

# Linear Aggregation with Common Trends and Cycles

Massimiliano Marcellino <sup>\*†</sup>

Istituto di Economia Politica, Università Bocconi,  
IGIER and European University Institute

## Abstract

We propose a general framework to study whether and how common trends and common cycles are still present when the original variables are linearly aggregated or only a subset of them is analysed. This is particularly important because of the adoption in empirical analysis of aggregated data on a limited number of variables.

*Key words:* Aggregation, Marginalization, Common Trends, Common Cycles.

*JEL Classification:* C32, C43, C5.

---

<sup>\*</sup>I wish to thank two anonymous referees, Clive Granger, Søren Johansen, Marco Lippi, Grayham Mizon and Pravin Trivedi for helpful suggestions and comments on an earlier version of this paper. The usual disclaimers apply.

<sup>†</sup>Please address correspondence to IGIER - Bocconi, Via Salasco 3/5, 20136 Milan, Italy. E-mail: [massimiliano.marcellino@uni-bocconi.it](mailto:massimiliano.marcellino@uni-bocconi.it)

## 1. Introduction

The possibility that the univariate behaviour of economic variables can be represented as a random walk plus a stationary component has received increasing attention and support since the early '80s, starting with the seminal paper by Nelson and Plosser (1982). Moreover, the existence of “equilibrium” relationships among the variables, as implied by economic theory and by the observation of similar patterns of growth, has led to the conclusion that only a limited number of random walks drives the joint long run evolution of the variables, i.e., they are cointegrated, see, e.g., Engle and Granger (1987), Stock and Watson (1988), and Johansen (1988). Similarly, the short run behaviour of the variables could be determined by a limited number of shocks, i.e., there could be serial correlation common features among the variables, cofeatures henceforth, see Engle and Kozicki (1993), Vahid and Engle (1993a, 1993b).

In the theoretical papers which introduced these concepts, the level of temporal and linear aggregation of the variables under analysis was not relevant, but it becomes very important in the applications. Actually, temporally and linearly aggregated data on a small number of variables are usually jointly studied. Hence, it is of particular interest to determine whether or not the presence of cointegration and cofeatures at the joint disaggregate level implies that they are still present in partial aggregate models, and vice versa.

We will assume that the frequency of data generation and observation coincide, in order to focus on the effects of linear aggregation.<sup>1</sup> An important particular case, aggregation over sectors in our notation, has been studied by Lippi (1988, 1989) and further considered by Gonzalo (1993) and Forni and Lippi (1997). In this paper, we extend former results and propose some new ones, consider more general aggregation schemes, and also deal with the effects of aggregation on cofeatures.

The required disaggregate probabilistic framework is introduced in Section 2. The effects of aggregation are studied in Section 3. Section 4 discusses particular cases of interest. Section 5 presents some statistical considerations. Section 6 concludes.

---

<sup>1</sup>Issues of temporal aggregation are considered, e.g., in Marcellino (1998, 1999).

## 2. The probabilistic framework

Let us consider the  $nm$  dimensional process

$$x = \{(x_{11t}, x_{12t}, \dots, x_{1nt}, x_{21t}, \dots, x_{2nt}, \dots, x_{m1t}, \dots, x_{mnt})'\}_{t=0}^{\infty},$$

whose evolution is governed by the stochastic difference equations

$$G(L)x_t = \varepsilon_{xt}, \quad (2.1)$$

where  $L$  is the lag operator,  $G(L) = I - G_1L - G_2L^2 - \dots - G_gL^g$ ,  $\varepsilon_{xt}$  is a vector white noise with  $V(\varepsilon_{xt}) = \Upsilon_x$ ,  $\Upsilon_x$  is non singular. Exogenous variables and correlated error terms could be considered, but they complicate the analysis without adding substantial information on the topic of interest.

We will interpret  $n$  as the number of agents in the economy and  $m$  as the number of sectors, even if a switching or other interpretations are possible. (2.1) can be considered either as a generic linear process, or as an economic theory based model. The former possibility seems preferable because it is more general. It will be also convenient to write the vector autoregressive (VAR) process in (2.1) in Error Correction (EC) form as:

$$\Delta x_t = \Pi x_{t-1} + \Gamma_1 \Delta x_{t-1} + \dots + \Gamma_{g-1} \Delta x_{t-g+1} + \varepsilon_t, \quad (2.2)$$

where  $\Pi = -G(1)$  and  $\Gamma_i = -\sum_{j=i+1}^g G_j$ ,  $i = 1, \dots, g-1$ .

Former studies usually impose restrictions on (2.1) or (2.2), see, e.g., Hendry *et al.* (1984), Nickell (1985), Granger (1987), Lippi (1988), and the review in Granger (1990). For example, it is sometimes maintained that the micro variables only react to the aggregated ones, e.g.:

$$\begin{pmatrix} x_{11t} \\ x_{12t} \\ \dots \\ x_{mnt} \end{pmatrix} = \begin{pmatrix} g_{11} \sum_{i=1}^n \sum_{j=1}^m x_{ijt-1} \\ g_{12} \sum_{i=1}^n \sum_{j=1}^m x_{ijt-1} \\ \dots \\ g_{mn} \sum_{i=1}^n \sum_{j=1}^m x_{ijt-1} \end{pmatrix} + \varepsilon_{xt} \Leftrightarrow G(L) = \begin{pmatrix} 1 - g_{11}L & g_{11}L & \dots & g_{11}L \\ g_{12}L & 1 - g_{12}L & \dots & g_{12}L \\ \dots & \dots & \dots & \dots \\ g_{mn}L & g_{mn}L & \dots & 1 - g_{mn}L \end{pmatrix}$$

Or that the coefficients do not depend on the characteristics of the agents, e.g.:

$$\begin{pmatrix} x_{11t} \\ x_{12t} \\ \dots \\ x_{mnt} \end{pmatrix} = \underset{1 \times 1}{g(L)} \begin{pmatrix} x_{11t-1} \\ x_{12t-1} \\ \dots \\ x_{mnt-1} \end{pmatrix} + \varepsilon_{xt} \Leftrightarrow G(L) = \begin{pmatrix} 1 - g(L) & 0 & \dots & 0 \\ 0 & 1 - g(L) & \dots & 0 \\ \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & 1 - g(L) \end{pmatrix},$$

where  $g(L) = g_1L + \dots + g_gL^g$ .

Or that there are no cross agents effects, e.g.,

$$\begin{pmatrix} x_{11t} \\ x_{12t} \\ \dots \\ x_{mnt} \end{pmatrix} = \begin{pmatrix} g_{11}x_{11t-1} + g_{21}x_{21t-1} + \dots g_{m1}x_{m1t-1} \\ g_{12}x_{11t-1} + g_{22}x_{21t-1} + \dots g_{m2}x_{m1t-1} \\ \dots \\ g_{1n}x_{11t-1} + g_{2n}x_{21t-1} + \dots g_{mn}x_{m1t-1} \end{pmatrix} + \varepsilon_{xt} \Leftrightarrow$$

$$\Leftrightarrow G(L) = \begin{pmatrix} (1 - g_{11}L, 0, \dots, 0) & (g_{11}L, 0, \dots, 0) & \dots & (g_{11}L, 0, \dots, 0) \\ (0, g_{12}L, 0, \dots, 0) & (0, 1 - g_{12}L, 0, \dots, 0) & \dots & (0, g_{12}L, 0, \dots, 0) \\ \dots & \dots & \dots & \dots \\ (0, \dots, 0, g_{mn}L) & (0, \dots, 0, g_{mn}L) & \dots & (0, \dots, 0, 1 - g_{mn}L) \end{pmatrix}.$$

We make milder hypothesis on  $G(L)$ . In particular, we only assume that:

*A1.* The roots of  $|G(z)| = 0$  lie in 1 or outside the unit circle.

*A2.* The rank of the  $nm \times nm$  matrix  $\Pi = -G(1) = \alpha\beta'$  equals  $p \leq nm$ , where  $\alpha$  and  $\beta$  are  $nm \times p$  matrices of rank  $p$

*A3.* The rank of the  $r \times r$  matrix  $\alpha'_\perp \Psi(1)\beta_\perp$  equals  $r = nm - p$ , where  $\Psi(1)$  is the derivative of  $\Pi(z)$  with respect to  $z$  evaluated at  $z = 1$ ,  $\Pi(z) = (1 - z)I - \sum_{j=1}^{g-1} \Gamma_j(1 - z)z^j - \Pi z$ , and  $\alpha_\perp$  and  $\beta_\perp$  are orthogonal to  $\alpha$  and  $\beta$ , and such that  $span(\alpha, \alpha_\perp) = span(\beta, \beta_\perp) = R^{nm}$ .

The assumption *A1* rules out the possibility of explosive processes.

The assumption *A2* is related to the presence of cointegration among the variables (e.g., Engle and Granger (1987), Johansen (1988,1995)). Actually, the rank of  $\Pi$  coincides with the number of cointegration vectors, i.e., of stationary linear combinations of the variables. They can be grouped in the columns of the matrix  $\beta$ . When  $p = 0$  there is no cointegration among the variables, when  $p = nm$  the variables are stationary.

The assumption *A3* implies that the variables are at most integrated of order one, i.e., they become stationary after first differencing. Actually, as Engle and Granger (1987), Stock and Watson (1988) and Johansen (1991) showed, if  $x$  is generated by (2.2) and *A1*, *A2*, *A3* hold, then an alternative representation is:

$$\Delta x_t = T(L)\varepsilon_t, \quad \frac{T(L)}{(1-L)}G(L) = I, \quad (2.3)$$

or

$$x_t = T(1) \sum_{i=1}^t \varepsilon_i + \bar{T}(L) \varepsilon_t + B_0, \quad (2.4)$$

where  $\bar{T}(L) = (T(L) - T(1))/(1 - L)$ ,  $T(1) = \beta_{\perp}(\alpha'_{\perp} \Psi(1) \beta_{\perp})^{-1} \alpha'_{\perp}$ ,  $B_0 = \sum_{i=1}^g \bar{T}_{t-i} (G_i x_0 + \dots + G_{i+g-1} x_{1-g})$ .<sup>2</sup>  $T(1)$  can be decomposed into two  $nm \times r$  matrices of full column rank:

$$T(1) = OP', \quad O = \beta_{\perp}(\alpha'_{\perp} \Psi(1) \beta_{\perp})^{-1}, \quad P = \alpha_{\perp}. \quad (2.5)$$

Therefore, only  $r$  linearly independent common trends,  $f_t = \alpha'_{\perp} \sum_{i=1}^t \varepsilon_i$ , drive the evolution of  $x$ .

Besides the common trends, the dynamics of the variables could be driven by only a limited number of shocks with temporary effects, e.g., temporary tax reductions or changes in monetary policy. Vahid and Engle (1993a) show that this requires the existence of linear combinations of the variables such that their first differences are innovations.<sup>3</sup> If we assume that there are  $q$  such so called cofeature combinations, and group their coefficients into the  $nm \times q$  matrix of cofeature vectors  $\gamma$ , then

$$\gamma' \Delta x_t = \gamma' \varepsilon_t. \quad (2.6)$$

Vahid and Engle (1993a) show that the following conditions are equivalent and sufficient for the existence of cofeature vectors:

$$\gamma' \bar{T}_i = 0, i \geq 0; \quad (2.7)$$

$$\gamma' T_i = 0, i > 0; \quad (2.8)$$

$$\gamma' \Gamma_i = 0, i = 1, \dots, g-1, \quad \text{and} \quad \gamma' \Pi = 0; \quad (2.9)$$

$$\gamma'(I - G_1) = 0, \quad \text{and} \quad \gamma' G_i = 0, i \geq 2. \quad (2.10)$$

The number of cointegration and cofeature vectors are linked by the relationship  $q \leq nm - p$ , and the two types of linear combinations of  $x$  are linearly independent. From (2.7) it follows that

$$\gamma' x_t = \gamma' O f_t. \quad (2.11)$$

---

<sup>2</sup>The assumption  $A\beta$  also guarantees that  $x_t$  can be given initial distributions such that  $\beta' B_0 = 0$ , so that  $\Delta x_t - E(\Delta x_t)$  and  $\beta' x_t - E(\beta' x_t)$  are stationary (Johansen (1995, p. 49)).

<sup>3</sup>Recall that an innovation process is a white noise process, but the reverse is not necessarily true, see e.g. Hendry (1995, p. 63) for an example.

Moreover, Vahid and Engle (1993a) prove that  $\bar{T}_i$  can be decomposed for all  $i$  into two  $nm \times nm - q$  matrices  $F$  and  $\bar{T}_i^*$  such that

$$\bar{T}(L)\varepsilon_t = F \bar{T}^{*'}(L)\varepsilon_t = Fg_t,$$

and they call common cycles (*CC*) the  $s = nm - q$  elements of  $g_t$ .

In summary,  $x$  admit the common trends - common cycles representation

$$x_t = Of_t + Fg_t, \quad (2.12)$$

and we are now going to exploit this representation to determine that for the aggregated (and) or marginalized variables.

### 3. Aggregation

In order to examine the effects of aggregation and marginalization on the presence of common trends and cycles among the resulting variables, we introduce a  $j \times nm$  aggregator matrix of full row rank,  $A$ , as, e.g., in Lütkepohl (1984, 1987), with

$$A = \begin{pmatrix} A_{11} & A_{12} & \dots & A_{1m} \\ A_{21} & A_{22} & \dots & A_{2m} \\ \dots & & & \\ A_{j1} & A_{j2} & \dots & A_{jm} \end{pmatrix},$$

where  $A_{ik}$  is a  $1 \times n$  vector. Application of  $A$  to  $x$  leads to the desired  $j$  dimensional aggregated or marginalized process

$$y = Ax.$$

For example, we are often interested in studying the behaviour of sectors of the economy, i.e., of the  $m$  dimensional process which is obtained by aggregating over the  $n$  agents in each sector. Hence,  $A$  has  $m$  rows and

$$A_A = \begin{pmatrix} e & 0 & \dots & 0 \\ 0 & e & \dots & 0 \\ \dots & & & \\ 0 & 0 & \dots & e \end{pmatrix},$$

where all the elements of the  $1 \times n$  vector  $e$  are equal to one. For aggregation over sectors, i.e. when we are interested in studying the aggregate behaviour over sectors of the  $n$  agents, under some homogeneity conditions such as the equality of the unit of measurement across sectors,  $A$  is the  $n \times nm$  matrix

$$A_S = \begin{pmatrix} I \\ I, \dots, I \end{pmatrix}_{n \times n}.$$

Instead, if the focus is on the behaviour of a limited set of agents or sectors of the economy, the process of interest,  $y$ , can be obtained by applying to  $x$  the matrices

$$A_{M1} = \begin{pmatrix} S_{n_1} & 0 & \dots & 0 \\ 0 & S_{n_1} & \dots & 0 \\ \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & S_{n_1} \end{pmatrix}_{n_1 m \times nm} \quad \text{or} \quad A_{M2} = \begin{pmatrix} I & 0 & \dots & 0 \\ 0 & I & \dots & 0 \\ \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & I \end{pmatrix}_{nm_1 \times nm},$$

where  $S_{n_1}$  in  $A_{M1}$  is a selection matrix with ones in correspondence of the  $n_1$  agents of interest for each sector, while the  $m_1$  identity matrices in  $A_{M2}$  identify the sectors of interest.

The VAR representation for  $y$  can be obtained following the technique developed by Lütkepohl (1984). In general, there will also be a moving average (MA) component, but analytical determination of its coefficients is quite complex. It should also be noticed that the aggregated error terms are not simply a linear combination of their disaggregated counterpart.<sup>4</sup> Yet, to study the effects of aggregation on common trends and cycles, a simpler representation can be exploited.

From (2.12) it follows that  $y$  satisfies:

$$y_t = AOf_t + AFg_t. \tag{3.1}$$

Then we have,

**Proposition 1.** A necessary and sufficient condition for the existence of cointegration among the elements of  $y$  is that the null space spanned by  $O'A'$  is not empty. Any of the following conditions are sufficient for this:

---

<sup>4</sup>Formally, they are the projections of the variables on the orthogonal complement of the closure of the space spanned by  $\{y_{t-1}, y_{t-2}, \dots\}$ . See Ansley *et al.* (1977) for a formal demonstration.

*P11:*  $r < j$ ;

*P12:*  $r \geq j$  but at least  $r - j + 1$  columns of  $AO$  are linear combinations of the other ones;

*P13:* at least one row of the matrix  $\begin{pmatrix} \beta' \\ A \end{pmatrix}$  is a linear combination of the other ones. ■

**Proposition 2.** A necessary and sufficient condition for the existence of cofeatures among the elements of  $y$  is that the null space spanned by  $F'A'$  is not empty. Any of the following conditions are sufficient for this:

*P21:*  $s < j$ ;

*P22:*  $s \geq j$  but at least  $s - j + 1$  columns of  $AF$  are linear combinations of the other ones;

*P23:* at least one row of the matrix  $\begin{pmatrix} \gamma' \\ A \end{pmatrix}$  is a linear combination of the other ones. ■

**Proof.** The necessary and sufficient conditions are obvious. The sufficient conditions in *P11-P12* and *P21-P22* ensure the existence of two matrices,  $\delta$  and  $\theta$ , such that  $\delta'AO = 0$  and  $\theta'AF = 0$ . The columns of  $\delta$  and  $\theta$  are the cointegration and cofeature vectors for  $y$ . Their exact number coincides with the dimension of the null space spanned by  $O'A'$  and  $F'A'$ . For *P13*, let us consider, for example, the case

$$a = \begin{matrix} H & \beta' \\ h \times nm & h \times nm - r & nm - r \times nm \end{matrix},$$

where  $a$  consists of  $h$  stacked rows of  $A$ . Then, it is immediate that  $ax_t$  is stationary. If instead

$$\beta'_k = \begin{matrix} K & A \\ k \times nm & k \times j & j \times nm \end{matrix},$$

where  $\beta'_k$  is made up of  $k$  stacked rows of  $\beta'$ , then  $Ky_t$  is stationary. A similar reasoning holds for cofeatures. ■

These conditions are also rather intuitive. *P11* and *P21* require the presence of fewer common trends and cycles than aggregated variables. When there are more, *P12* and *P22* require them to be groupable in fewer than  $j$  trends and cycles. *P13* and *P23* require at least one cointegration or cofeature vector to be a linear combination of the rows of  $A$ , or vice versa.

The formal similarity of Propositions 1 and 2 is due to the similar existence requirement of reduced rank matrices for the presence of cointegration and cofeatures among



the variables. Notice that the two propositions, and the three different conditions in each of them, are not mutually exclusive but can coexist.

Two extreme cases of interest are when the disaggregate variables are stationary or integrated but not cointegrated, i.e., when  $T(1)$  in (2.4) has zero or full rank. These properties are invariant to aggregation, because in the former case it is  $r(AT(1)) = 0$ , while in the latter  $r(AT(1)) = r(A) = j$ .<sup>5</sup>

A further implication of Propositions 1 and 2 is that if the aggregate variables share common trends or cycles, this must be true also for the disaggregate variables. Actually, if this were not true,  $AO$  or  $AF$  would have full row rank and the null space spanned by  $O'A'$  or  $F'A'$  would be empty.<sup>6</sup>

## 4. Examples

We now provide examples where the conditions in Propositions 1 and 2 are satisfied in the presence of aggregation over agents or sectors, or marginalization. Actually, for the usual finding of cointegration and cofeatures among aggregated variables to be reliable, it seems important to analyse whether the conditions in Propositions 1 and 2 are implied by reasonable economic assumptions. We will also relate ours to former results in the literature, which were obtained for particular choices of  $G(L)$  and  $A$ .

### 4.1. Aggregation over agents

Due to the structure of the aggregator matrix  $A_A$ , the coefficients of  $k^{th}$  original common trend and cycle in the equation of the  $i^{th}$  sector are equal to the sum of the

---

<sup>5</sup>An alternative approach to linear aggregation consists in assigning a continuous density function to characterize the distribution of the parameters across the agents, see, e.g., Granger (1980), Trivedi (1985). This can lead to rather different results, e.g., Granger shows that when a Beta distribution is adopted, stationary micro variables can be turned into fractionally integrated macro variables.

<sup>6</sup>If we define  $C = AO$ , it is  $rank(C) \leq \min\{rank(A), rank(O)\}$ , which is equal to  $rank(A)$ , i.e. to  $j$ , when  $O$  has full rank. But it is also  $rank(A) \leq \min\{rank(C), rank(O^{-1})\} = rank(C)$ , so that  $AO = C$  would have full row rank,  $j$ . Hence, there could not be cointegration among the aggregated variables.

coefficients of these trend and cycle for all the agents in that sector. Formally,

$$A_A O = \begin{pmatrix} \sum o_{11j} & \sum o_{12j} & \dots & \sum o_{1rj} \\ \sum o_{21j} & \sum o_{22j} & \dots & \sum o_{2rj} \\ \dots & & & \\ \sum o_{m1j} & \sum o_{m2j} & \dots & \sum o_{mrj} \end{pmatrix}, \quad A_A F = \begin{pmatrix} \sum f_{11j} & \sum f_{12j} & \dots & \sum f_{1sj} \\ \sum f_{21j} & \sum f_{22j} & \dots & \sum f_{2sj} \\ \dots & & & \\ \sum f_{m1j} & \sum f_{m2j} & \dots & \sum f_{msj} \end{pmatrix},$$

where the sums are for  $j = 1, \dots, n$ , and  $o_{ikj}$  and  $f_{ikj}$  are the coefficients of the  $k^{th}$  common trend and cycle in the equation of the  $j^{th}$  agent in the  $i^{th}$  sector.

We now discuss some cases where Proposition 1 holds, and they can be simply modified to provide examples where the conditions in Proposition 2 are satisfied.

To start with, if a limited number of shocks with permanent effects drives the long run evolution of the economy, so that  $r < m$ , then the aggregated variables are also cointegrated (P11 is satisfied).

As an alternative, we can imagine that there are more agents than sectors, that each agent is affected by a different common trend, but the effect of the trend on the agents in the same sector is the same. For example, if the agents are consumers and the common trends are interpreted as income, then we are requiring the marginal propensity to consume to be equal in each sector for all the agents. Formally, it is  $n = r > m$ ,

$$O = \begin{pmatrix} c_1 I \\ 1 \times 1 \ n \times n \\ c_2 I \\ \dots \\ c_m I \end{pmatrix} \quad \text{and} \quad A_A O = \begin{pmatrix} c_1 & c_1 & \dots & c_1 \\ c_2 & c_2 & \dots & c_2 \\ \dots & & & \\ c_m & c_m & \dots & c_m \end{pmatrix}.$$

This implies that the  $n$  original common trends  $f_1, f_2, \dots, f_n$  can be grouped into  $\bar{f} = f_1 + f_2 + \dots + f_n$ , and the  $m$  aggregated variables are driven by one common trend:

$$y_t = \begin{pmatrix} c_1 \\ c_2 \\ \dots \\ c_m \end{pmatrix} \bar{f}_t + \text{stationary terms.}$$

This is an illustration of condition P12, i.e., of the possibility that the columns of  $AO$  become collinear.

A well known and obvious condition for aggregation to preserve cointegration and, more generally, all the characteristics of the micro equations, is that the coefficients of the disaggregate equations are the same for all agents. Actually, equality across agents of the cointegration vectors is sufficient to preserve cointegration. For example, if

$$\beta' = \begin{pmatrix} (\beta_1, 0, \dots, 0) & (\beta_2, 0, \dots, 0) & \dots & (\beta_m, 0, \dots, 0) \\ \text{1} \times \text{n} & & & \\ (0, \beta_1, \dots, 0) & (0, \beta_2, \dots, 0) & \dots & (0, \beta_m, \dots, 0) \\ \dots & & & \\ (0, 0, \dots, \beta_1) & (0, 0, \dots, \beta_2) & \dots & (0, 0, \dots, \beta_m) \end{pmatrix},$$

then the aggregated cointegration vector is  $\bar{\beta}' = (\beta_1, \beta_2, \dots, \beta_m)$ . In this case, e.g., the first row of  $\beta'$  can be written as a linear combination of the remaining ones and of those of  $A_A$ , so that P13 is satisfied. Notice that equality of the loadings of the common trends across agents is sufficient but not necessary for the cointegration vectors to be equal, as simple examples would show.

Finally, cointegration can be preserved and even increased also with extreme heterogeneity of the agents. As a limit example, let us assume that only the first stochastic trend affects the agents in the first sector, i.e.,  $o_{12j} = o_{13j} = \dots = o_{1rj} = 0$ , that its related coefficients,  $o_{11j}$ , are all different, and that there is a cointegration vector equal to  $(e, 0, \dots, 0)$ . The last hypothesis implies that  $\sum_{j=1}^n o_{11j} = 0$ , i.e., even if the behaviour of each agent is driven by the trend, on average its effect is null. It follows that the first element of the vector  $y$  is stationary. Actually, the first row of  $A$  coincides with the cointegration vector, and P13 is again satisfied. This type of restrictions can have a theoretical justification in a general equilibrium model, and such a possibility requires further examination.

## 4.2. Aggregation over sectors

In the case of aggregation over sectors the coefficients of the  $k^{th}$  original common trend and cycle in the aggregate equation of the  $j^{th}$  agent are equal to the sum over the sectors of the coefficients of these trend and cycle in the equations for the  $j^{th}$  agent.

Formally,

$$A_S O = \begin{pmatrix} \sum o_{i11} & \sum o_{i21} & \dots & \sum o_{ir1} \\ \sum o_{i12} & \sum o_{i22} & \dots & \sum o_{ir2} \\ \dots & & & \\ \sum o_{i1n} & \sum o_{i2n} & \dots & \sum o_{irn} \end{pmatrix}, \quad A_S F = \begin{pmatrix} \sum f_{i11} & \sum f_{i21} & \dots & \sum f_{ir1} \\ \sum f_{i12} & \sum f_{i22} & \dots & \sum f_{ir2} \\ \dots & & & \\ \sum f_{i1n} & \sum f_{i2n} & \dots & \sum f_{irn} \end{pmatrix},$$

where the sums are for  $i = 1, \dots, m$ .

Thus, e.g., if there are fewer common trends than agents, the aggregated variables will be cointegrated. Instead, notice that equality of the cointegration vectors among agents is no longer sufficient for this property, while their equality across sectors is. For example,<sup>7</sup> if

$$\Delta x_t = \begin{pmatrix} \alpha_1 & 0 \\ 0 & \alpha_2 \end{pmatrix} \begin{pmatrix} \beta & 0 \\ 0 & \beta \end{pmatrix}' x_{t-1} + \varepsilon_t,$$

then

$$x_t = \begin{pmatrix} \beta_{\perp}(\alpha'_{1\perp}\beta_{\perp})^{-1}\alpha'_{1\perp} & 0 \\ 0 & \beta_{\perp}(\alpha'_{2\perp}\beta_{\perp})^{-1}\alpha'_{2\perp} \end{pmatrix} \sum_{i=1}^t \varepsilon_i + \bar{T}(L)\varepsilon_t + B_0,$$

where

$$\bar{T}_0 = I - T(1), \quad \bar{T}_1 = T_1 - \bar{T}_0, \quad \bar{T}_j = T_j + \bar{T}_{j-1}, \quad j \geq 2$$

and

$$T_0 = I, \quad T_1 = \begin{pmatrix} -\alpha_1\beta & 0 \\ 0 & -\alpha_2\beta \end{pmatrix}, \quad T_i = T_{i-1} \begin{pmatrix} 1 + \alpha_1\beta & 0 \\ 0 & 1 + \alpha_2\beta \end{pmatrix}^{i-1}.$$

After aggregation with  $A_S = \begin{pmatrix} I & : & I \\ n \times n & & n \times n \end{pmatrix}$ , we get

$$y_t = (\beta_{\perp}(\alpha'_{1\perp}\beta_{\perp})^{-1}\alpha'_{1\perp} + \beta_{\perp}(\alpha'_{2\perp}\beta_{\perp})^{-1}\alpha'_{2\perp}) \sum_{i=1}^t \varepsilon_i + A_S \bar{T}(L)\varepsilon_t + A_S B_0,$$

from where it follows that  $\beta' y_t$  is stationary.

Notice that even for this very simple example it is not possible to find a finite order VAR representation for the aggregated process. Actually, we have

$$\Delta y_t = \alpha_1 \beta' y_{t-1} + u_t,$$

---

<sup>7</sup>I owe this example to a referee.

where  $u_t$  has the same autocovariance function as  $(\alpha_2 - \alpha_1)\beta' x_{2t-1} + \varepsilon_{1t} + \varepsilon_{2t}$ .  $u_t$  is instead a white noise process for  $\alpha_1 = \alpha_2$ .

The effects of this type of aggregation on common trends, but not on common cycles, were analysed by Gonzalo (1993) in a different framework. He considered each sector separately, and for each of them adopted the representation in equation (2.4), i.e.,

$$x_{it} = T_i(1) \sum_{k=1}^t u_{ik} + \bar{T}_i(L)u_{it}, \quad i = 1, \dots, m, \quad (4.1)$$

where  $x_{it} = (x_{i1t}, x_{i2t}, \dots, x_{int})'$  and  $u_{it}$  is an  $n$  dimensional vector white noise process. The first condition for cointegration among the variables aggregated over sectors in Gonzalo's Theorem 1 is equal to P11: in our notation, the total number of common trends driving the set of  $m$  systems must be smaller than  $n$ . Notice that this number is in general smaller than the sum of trends driving each system, because of the possible presence of inter sector cointegration. An alternative second condition is that the intersection of the null spaces spanned by  $T_i(1)'$  is not empty. This is a particular case of P13, boiling down to requiring the existence of at least one common cointegration vector. The latter is also the main requirement in Lippi (1988) for cointegration to be preserved through aggregation. The third condition that Gonzalo suggested is a mixture of the first two.

Forni and Lippi (1997) adopted a similar framework and further refined the analysis, showing that either cointegration is always transmitted through aggregation, as in our proposition P11, or it holds at the aggregate level only for a set of disaggregate parameters of measure zero, as in P12 and P13.

While both Gonzalo and Forni and Lippi started with a model for each sector, we prefer to adopt a joint model for all the variables under analysis, compare equations (4.1) and (2.4). The similarity in the conditions for cointegration to be preserved through cross sector aggregation is a positive result, it is due to the particular properties of first order integrated processes, as we will see in more detail in next section. But their approach is not ideal for the analysis of the effects of aggregation on short-run characteristics, such as the presence of common cycles. This requires to specify *all* the relationships among the variables, not only those within each sector.

### 4.3. Marginalization

When only  $n_1$  agents per sector, or only  $m_1$  sectors, are studied, if  $n_1 m < r$  or  $n m_1 < r$  there is still cointegration among the marginalized variables  $y$ . Notice also that if there are  $(n - n_1)m$  or  $n(m - m_1)$  common trends that only affect the variables not of interest, P12 is satisfied, and the number of cointegration vectors remains  $p$ . Similarly, cointegration vectors that only involve the variables of interest are preserved, because of P13.

To conclude, aggregation and marginalization can take place contemporaneously. It is simple to determine the proper  $A$  matrix for these cases, e.g., for aggregation over agents and marginalization over sectors there must be only  $m_1$   $e$  vectors in the  $A_A$  matrix in section 3, and to provide examples where cointegration and cofeatures are preserved.

## 5. Some statistical considerations

We now briefly discuss some of the issues that arise when verifying in practice whether the conditions in Propositions 1 and 2 are satisfied or not, assuming that the aggregator matrix  $A$  is known.

The starting point is estimation of the VAR in equation (2.1). If enough observations are available, this poses no problems: equation by equation OLS suffices. Then the number and composition of cointegration and cofeature vectors can be determined by applying standard techniques, see e.g. Johansen (1995) and Engle and Kozicki (1993). This is for example the case when studying the interaction among a small set of variables across a few countries.

Yet, the number of parameters of the VAR grows with the square of that of the variables, so that there will typically be more parameters than observations. Therefore, a priori restrictions on the parameters, such as those considered in section 2, are in general needed. But if they are not valid, the resulting estimators will be biased. This bias can then influence the following analysis, i.e., the determination of the number and composition of the cointegration and cofeature vectors.

There appears to be no solution to the trade-off between estimability and consistency, but it is important to be aware of the problem. Yet, we want to recall the

method developed by Gonzalo and Granger (1995) to get the total number of common trends driving a set of variables from analysis of subsets of them, because it is useful to check the condition P11, i.e., to evaluate whether  $r < j$ .

They propose a two step procedure. In our notation, in the first step cointegration analysis is applied to the  $n$  dimensional systems for  $x_{it}$ ,  $i = 1, \dots, m$ , and the common trends driving each system are determined. In the second step, the number of cointegration relationships (say  $q$ ) among all the estimated common trends (say  $w$ ) is determined. Then, the total number of linearly independent common trends in the full system is  $r = q - w$ . This method is applied, e.g., by Gallo *et al.* (1998) to study the relationships among a set of commodity price indices and the general price level.

If the disaggregated cointegration vectors,  $\beta'$ , can be estimated, P13 requires  $A\beta_{\perp} = 0$  or  $\beta' A'_{\perp} = 0$ , that can be tested by standard likelihood ratio tests, see e.g. Johansen (1995, Ch. 7).

The condition in P12 requires to impose restrictions on the loadings of the common trends in the disaggregated MA representation. The statistical theory to test these hypotheses has to be developed yet, and is an interesting topic for further research.

Finally, when the disaggregated model can be estimated and once the number and composition of the cofeature vectors are determined, the conditions P21, P22, and P23 can be verified by means of standard testing procedures.

## 6. Conclusions

In this paper we have analyzed the effects of linear aggregation of variables that share common trends and, possibly, cycles. Casting the problem in a general framework and for general aggregation schemes, we have provided conditions for cointegration and cofeatures to be still present among the aggregated variables. They require either “enough cointegration” at the disaggregate level, or particular restrictions on the disaggregate equations, some of which can be given an economic justification. Testing for the validity of these conditions is feasible, but it is often complicated by lack of observations.

The results are also easily extended to the case where the common trends and cycles in (2.12) are substituted with the nonlinear common trends in Granger *et al.* (1994) and the non synchronous common cycles in Vahid and Engle (1993b). Instead,

the analysis of the effects of aggregation in more general nonlinear models seems to be an interesting subject for future research.

## References

- [1] Ansley, C.F., Spivey, W.A., and Wroblewski, W.J. (1977), “On the Structure of Moving Average Processes”, *Journal of Econometrics*, 6, pp. 121-134.
- [2] Engle, R.F. and Granger, C.W.J. (1987), “Cointegration and Error Correction: Representation, Estimation and Testing”, *Econometrica*, 55, pp. 251-276.
- [3] Engle, R.F. and Kozicky, S. (1993), “Testing for Common Features”, (with comments), *Journal of Business and Economic Statistics*, 11, pp. 369-395.
- [4] Forni, M. and Lippi, M. (1997), *Aggregation and the Microfoundations of Macroeconomics*, Oxford: Oxford University Press.
- [5] Gallo, G.M., Marcellino, M., and Trivedi, P.K. (1998), “Inflation: the message in commodity price indices”, *mimeo*, European University Institute.
- [6] Gonzalo, J. (1993), “Cointegration and Aggregation”, *Ricerche Economiche*, 47, pp. 281-291.
- [7] Gonzalo, J. and Granger, C.W.J. (1995), “Estimation of Common Long Memory Components in Cointegrated Systems”, *Journal of Business and Economic Statistics*, 13, pp. 27-35.
- [8] Granger, C.W.J. (1980), “Long Memory Relationships and the Aggregation of Dynamic Models”, *Journal of Econometrics*, 14, pp. 227-238.
- [9] Granger, C.W.J. (1987), “Implications of Aggregation with Common Factors”, *Economic Theory*, 3, pp. 208-222.
- [10] Granger, C.W.J. (1990), “Aggregation of Time-Series Variables: A Survey”, in Barker, T. and Pesaran M.H. (eds.), *Disaggregation in Econometric Modelling*, London: Routledge.



- [11] Granger, C.W.J., Inoue, T., and Morin, N. (1994), “Non-linear Stochastic Trends”, *mimeo*, University of California at San Diego.
- [12] Hendry, D.F., Pagan, A.R., and Sargan, J.D. (1984), “Dynamic Specification”, *Handbook of Econometrics*, vol. II, New York: van Nostrand.
- [13] Hendry, D.F. (1995), *Dynamic Econometrics*, Oxford: Oxford University Press.
- [14] Johansen, S. (1988), “Statistical Analysis of Cointegration Vectors”, *Journal of Economic Dynamics and Control*, 12, pp. 231-254.
- [15] Johansen, S. (1991), “Estimation and Hypothesis Testing of Cointegration Vectors in Gaussian Vector Autoregressive Models”, *Econometrica*, 59, pp. 1551-1580.
- [16] Johansen, S. (1995), *Likelihood-Based Inference in Cointegrated Vector Autoregressive Models*, Oxford: Oxford University Press.
- [17] Lippi, M. (1988), “On the Dynamic Shape of Aggregated Error Correction Models” *Journal of Economic Dynamics and Control*, 12, pp. 561-585.
- [18] Lippi, M. (1989), “A Short Note on Cointegration and Aggregation”, W.P. 49, Università di Modena.
- [19] Lütkepohl, H. (1984), “Linear Transformations of Vector ARMA Processes”, *Journal of Econometrics*, 26, pp. 283-293.
- [20] Lütkepohl, H. (1987), *Forecasting Aggregated Vector ARMA Process*, Berlin: Springer-Verlag.
- [21] Marcellino, M. (1998) “Temporal Disaggregation, Missing Observations, Outliers, and Forecasting: A Unifying Non-Model Based Approach”, *Advances in Econometrics*, vol. 13, 181-202.
- [22] Marcellino, M. (1999) “Some Temporal Aggregation Issues in Empirical Analysis”, *Journal of Business and Economic Statistics*, 17, 129-136.
- [23] Nelson, C.R. and Plosser, C.I. (1982), “Trends and Random Walks in Macroeconomic Time Series: Some Evidence and Implication”, *Journal of Monetary Economics*, 10, pp. 139-162.

- [24] Nickell, S. (1985) "Error Correction, Partial Adjustment and All That: An Expository Note", *Oxford Bulletin of Economics and Statistics*, 47, pp. 119-129.
- [25] Stock, J.H. and Watson, M.W. (1988), "Testing for Common Trends", *Journal of the American Statistical Association*, 83, pp. 1097-1107.
- [26] Trivedi, P.K. (1985), "Distributed Lags, Aggregation and Compounding: Some Econometric Implications", *Review of Economic Studies*, 52, pp. 19-35.
- [27] Vahid, F. and Engle, R.F. (1993a), "Common Trends and Common Cycles", *Journal of Applied Econometrics*, 8, pp. 341-360.
- [28] Vahid, F. and Engle, R.F. (1993b), "Non-Synchronous Common Cycles", W.P. ECO 9355, University of California at San Diego.