(A_1, A_2; u) = span\{W(A_1, A_2)u : \text{degree}(W) \le m\}$$

is called the mth generalized Krylov space and its dimension is denoted by l_m. Call $w_m = l_m - l_{m-1}$ the width of the mth layer and set $w_0 = 1$. Clearly, $w_m \le m+1$. Assume also $l_p = n$. We need another definition. The vectors

$$x_1, x_2, \dots, x_n \tag{a4}$$

are said to be a graded basis of $\square^n$, if for $q = 0,1,\dots, p$, the first l_q vectors in (a4) form a basis of the generalized Krylov subspace L_q.

The positions of zeros in the ith column can be determined by finding the minimal index s of the subspace L_q to which the vector x_i belongs and set

$$N_i = i - 1 + w_s + w_{s+1}. \tag{a5}$$

Then the entries with row indices exceeding N_i are zero. It follows from (a5) that among the subdiagonal entries of the ith column, only M_i entries can be nonzero, where $M_i = w_s + w_{s+1} - 1$. In sum, the procedure shows a way to obtain additional zeros in the diagonal blocks of a condensed form of partially commuting matrices A_1 and A_2.

References

Ahn, Sung K. and Eui Yong Lee, 2000, A note on testing the nested structure in multivariate regression models, Oxford Bulletin of Economics and Statistics, 62, 3, 451-458

Ahn, Sung K. and Gregory C. Reinsel, 1990, Estimation for partially nonstationary multivariate autoregressive models, Journal of the American Statistical Association, 85, 411, 813-823

Alpin, Yu.A., L. Elsner, and Kh.D. Ikramov, 2000, On condensed forms for partially commuting matrices, Linear Algebra and its Applications, 306, 165-182

Bekdache, Basma, 1999, The time-varying behavior of real interest rates: a re-evaluation of the recent evidence, Journal of Applied Econometrics, 14, 171-190

Bidarkota, Prasad, 1998, The comparative forecast performance of univariate and multivariate models: am application to real interest rate forecasting, International Journal of Forecasting, 14, 457-468

Bodo, Giorgio, Roberto Golinelli, and Giuseppe Parigi, 2000, Forecasting industrial production in the Euro area, Empirical Economics, 25, 541-561

Campbell, John Y., 1987, Does saving anticipate declining labor income? An alternative test of the permanent income hypothesis, Econometrica, 55, 6, 1249-1273

Cappuccio, Nunzio and Diego Lubian, 1996, Triangular representation and error correction mechanism in cointegrated systems, Oxford Bulletin of Economics and Statistics, 58, 409-415

Clark, Todd E., 1995, Forecasting an aggregate of cointegrated disaggregates, Research Division Working Paper 95-13, Federal Reserve Bank of Kansas City

Clark, Todd E., 2000, Forecasting an aggregate of cointegrated disaggregates, Journal of Forecasting, 19, 1-21

Clarida, Richard H, and Mark P. Taylor, 1997, The term structure of forward exchange premiums and the forecastability of spot exchange rates: correcting the errors, The Review of Economics and Statistics, 79, 353-361

Clements, Michael P. and David F. Hendry, 1995, Forecasting in cointegrated systems, Journal of Applied Econometrics, 10, 127-146

Cochrane, John, 1994, Permanent and transitory components of GNP and stock prices, Quarterly Journal of Economics, 109, 241-265

Cohen-Tannoudji, Claude, Bernard Diu, and Franck Laloë, 1977, Quantum Mechanics, translated by Susan Reid Hemley, Nicole Ostrowsky and Dan Ostrowsky, New York: John Wiley & Sons

Elsner, L. and Kh.D. Ikramov, 1997, On a condensed form for normal matrices under finite sequences of elementary unitary similarities, Linear Algebra and its Applications, 254, 79-98

Engle, Robert F. and Clive W.J. Granger, 1987, Cointegration and error correction: Representation, estimation, and testing, Econometrica, 55, 251-276

Engle, Robert F. and Clive W.J. Granger (eds.), 1991, Long-run economic relationships: Readings in cointegration (Advanced Texts in Econometrics), Oxford, U.K.: Oxford University Press

Engle, Robert F. and Sharon Kozicki, 1993, Testing for common features, with comments and reply, Journal of Business & Economic Statistics, 11, 4, 369-395

Ericsson, Neil R., 1993, Comment, Journal of Business & Economic Statistics, 11, 4, 380-383

Fair, Ray C. and Robert J. Shiller, 1990, Comparing information in forecasts from econometric models, American Economic Review, 80, 375-389

Gonzalo, J. and Clive W.J. Granger, 1995, Estimation of common long-memory component in cointegrated systems, Journal of Business & Economic Statistics, 13, 27-35

Granger, Clive W.J., 1969, Investigating causal relationships by econometric models and cross-spectral methods, Econometrica, 37, 424-438

Granger, Clive W.J. and Jin-lung Lin, 1995, Causality in the long-run, Econometric Theory, 11, 3, 530-536

Hall, Stephen G. and Alistair Milne, 1994, The relevance of P-star analysis to UK monetary policy, Economic Journal, 104, 597-604

Hallman, Jeffrey J., Richard R. Porter, and David H. Small, 1991, Is the price level tied to the M2 monetary aggregate in the long run? American Economic Review, 81, 841-858

Hoffman, Dennis L. and Robert H. Rasche, 1996, Assessing forecasting performance in a cointegrated system, Journal of Applied Econometrics, 11, 495-517

Johansen, Søren, 1991, Estimation and hypothesis testing of cointegration in Gaussian vector autoregressive models, Econometrica, 59, 1551-1580

Johansen, Søren, and Anders Rygh Swensen, 1999, Testing exact rational expectations in cointegrated vector autoregressive models, Journal of Econometrics, 93, 73-91

Kohn, R., 1982, When is an aggregate a time series efficiently forecast by its past? Journal of Econometrics, 18, 337-350

Lütkepohl, Helmut, 1984, Linear transformations of vector ARMA processes, Journal of Econometrics, 26, 283-293

Lütkepohl, Helmut, 1991, Introduction to Multiple Time Series Analysis, Springer-Verlag: Berlin

Mishkin, Frederic, 1995, Nonstationarity of regressors and tests on real-interest-rate behavior, Journal of Business & Economic Statistics, 13, 47-51

Penm, Jammie H., Jack H.W. Penm, and R.D. Terrell, 1997, The selection of zero-non-zero patterned cointegrating vectors in error-correction modelling, Econometric Review, 16, 281-304

Phillips, Peter C.B., 1991, The optimal inference in cointegrated systems, Econometrica, 59, 283-306

Proietti, Tommaso, 1997, Short-run dynamics in cointegrated systems, Oxford Bulletin of Economics and Statistics, 59, 405-422

Shemesh, Dan, 1980, Common eigenvectors of two matrices, Linear Algebra and its Applications, 62, 11-18

Shemesh, Dan, 1987, Common eigenvectors and commutative subspaces, Linear Algebra and its Applications, 104, 227-231

Sims, Christopher, 1980, Macroeconomics and reality, Econometrica, 48, 1-48

Zivot, Eric, 2000a, Cointegration and forward and spot exchange rate regressions, Journal of International Money and Finance, 19, 785-812

Zivot, Eric, 2000b, The power of single equation tests for cointegration when the cointegrating vector is prespecified, Econometric Theory, 16, 407-439