December 1995

Separation in Cointegrated Systems, Long Memory Components and Common Stochastic Trends.

Clive W. J. GRANGER

Department of Economics, University of California, San Diego, La Jolla, CA-92093, USA.

Niels HALDRUP

Department of Economics and Centre for Non-linear Modelling in Economics, Aarhus University, DK-8000 Aarhus C, Denmark.

ABSTRACT

The notion of separation in cointegrated systems helps identifying possible sub-system structures that may reduce the complexity of larger systems by yielding a more parsimonous representation of the time series. In this paper we demonstrate that although the subsystem cointegration analysis in such systems can be conducted in case of both completely and partially separated systems, the dual appraoch, i.e. calculation of the common stochastic trends, may turn out to yield properties of the trends that differ depending upon the type of separation under consideration. In particular, we demonstrate how persistent-transitory decompositions and long- and short-memory factorisations of a multivariate time series will be affected when considering different types of separation. Generalisations to non-linear error correction models are also discussed.

JEL-codes: C32, C40

Key-words: Separation, sub-system cointegration, persistent-transitory decomposition, non-linear error correction.

The first author acknowledges support from NSF grant SBR 93-08295. The research was undertaken while the second author was visiting the UCSD during fall, 1995.

1. Introduction and some motivation.

It is a frequent empirical finding in macroeconomics, that several cointegration relations may exist amongst economic variables but in the particular way that the single relations appear to have no variables in common. It is also sometimes found in such systems that the error correction terms or other stationary variables from one set of variables may have important explanatory power for variables in another set. For example Konishi *et al*. (1993) considered three types of variables of US data: real, financial and interest rate variables. They found that cointegration existed between variables in each subset but not across the variables such that the different sectors did not share a common stochastic trend. On the other hand, it was also found that the error correction terms of the interest rate relation and the sector of financial aggregates had predictive power with respect to the real variables of the system. As argued by Konishi *et al*. (1993) the situation sketched above may extend the usual ’partial equilibrium’ cointegration set-up to a more ’general equilibrium’ setting although in a limited sense.

The notion of separation initially developed by Konishi and Granger (1992) and Konishi (1993) provides a useful way of describing formally the above possibility: Consider two groups of I(1)-variables, X_{1t} and X_{2t} of dimension p_1 and p_2, respectively. X_1 and X_2 are assumed to have no variables in common and in each sub-system there is cointegration with the cointegration ranks being $r_1<p_1$ and $r_2<p_2$. Hence it follows that the dimensions of the associated common stochastic trends of each system are p_1-r_1 and p_2-r_2. Denote the two sets of I(1) stochastic trends W_{1t} and W_{2t}. It follows from Stock and Watson (1988) that each sub-system can be given the representation

$$\begin{aligned} X_{1t} &= G_1 W_{1t} + \tilde{X}_{1t} \\ X_{2t} &= G_2 W_{2t} + \tilde{X}_{2t} \end{aligned} \tag{1.1}$$

where G_i, i=1,2, are $p_i \times (p_i - r_i)$ matrices and the $\tilde{X}_{it}$ components are stationary I(0) relations. Separation of the two sub-systems means that the components W_{1t} and W_{2t} are not cointegrated so that there is no long-run relationship between the X_{1t} and X_{2t} variables. As a consequence the stacked time series $X_t=(X_{1t}', X_{2t}')'$ will be of dimension $p=p_1+p_2$ and have cointegration rank $r=r_1+r_2$. The full system stochastic trend component will have the dimension p-r.

Despite this separation of variables it may well occur that a relationship exists between X_{1t} and X_{2t} in the short run. Essentially there are two ways this can happen: ΔX_{2t} (ΔX_{1t}) may appear in the transitory I(0) component $\widetilde{X}_{1t}$ ($\widetilde{X}_{2t}$) and/or error correction terms from one system may enter the second. Both possibilities will be of interest in this paper and we will refer to these situations as *partial separation of types A and B*, respectively. If the error correction terms from each subsystem and the long-run impact of first differences only have explanatory power in their own system, we denote this *complete separation*.

What will be of concern in this paper is to consider a decomposition of the vector time series X_t in persistent-transitory (P-T) components for separated cointegration models. Identification of these components are generally non-unique since any I(1) process can be contaminated with an I(0) process and still have the I(1) property. Various additional requirements have been suggested in the literature to identify the components and more recently Gonzalo and Granger (1995), using a factor model approach, suggest that the temporary component be defined in terms of the error correction relations such that it will have no explanatory power on the series in the long run. Moreover, the single factors can be measured in terms of the observed variables X_t. One of the findings of the present paper (which in some respects actually goes beyond the particular decomposition suggested by Gonzalo and Granger), is that if the decomposition $X_t=P(X_t)+T(X_t)$ is considered where $P(X_t)$ and $T(X_t)$ are the persistent (long-memory) and transitory (short-memory) components, respectively, the persistent component P_{1t} associated with the X_1-system, for instance, can be expressed as $P_{1t}=P_1(X_{1t},X_{2t})$ in case of both complete and partial separation. Hence, in order to extract *observable* persistent components in a separated system, it is not generally sufficient to consider each sub-system in separation, since all system variables may be needed to define the components. For example, if one (wrongly) treats a partially separated system as a complete one, we will demonstrate that both the persistent *and* the transitory components of the other system may turn out to affect the persistent component of the sub-system under consideration. Only when the entire system is completely separated will it be sufficient to look at the submodels to find the long-memory components and the common stochastic trends. This result is interesting because it suggests that cointegration analysis, and especially common stochastic trends analysis and P-T decompositions, may suffer from only looking at small partial models. In the interpretation of common stochastic trends the idea of ’general equilibrium’ cointegration is therefore relevant since persistent and transitory components may interact across systems.

The plan of the paper is the following. In section two we provide a formal definition of the various separation concepts and we briefly review some of the literature concerned with decomposition of a series into persistent and transitory components. The following section focuses on the decomposition in the context of separated cointegrated models. We demonstrate that if a partially separated system is treated as if it is complete, both the long- and short-memory components of the neglected system may potentially affect the persistent component of the system being analyzed. However, the problem can be avoided by considering the full system but with the implication that the (true) long and short memory factors may depend upon *all* the model variables of each sub-system. In section 4 possible extension to non-linear error correction models are considered and we demonstrate that fairly strong restrictions need to be imposed on the functional forms across systems in order to ensure stability. In the final section we conclude.

2. Definition of the Concepts.

We shall here define formally the different concepts that will be used in the sequel.

2.1 Complete and partial separation in cointegrated systems.

The definition of separation provided below extends Konishi and Granger (1992) and Konishi (1993).

DEFINITION. *Consider the p-dimensional cointegrated vector time series* $X_t=(X_{1t}', X_{2t}')'$ *where* X_{1t} *and* X_{2t} *are of dimension* p_1 *and* p_2 ($p=p_1+p_2$) *and have no variables in common. Then the associated error correction model reads*

$$\Delta X_t = \underset{p\times r}{\gamma}\,\underset{r\times p}{\alpha'} X_{t-1} + \underset{p\times p}{\Gamma(L)}\Delta X_{t-1} + \varepsilon_t \qquad (2.1)$$

where r is the cointegration rank. If the matrix of cointegration parameters can be factored as

$$\alpha' = \begin{pmatrix} \alpha'_{11} & 0 \\ 0 & \alpha'_{22} \end{pmatrix} \tag{2.2}$$

where α_{ii}' *is* $p_i \times r_i$*, i=1,2, the system is said to have separate cointegration. Conformably with this partitioning, consider also the matrices*

$$\gamma = \begin{pmatrix} \gamma_{11} & \gamma_{12} \\ \gamma_{21} & \gamma_{22} \end{pmatrix} \quad \textit{and} \quad \Gamma(L) = \begin{pmatrix} \Gamma_{11}(L) & \Gamma_{12}(L) \\ \Gamma_{21}(L) & \Gamma_{22}(L) \end{pmatrix}. \tag{2.3}$$

Given separate cointegration; if $\gamma_{12}=\gamma_{21}=\Gamma_{12}(1)=\Gamma_{21}(1)=0$ *the system is said to be completely separated. The case with* γ_{12}, $\gamma_{21}\neq 0$ *refers to partial separation of Type A, and the case where* $\Gamma_{12}(1)$, $\Gamma_{21}\neq 0$ *will refer to partial separation of Type B.*

Notice that in the definition of complete separation there is no feedback from the error correction terms across each sub-system. Neither is there feedback, in the long run, from the first differenced variables across the systems. However, we do not preclude the possibility that the first differences of the variables in one system may have explanatory power in the other system in the short run.

2.2. P-T decomposition of a vector time series.

It is frequently of interest to decompose a time series into components that may have different characteristics, for instance a Persistent-Transitory (P-T) decomposition may be relevant, see e.g. Beveridge and Nelson (1981) and Quah (1992). For a vector time series similar decompositions may be considered, see e.g. Stock and Watson (1988), Kasa (1992) Mellander *et al.* (1992), Gonzalo and Granger (1995), and Proietti (1995). However, since identification of such factors is generally non-unique, additional identifying requirements are needed. Gonzalo and Granger have suggested that the persistent I(1) factors should 1) be observable, i.e. such that the persistent components be expressed in terms of the original variables X_t and 2), the shocks to the transitory part should have no impact on the persistent components in the long run, see e.g. Hosoya (1991) and Granger and Lin (1995) for a definition of causality at different frequencies. Essentially, this is why the two types of factors for this particular decomposition may be given the economic interpretation of long-memory and short-memory components. The second condition stated above says that if we let $X_t=P_t+T_t$ be the factorization of X_t into a persistent and a transitory component, then the components can be given the VAR representation

$$\begin{pmatrix} H_{11}(L) & H_{12}(L) \\ H_{21}(L) & H_{22}(L) \end{pmatrix}\begin{pmatrix} \Delta P_t \\ T_t \end{pmatrix}=\begin{pmatrix} \varepsilon_{Pt} \\ \varepsilon_{Tt} \end{pmatrix} \tag{2.4}$$

such that T_t does not cause ΔP_t in the long run if $H_{12}(1)=0$.

Observability of the factors can be achieved by considering the expression

$$X_t=P(X_t)+T(X_t) \tag{2.5}$$

where $P(X_t)=A_1 f_t$ and $T(X_t)=A_2 z_t$ with $f_t=\gamma_\perp' X_t$ and $z_t=\alpha' X_t$ and where $A_1=\alpha_\perp(\gamma_\perp'\alpha_\perp)^{-1}$ and $A_2=\gamma(\alpha'\gamma)^{-1}$. The matrices $\alpha_\perp$ and $\gamma_\perp$ are orthogonal complements of α and γ, i.e. such that $\gamma_\perp'\gamma=0$ and $\alpha_\perp'\alpha=0$. Throughout the symbol "$\perp$" will indicate the orthogonal complement of the associated matrix. Notice that the orthogonal matrices in the present case are both $p\times(p-r)$ and that the factorisation of the vector process exists when $\alpha'\gamma$ is invertible. In fact, this will always be the case when the cointegration rank is r. The persistent (or long-memory) component is given by $P(X_t)$ which can be seen to be expressed in terms of the $(p-r)$ common stochastic trends f_t, and similarly the short-memory components can be expressed by the r error correction terms in a particular way. Observe that $P(X_t)$ and $T(X_t)$ do not necessarily constitute an orthogonal factorisation; this will only happen in special situations. The Gonzalo-Granger decomposition has similarities with other decompositions in the literature. For instance the f_t term is identical to the common stochastic trends of Stock and Watson (1988) which is a multivariate generalisation of the Beveridge-Nelson decomposition of a univariate time series. In a recent paper Proietti (1995) compares the various representations in a common set-up and he demonstrates that the Gonzalo-Granger decomposition can be obtained from the Beveridge-Nelson decomposition by adding a particular distributed lag polynomial of the first differences of the series to the long-memory component. The reason why this can be done is, of course, that any stationary component can be added to the stochastic trend (or I(1)) component without altering the over all I(1) characteristics.

Each term in the Gonzalo-Granger decomposition can easily be calculated in terms of the X_t series given that estimates of α and γ are available. The factorisation into the two different components is also seen to be in correspondance with the distinction between partial separation of Type A and complete separation: It is the error correction terms which contribute to the short memory component of the Gonzalo-Granger decomposition, and it is also the impact of the error correction terms across sub-systems which distinguishes between partial separation of Type A and complete separation.

3. P-T decompositions in separated cointegrating systems.

In this section we focus our attention on different types of separated models to see how the long and short memory components will depend upon the particular type of separation.

3.1. Erroneously treating a partially separated system as completely separated.

In order to interpret the outcome of cointegration analysis it is frequently an advantage to consider systems of low dimension. Assume that the econometrician correctly considers a separated cointegrated system, but wrongly assumes that separation is complete rather than partial. The difference is, naturally, that the feedback from other cointegrating relations through the error correction terms and the first differenced variables from the other system are ignored in the analysis. For simplicity, assume that the model considered is (2.1)-(2.3) with γ_{21}=0 and $\Gamma_{21}(L)$=0 such that we have a recursive system. This assumption is with no loss of generality but it will make the subsequent arguments clearer. The econometrician is concerned with the error correction model associated with the X_1 system, that is

$$\Delta X_{1t} = \gamma_{11}\alpha_{11}' X_{1,t-1} + \Gamma_{11}(L)\Delta X_{1,t-1} + u_{1t} \ . \qquad (3.1)$$

Naturally, the error term captures whatever may have been left out from the analysis[1], so in this case we have that $u_{1t}=\gamma_{12}\alpha_{22}'X_{2,t-1}+\Gamma_{12}(L)\Delta X_{2,t-1}+\varepsilon_{1t}$.

The X_2 system reads

[1] Naturally, it may generally occur that the errors u_t are correlated with the regressors in (3.1). Hence, with respect to estimation this should be reflected by choice of an appropriate estimation procedure. The topic of the present paper is on representation and hence we will not consider estimation issues here.

$$\Delta X_{2t}=\gamma_{22}\alpha_{22}^{\prime}X_{2,t-1}+\Gamma_{22}(L)\Delta X_{2,t-1}+\varepsilon_{2t}\ . \tag{3.2}$$

By treating (3.1) as an isolated system the common stochastic trends are given by premultiplication of the error correction model (3.1) by the $p_1\times(p_1-r_1)$ orthogonal complement of γ_{11}, i.e. $\gamma_{11}^{\perp\prime}$ where $\gamma_{11}^{\perp\prime}\gamma_{11}=0$. This yields

$$f_{1t}=\gamma_{11}^{\perp\prime}X_{1t}=\gamma_{11}^{\perp\prime}\sum_{j=0}^{\infty}u_{1,t-j}\ , \tag{3.3}$$

and similarly we have for system 2, that the stochastic trends are given by

$$f_{2t}=\gamma_{22}^{\perp\prime}X_{2t}=\gamma_{22}^{\perp\prime}\sum_{j=0}^{\infty}\varepsilon_{2,t-j}\ . \tag{3.4}$$

Since f_{1t} is seen to depend upon variables from the second system through the errors u_{1t}, it will be useful to write the trends in terms of the persistent and transitory components of each system. In accordance with the Gonzalo-Granger decomposition we can define (with an obvious notation)

$$\begin{aligned} X_{1t}&=P_{1t}+T_{1t}=A_{11}\gamma_{11}^{\perp\prime}X_{1t}+A_{21}\alpha_{11}^{\prime}X_{1t} \\ &=A_{11}f_{1t}+A_{21}Z_{1t} \end{aligned} \tag{3.5}$$

and

$$\begin{aligned} X_{2t}&=P_{2t}+T_{2t}=A_{12}\gamma_{22}^{\perp\prime}X_{2t}+A_{22}\alpha_{22}^{\prime}X_{2t} \\ &=A_{12}f_{2t}+A_{22}Z_{2t} \end{aligned} \tag{3.6}$$

where $A_{11}=\alpha_{11}^{\perp}(\gamma_{11}^{\perp\prime}\alpha_{11}^{\perp})^{-1}$, $A_{21}=\gamma_{11}(\alpha_{11}^{\prime}\gamma_{11})^{-1}$, $A_{12}=\alpha_{22}^{\perp}(\gamma_{22}^{\perp\prime}\alpha_{22}^{\perp})^{-1}$ and $A_{22}=\gamma_{22}(\alpha_{22}^{\prime}\gamma_{22})^{-1}$. From (3.3) it now follows that

$$\begin{aligned} \Delta P_{1t}&=A_{11}\gamma_{11}^{\perp\prime}\Delta X_{1t} \\ &=A_{11}\gamma_{11}^{\perp\prime}\Gamma_{11}(L)(\Delta P_{1,t-1}+\Delta T_{1,t-1})+A_{11}\gamma_{11}^{\perp\prime}u_{1t}\ . \end{aligned} \tag{3.7}$$

As seen the transitory component T_{1t} will have no impact on ΔP_{1t} in the long run as it is required in the identification of the single components suggested by Gonzalo and Granger. However, consider the remainder term $A_{11}\gamma_{11}^{\perp\prime}u_{1t}$. This can be written as

$$A_{11}\gamma_{11}^{\perp\prime}u_{1t}=A_{11}\gamma_{11}^{\perp\prime}\{\gamma_{12}Z_{2,t-1}+\Gamma_{12}(L)(\Delta P_{2,t-1}+\Delta T_{2,t-1})+\sum_{j=0}^{\infty}\varepsilon_{1,t-j}\}\ . \tag{3.8}$$

This result has the following interesting implications: If partial separation is of Type B the long-memory component of system 1 will depend upon the long-memory component of system 2. Only when $\Gamma_{12}(1)=0$ is this possibility excluded. Furthermore, since $\gamma_{12}Z_{2,t-1}=\gamma_{12}\alpha_{22}'T_{2,t-1}$ the short memory component of system 2 will have an impact on the long-memory component of system 1, unless $\gamma_{12}=0$. In other words, if there is partial separation of Type A, the transitory component from the neglected system will affect the persistent component of the system that is examined. The interaction between the two subsystems is absent, however, if separation is complete.

3.1. Partial separation and P-T decomposition.

The proper way to proceed in order to avoid the caveat emphasized in the previous section, is to treat the two sub-systems jointly. Again we assume for simplicity that $\gamma_{21}=0$ and let $\Gamma_{21}(L)=0$. Define the matrix

$$\gamma_{\perp}'=\begin{pmatrix}\gamma_{11}^{\perp\,\prime} & \gamma_{12}^{*\,\prime}\\ 0 & \gamma_{22}^{\perp\,\prime}\end{pmatrix} \tag{3.9}$$

such that $\gamma_{\perp}'\gamma=0$, and hence $\gamma_{11}^{\perp\prime}\gamma_{12}+\gamma_{12}^{*\prime}\gamma_{22}=0$ and with $\gamma_{11}^{\perp\prime}$ and $\gamma_{22}^{\perp\prime}$ defined as before. The common stochastic trends of the full system can now be written as

$$\begin{aligned}f_{1t}&=\gamma_{11}^{\perp\,\prime}X_{1t}+\gamma_{12}^{*\,\prime}X_{2t}=\gamma_{11}^{\perp\,\prime}\Gamma_{11}(L)X_{1,t-1}+\gamma_{11}^{\perp\,\prime}\Gamma_{12}(L)X_{2,t-1}+\gamma_{11}^{\perp\,\prime}\sum_{j=0}^{\infty}\varepsilon_{1t}+\gamma_{12}^{\perp\,\prime}\sum_{j=0}^{\infty}\varepsilon_{2t}\\ f_{2t}&=\gamma_{22}^{\perp\,\prime}X_{2t}=\gamma_{22}^{\perp\,\prime}\Gamma_{22}(L)X_{2,t-1}+\gamma_{22}^{\perp\,\prime}\sum_{j=0}^{\infty}\varepsilon_{2t}\ .\end{aligned} \tag{3.10}$$

In this case, by construction, the common stochastic trends and the Gonzalo-Granger decomposition effectively separates the adjustment of error correction errors from the long-memory component as intended. However, an interesting thing to observe is that generally the decomposition for each sub-system will take the form

$$\begin{aligned}X_{1t}&=P_1(X_{1t},X_{2t})+T_1(X_{1t},X_{2t})\\ X_{2t}&=P_2(X_{2t})+T_2(X_{2t}).\end{aligned} \tag{3.11}$$

The variables of the full system are thus needed in both the long and the short memory components of the X_1-system. Notice that P_{1t} and P_{2t} are *not* cointegrated. Since P_{1t} is I(1) plus I(0) in a particular way, it can also be seen that P_{1t}, which essentially is determined by f_{1t} in (3.10), will have X_{1t} as the only factor if $\Gamma_{12}(1)=0$. In general, however, X_{2t} will contribute to both the I(1) and the I(0) components.

An alternative way to proceed yielding further insights is to study the interaction of the persistent and transitory components across sub-systems. The long-memory components read

$$\Delta P_t = A_1 \gamma_\perp' \Delta X_t. \tag{3.12}$$

By straightforward matrix operations, using rules of partitioned inverse, it can be shown that

$$\begin{aligned} \Delta P_{1t} &= A_{11}\gamma_{11}^{\perp\prime}\Delta X_{1t} + A_{11}\gamma_{12}^{*\prime}A_{22}\alpha_{22}^{\prime}\Delta X_{2t} \\ \Delta P_{2t} &= A_{12}\gamma_{22}^{\perp\prime}\Delta X_{2t}\ . \end{aligned} \tag{3.13}$$

By using the error-correction model (2.1)-(2.3) for ΔX_{1t} and ΔX_{2t} in the present set-up and using the fact that $\gamma_{11}^{\perp\prime}\gamma_{12}+\gamma_{12}^{*\prime}\gamma_{22}=0$, it can be easily proved that the single components are related in the following way:

$$\begin{aligned} \Delta P_{1t} = {} & A_{11}\gamma_{11}^{\perp\prime}\Gamma_{11}(L)\,(\Delta P_{1,t-1}+\Delta T_{1,t-1}) + \\ & \{A_{11}\gamma_{11}^{\perp\prime}\Gamma_{12}(L)+A_{11}\gamma_{12}^{*\prime}A_{22}\alpha_{22}^{\prime}\Gamma_{22}(L)\}(\Delta P_{2,t-1}+\Delta T_{2,t-1}) + \\ & A_{11}\gamma_{11}^{\perp\prime}\varepsilon_{1t}+A_{11}\gamma_{12}^{*\prime}A_{22}\alpha_{22}^{\prime}\varepsilon_{2t} \end{aligned} \tag{3.14}$$

$$\Delta P_{2t} = A_{12}\gamma_{22}^{\perp\prime}\Gamma_{22}(L)(\Delta P_{2,t-1}+\Delta T_{2,t-1}) + A_{12}\gamma_{22}^{\perp\prime}\varepsilon_{2t}\ .$$

This way of writing the autoregressive representation of the components demonstrates that in the long run T_{1t} and T_{2t} will not have any explanatory power with respect to the long-memory components in either system. This is fully consistent with their definition of course. However, the long-memory component of the X_2-system will cause the corresponding component of the X_1-system, but without being cointegrated. Trivial exceptions where ΔP_{2t} does not cause ΔP_{1t} in the long run exist, of course, for instance when $\Gamma_{12}(1)=\Gamma_{22}(1)=0$.

The analysis of the past two sections demonstrates the importance of considering whether error correction terms and other short run dynamics from other systems may have an impact on the system of interest when cointegration is separate. Although it is not going to affect the cointegration properties of the data, it clearly becomes of importance in extracting and interpreting the common stochastic trends and the long and short memory components of the multivariate system. In this sense it is of interest to consider the notion of cointegration in a general (rather than a partial) equilibrium framework. After all, it can be seen that examinations including common stochastic trends analysis should be done with care due to the dependence of such trends with respect to the information set.

4. Extensions to non-linear error correction models.

Cointegrated models with non-linear error correction mechanisms have recently attracted much attention in the literature, compare e.g. Granger and Swanson (1995), and Granger and Teräsvirta (1993) and the references therein. The types of non-linearity entering such systems need to be restricted, however, in order to ensure stability of the model. In this section we demonstrate how the restrictions required in one system may or may not restrict the other system when cointegration is separate.

Non-linear error correction models may take many different forms. Consider, for example, a single system with the non-linear error correction mechanism entering as follows:

$$\Delta X_t = \gamma\theta(\beta' Z_{t-1}) + \Gamma(L)\Delta X_t + \varepsilon_t \ , \qquad (4.1)$$

where $Z_t=\alpha' X_t$. As usual X_t is a p-vector time series and we let $\theta(\beta' Z_{t-1})$ be a (p-r) vector of non-linear functions of the lagged error correction terms; notice that since β is $r\times 1$, $\beta' Z_t$ is assumed to be a scalar variable. Here we want to emphasize the non-linear property and assume for simplicity that $\Gamma(L)=0$.

Multiplying (4.1) by α' we obtain

$$\Delta Z_t = \alpha'\gamma\theta(\beta' Z_{t-1}) + \alpha'\varepsilon_t . \qquad (4.2)$$

which is a non-linear VAR(1) process. In defining

$$Z_t = h(\beta' Z_{t-1}) + \eta_t \qquad (4.3)$$

where

$$h(Z)=Z+\alpha'\gamma\theta(\beta'Z) \tag{4.4}$$

the admissable class of functions ensuring stability should satisfy the necessary and sufficient stability conditions, see Tweedie (1975), Lasota and Mackey (1989), and Granger and Teräsvirta (1993),

$$\|h(Z)\|\leq a\|Z\| \text{ for } \|Z\|\geq c \text{ and } |a|<1. \tag{4.5}$$

and

$$\|h(Z)\| \text{ is finite for all finite } Z. \tag{4.6}$$

$\|.\|$ can be any norm, not necessarily the Euclidean norm[2]. It follows that for the case of one dimension, the functions satisfying stability must be dominated by a linear function with slope less than one, for instance, if $\theta(Z)$ is one dimensional the function could be logistic in Z or $\log(Z)$. The stability condition above applies to the vector Z. If this is stable, so are the single components, but it is not generally possible to provide conditions on the stability of each element in $\theta(Z)$.

The restrictions above can be weakened in some cases meaning that only a subset of the functions in $\theta(Z)$ need to be restricted, i.e. the functions for which the adjustments lie in the space spanned by $\alpha_\perp$ we need no restrictions to be imposed to ensure stability. Assume for simplicity that this space is empty such that each element of $\theta(Z)$ should be considered in derivation of the stability conditions.

Despite non-linearity in the adjustment and error correction terms, the common stochastic trends f_t, in the Stock-Watson and Gonzalo-Granger sense, turn out to behave linearly since $f_t=\gamma_\perp' X_t=\gamma_\perp'\Delta^{-1}\varepsilon_t$ in the present situation. In other words, the common stochastic trends will have *no* non-linear feature.

Assume now that cointegration is separate, using the terminology of section 2, and that error correction is non-linear in the following way,

[2] Strictly speaking, the above conditions also require that the non linear VAR model is of first order.

$$\begin{pmatrix}\Delta X_{1t}\\ \Delta X_{2t}\end{pmatrix} = \begin{pmatrix}\gamma_{11} & \gamma_{12}\\ 0 & \gamma_{22}\end{pmatrix}\begin{pmatrix}\theta_1(\beta_1'\alpha_{11}'X_{1t-1})\\ \theta_2(\beta_2'\alpha_{22}'X_{2t-1})\end{pmatrix} + \begin{pmatrix}\varepsilon_t\\ \varepsilon_t\end{pmatrix} \tag{4.7}$$

using an obvious notation. In case of complete separation, which in the present set-up means that γ_{12}=0, the common stochastic trends (with no non-linear feature) are easily calculated for each sub-system. This case is rather trivial. So is the situation where $X_t=(X_{1t}',X_{2t}')'$ is treated jointly and separation is partial ($\gamma_{12}\neq 0$). In this case $\gamma_\perp'$ effectively kills both the non-linear error correction terms.

Consider instead the case where system 1 is treated as completely separated although it is only partially separated. In this case the common stochastic trends of the X_1-system read

$$\Delta f_{1t}=\gamma_{11}^{\perp\,\prime}\Delta X_{1t}=\gamma_{11}^{\perp\,\prime}\gamma_{12}\theta_2(\beta_2'\alpha_{22}'X_{2,t-1})+\gamma_{11}^{\perp\,\prime}\varepsilon_{1t} \ . \tag{4.8}$$

Hence, although the common stochastic trends of the X_2-system are linear, the correspon-ding trends of the X_1-system will generally have a non-linear feature.

What restrictions are needed on $\theta_1(.)$ and $\theta_2(.)$ in the partially separated system to ensure stability ? We have that

$$\begin{aligned}\Delta Z_{1t}&=\alpha_{11}'\gamma_{11}\theta_1(\beta_1'Z_{1,t-1})+\alpha_{11}'\gamma_{12}\theta_2(\beta_2'Z_{2,t-1})+\alpha_{11}'\varepsilon_{1t}\\ \Delta Z_{2t}&=\alpha_{22}'\gamma_{22}\theta_2(\beta_2'Z_{2,t-1}) \ .\end{aligned} \tag{4.9}$$

So, the stability requirements in this case are not affected: As long as the stability conditions of system 2 are satisfied, the stability conditions that are necessary for system 1 will be unaffected by system 2. Observe, however, that if we introduce $\gamma_{21}\neq 0$ such that $\alpha_{22}'\gamma_{21}\theta_1(Z_{1,t-1})$ will appear in the expression for ΔZ_{2t} in (4.9), the stability conditions for the single systems cannot be calculated in isolation. The systems have to be treated jointly in this case, i.e. by letting $Z_t=(Z_{1t}',\ Z_{2t}')'$ and considering the system (4.2). The joint stability requirements of $\theta_1(.)$ and $\theta_2(.)$ are given by (4.5) and (4.6).

It is clearly a restriction implied by the particular non-linear model considered above, that the functional forms of the error correction terms associated with the X_2-system, and entering in the X_1-system, must be the same as those arising in the X_2-system with respect to the the same error correction terms. Many other model constructions could be considered. For instance, the model

$$\begin{pmatrix}\Delta X_{1t}\\ \Delta X_{2t}\end{pmatrix}=\begin{pmatrix}\gamma_{11}\theta_{11}(\beta'_{11}Z_{1,t-1})+\gamma_{12}\theta_{12}(\beta'_{12}Z_{2,t-1})\\ \gamma_{21}\theta_{21}(\beta'_{21}Z_{1,t-1})+\gamma_{22}\theta_{22}(\beta'_{22}Z_{2,t-1})\end{pmatrix}+\begin{pmatrix}\varepsilon_{1t}\\ \varepsilon_{2t}\end{pmatrix} \tag{4.10}$$

could be analyzed. This class of model is probably more relevant in practice, but its increased flexibility adds to the complexity of deriving common stochastic trends and P-T decompositions. No results are presently available for this type of non-linear error correction models, but is is certainly a class of dynamical models that will be of interest for future research.

5. Conclusion.

Separation in cointegrated systems is a useful notion which helps to reduce the complexity of large systems and eases their interpretation. Within a cointegrated VAR set-up, c.f. Johansen (1988, 1991), both partially and completely separated models can be easily tested by considering particular hypotheses on the cointegration vectors and the adjustment coefficients, see Konishi and Granger (1992).

The possibility of partial separation, i.e. where error correction terms and stationary variables from other systems may enter the model, is an important possibility to consider; not only because it may improve the model for forecasting purposes , but also, as we have demonstrated in this paper, because the implied short run dynamics actually may add to our understanding of the stochastic trends driving the system as well as the complex dynamical interaction that may exist across systems. It is therefore our suggestion for empirical practice that the applied econometrician is aware of such important links rather than just focusing on the long properties of the data in terms of cointegration.

Generalisations to non-linear models, and in particular, non-linear error correction models, is still in its infancy, but potentially a rich class of dynamical systems can be analyzed within this set-up. However, much more research needs to be done in order to obtain results that are useful for the practitioner.

6. References.

Beveridge, S. and C.R. Nelson, 1981, "A New Approach to the Decomposition of Economic Time Series into Permanent and Transitory Components with Particular Attention to the Measurement of the Business Cycle", *Journal of Monetary Economics*, 7, 151-174.

Gonzalo, J. and C.W.J. Granger, 1995, "Estimation of Common Long-Memory Components in Cointegrated Systems", *Journal of Business and Economic Statistics*, 13, 1, 27-35.

Granger, C. W. J., and T. Lee, 1989, "Investigation of Production, Sales and Inventory Relations using Multicointegration and Non-symmetric Error Correction Models", *Journal of Applied Econometrics*, 4, S145-S159.

Granger, C. W. J., and T. Lee, 1990, "Multicointegration", in Advances in Econometrics, Vol. 8, ed. by G. F. Rhodes and T. B. Fomby, Greenwich, CT: JAI Press, 71-84.

Granger, C. W. J., and J.Lin, 1995, "Causality in the Long-Run", *Econometric Theory*, 11, 530-536.

Granger, C. W. J., and N. Swanson, 1995, "Further Developments in the Study of Cointegrated Variables", forthcoming in *Oxford Bulletin of Economics and Statitics.*

Granger, C. W. J., and T. Teräsvirta, 1993, Modelling Nonlinear Economic Relationships, Oxford University Press.

Hosoya, Y. 1991, "The Decomposition and Measurement of the Interdependence between Second-order Stationary Processes", *Probability Theory and Related Fields*, 88, 429-444.

Johansen, S. 1988, "Statistical Analysis of Cointegration Vectors", *Journal of Economic Dynamics and Control*, 12, 231-254.

Johansen, S. 1991, "Estimation and Hypothesis Testing of Cointegration Vectors in Gaussian Vector Autoregressive Models." *Econometrica*, 59, 1551-1580.

Kasa, K., 1992, "Common Stochastic Trends in International Stock Markets", *Journal of Monetary Economics*, 29, 95-124.

Konishi, T., 1993, "Separation and Long-Run Non-causality in a Cointegrated System", PhD Dissertation, UCSD.

Konishi, T., and C.W.J. Granger, 1992, "Separation in Cointegrated Systems", manuscript, Department of Economics, UCSD.

Konishi, T., V.A. Ramey, and C.W.J. Granger, 1994, "Stochastic Trends and Short-Run Relationships between Financial Variables and Real Activity", manuscript, Department of Economics, UCSD.

Lasota, A., and MacKey, M. C. (1989), "Stochastic Pertubation of Dynamical Systems: the Weak Convergence of Measures", *Journal of Mathematical Analysis and Applications*, 138, 232-248.

Mellander, E., A. Vredin and A. Warne, 1992, "Stochastic Trends and Economic Fluctuations in a Small Open Economy", *Journal of Applied Econometrics*, 7, 369-394.

Proietti, T., 1995, "Short Run Dynamics in Cointegrated Systems", forthcoming in *Oxford Bulletin of Economics and Statistics.*

Quah, D., 1992, "The Relative Importance of Permanent and Transitory Components: Identification and Some Theoreticl Bounds", *Econometrica*, 60, 107-118.

Stock, J.H., and M.W. Watson, 1988, "Testing for Common Trends", *Journal of the American Statistical Association*, 83, 1097-1107.

Tweedie, R. L. (1975), "Sufficient Conditions for Ergodicity of Spectra", in U. Grenander (ed.), *Probability and Statistics*, (New York, Wiley), 300-330.