



Studying co-movements in large multivariate data prior to multivariate modelling[☆]

Gianluca Cubadda^a, Alain Hecq^b, Franz C. Palm^{b,*}

^a *Università di Roma "Tor Vergata", Italy*

^b *Maastricht University, The Netherlands*

ARTICLE INFO

Article history:

Received 3 September 2007

Received in revised form

20 May 2008

Accepted 26 August 2008

Available online 27 September 2008

JEL classification:

C32

Keywords:

Interactions

Multiple time series

Co-movements

ARIMA

Cointegration

Common cycles

Dynamic panel data

ABSTRACT

For non-stationary vector autoregressive models (VAR hereafter, or VAR with moving average, VARMA hereafter), we show that the presence of common cyclical features or cointegration leads to a reduction of the order of the implied univariate autoregressive-integrated-moving average (ARIMA hereafter) models. This finding can explain why we identify parsimonious univariate ARIMA models in applied research although VAR models of typical order and dimension used in macroeconometrics imply non-parsimonious univariate ARIMA representations.

Next, we develop a strategy for studying interactions between variables prior to possibly modelling them in a multivariate setting. Indeed, the similarity of the autoregressive roots will be informative about the presence of co-movements in a set of multiple time series. Our results justify both the use of a panel setup with homogeneous autoregression and heterogeneous cross-correlated vector moving average errors and a factor structure, and the use of cross-sectional aggregates of ARIMA series to estimate the homogeneous autoregression.

© 2008 Elsevier B.V. All rights reserved.

1. Introduction

The analysis of interactions between individuals, groups of agents, organisations, regions or countries is a key objective of scientific inquiry. Huge collections of data (national accounts for virtually all countries, unemployment data for hundreds of types of jobs, prices for thousands of products, assets and services, purchases by individual consumers, environmental data on CO₂ emissions and climate change, on health care, education, ...) are almost instantaneously available at relatively low cost. The IMF and the World Bank publish statistical information for 150 countries. In the European Union, starting for the six founding countries, Eurostat now collects and analyses hundreds of variables

for the 27 member countries. The availability of sets of data for a large number of individual entities and the importance of understanding individual behavior and interaction between individuals have created a demand for methods to analyse these types of phenomena using large sets of data.

To capture interactions between a large set of time series requires imposing a structure on the models used in the analysis. Examples of structure often imposed are the dynamic factor models put forward by Forni et al. (2000) and by Stock and Watson (2002). Another example is the approach of using the common factor structure adopted in the recent literature on dynamic macro-panels to account for cross-sectional correlation (e.g. Bai and Ng (2004)). In these approaches, the common factors usually account for the impact of changes in the environment that is largely exogenous and in which individual entities have to take decisions and to interact. Individual specificities are often accounted for by including individual effects.

Alternatively, partial systems are built for country-specific analyses or for a specific variable such as for instance GDP for a subset of countries or regions. An often adopted framework to analyse a limited number of time series is the vector autoregressive model (VAR hereafter). A VAR, allowing all variables to be endogenous, can only be implemented for small systems, with typically five or six variables. Still, in such cases, the number

[☆] Previous drafts of this paper were presented at the Joint Statistical Meetings in Salt Lake City, 62nd European Meeting of the Econometric Society in Budapest, and EC2 Conference on Advanced Time Series Econometrics in Faro. We wish to thank two anonymous referees and the editor Arnold Zellner for useful comments. The usual disclaimers apply.

* Corresponding address: Maastricht University, Department of Quantitative Economics, P.O.Box 616, 6200 MD Maastricht, The Netherlands. Tel.: +31 43 388 3833.

E-mail address: f.palm@ke.unimaas.nl (F.C. Palm).

of parameters is large. To deal with the dimensionality problem in VARs, different approaches have been adopted such as inter alia Bayesian analysis, simulation-based techniques, separability assumptions (as for instance used in panel data models when they are accompanied by the homogeneity assumption of the autoregressive parameters across units), automatic selection by deleting non-significant coefficients and reduced-rank regressions.

In this paper, we take a different route and make two new contributions to the literature. Instead of designing a system which captures interactions between entities, we first address the question what the implications are of the presence of common factors in a possibly large VAR system on the marginal processes for individual entities. The common factor structure could arise from the presence of common cyclical or other features (see e.g. Engle and Kozicki (1993)). For non-stationary series, a common factor structure could result from the presence of cointegration, that is from the occurrence of common stochastic trends. The first main contribution of the paper is that the presence of common factors or common features in a VAR leads to a reduction of the order of the implied univariate autoregressive-moving average (ARIMA hereafter) schemes for the individual series. In a way, individual series keep a print of the system as a whole. Second, we propose a strategy, that allows to identify the print of common features from individual data, that is to study aspects of interactions between individual variables without or prior to modelling these interactions in a multivariate framework.

Before implementing this strategy, we derive the implications of the presence of a common factor structure for the univariate processes implied by a VAR, both in absence of and under exogeneity restrictions. More precisely, in Section 2, we extend results by Zellner and Palm (1974) by showing that VARs with reduced rank restrictions coming from short- or long-run co-movements (see inter alia Vahid and Engle (1993) and Ahn (1997)) offer an alternative explanation to the well known paradox in time series: on one hand, small multivariate systems imply non-parsimonious individual ARIMA processes whereas on the other hand empirically selected and estimated univariate models are generally of low order.

Section 3 proposes and evaluates an estimation strategy based upon the common univariate autoregressive parameters of series generated by the same VAR system. Indeed, a homogeneous panel framework with common autoregressive parameters is the by-product of a large VAR. We propose to use cross-sectional aggregates to estimate these common autoregressive parameters, and we evaluate this estimation strategy in a small Monte Carlo experiment. Section 4 focuses on the link between the growth rates of GDP among Latin American economies. The occurrence of common autoregressive parts for the individual series is used as an indication that these series exhibit the same common features and therefore are subject to the same co-movements. This indication is then formally tested and found to be supported by the information in the data. A final section concludes and provides suggestions for further research.

2. Multivariate and implied univariate time series models

The aim of this section is to present new results on the relations between univariate and multivariate time series models. In particular, Section 2.1 briefly reviews the paradox, namely that VARs of typical order and dimension used in macroeconometrics imply non-parsimonious univariate ARMA representations, which are rarely observed in empirical applications. Section 2.2 shows how VARs with certain reduced rank structures can solve this paradox, and Section 2.3 specialises the results to the case of block-diagonal and block-triangular VARs. Moreover, Section 2.4 explores the implications of short-run co-movements for the

vector moving average (VMA hereafter) component of the so-called final equation representation of a VAR model (Zellner and Palm, 1974). The findings may explain why MA orders of univariate models are often found to be smaller than those theoretically implied by multivariate models. Finally, Section 2.5 illustrates concepts and methods with an empirical analysis of the relationship between quarterly growth rates of the industrial production indexes in the US and Canada.

2.1. The paradox

The literature on the link between multivariate time series and the behavior of univariate variables stresses the fact that univariate ARIMA analyses provide tools for the diagnostic testing of VAR (or VARMA) models. Important contributions in this area are the monograph by Quenouille (1957) and papers by Zellner and Palm (1974, 1975, 2004), Palm and Zellner (1980), Palm (1977) or Maravall and Mathis (1994). Some textbooks devote a few pages to that issue (e.g. Franses (1998, 198–199) or Lütkepohl (2005, 494–495)). More formally, let us first consider the stationary¹ multivariate VARMA process for an n vector of time series $z_t = (z_{1t}, \dots, z_{nt})'$, and without deterministic terms for simplicity:

$$\Phi(L)z_t = \Theta(L)\varepsilon_t, \quad t = 1, \dots, T, \quad (1)$$

where $\Phi(L)$ and $\Theta(L)$ are finite order polynomial coefficient matrices with the usual lag operator L such that $Lz_t = z_{t-1}$. The sequence of ε_t is a multivariate white noise process with each $\varepsilon_t \sim N(0, \Sigma)$. The VARMA model in (1) encompasses several useful specifications. For example, if $\Phi(L) \equiv I$ we have the vector moving average representation or VMA; if $\Theta(L) \equiv I$ we have a VAR(p). Let us concentrate the analysis on VAR models of order p , denoted VAR(p), one of the cornerstone specification in empirical macroeconometrics:

$$\Phi(L)z_t = \varepsilon_t. \quad (2)$$

Following Zellner and Palm (1974), the univariate representation of elements of z_t can be obtained by premultiplying both sides of (2) by $\Phi(L)^{adj}$, the adjoint matrix (or the adjugate) associated with $\Phi(L)$, in order to obtain the “final equations” (FEs henceforth):

$$\det[\Phi(L)]z_t = \Phi(L)^{adj}\varepsilon_t, \quad (3)$$

where the determinant $\det[\Phi(L)]$ is a scalar finite order polynomial in L . This means that each series is a finite order ARMA(p^* , q^*), with the same lag structure and the same coefficients for the autoregressive part for every series, although the multivariate system is a finite order VAR(p). For instance for the i th element of z_t we have

$$\det[\Phi(L)]z_{it} = \Phi_i(L)^{adj}\varepsilon_t = \theta_i(L)u_{it},$$

$\Phi_i(L)^{adj}$ denoting the i th row of the matrix $\Phi(L)^{adj}$, $\theta_i(L)$ is a scalar polynomial in L and u_{it} is a scalar innovation with respect to the past of z_{it} . This recognition and the compatibility of these p^* and q^* with n and p of the VAR(p) is the first step of an approach that has been developed in Zellner and Palm (1974, 1975) as a general modelling strategy called SEMTSA (*Structural Economic Modelling Time Series Analysis*). In the first stage of this approach, information from univariate schemes is used to restrict the dynamics of the structural model. However, given the conditions associated with that first diagnostic checking, one often faces a paradoxical situation: a small order VAR with few series already generates univariate ARMA with large p^* and q^* , an implication that is

¹ The generalisation to non-stationarity processes will be considered later in this section.

rejected when tested on economic data where one usually finds quite parsimonious models with low order autoregressive and moving average polynomials.

Indeed, an n dimensional VAR(p) would imply **at most** individual ARMA(np , $(n-1)p$) processes. This well known result is simply due to the fact that $\det[\Phi(L)]$ contains by construction up to L^{np} terms and the adjoint matrix is a collection $(n-1) \times (n-1)$ cofactor matrices, each of the matrix elements can contain the terms $1, L, \dots, L^p$. But whatever the simplicity of that result, it leads to implausible univariate models for most economic series.² This is illustrated in the following numerical example.

Example 1. Let us consider the bivariate VAR(2), $\Phi(L)z_t = (I - \Phi_1L - \Phi_2L^2)z_t = \varepsilon_t$ with the following numerical values

$$\Phi(L) = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} - \begin{bmatrix} -0.1 & -0.21 \\ 0.7 & 0.76 \end{bmatrix}L - \begin{bmatrix} 0.3 & 0.8 \\ 0.3 & -0.8 \end{bmatrix}L^2,$$

the determinant of which is $-0.48L^4 - 0.189L^3 + 0.571L^2 - 0.66L + 1$ and the adjoint matrix such that

$$\Phi(L)^{adj} = \begin{bmatrix} 1 - 0.76L + 0.8L^2 & -0.21L + 0.8L^2 \\ 0.7L + 0.3L^2 & 1 + 0.1L - 0.3L^2 \end{bmatrix}.$$

The two implied individual series are ARMA(4,2).

Wallis (1977) has proposed two explanations to solving this paradox. The first solution is to unravel and get rid of common roots in $\det[\Phi(L)]$ and $\Phi(L)^{adj}$. This requirement of minimal representation is not always trivial to implement in practice. The cancelling of common factors occurs for instance when the autoregressive matrix $\Phi(L)$ is diagonal, block-diagonal, triangular or block-triangular as pointed out by Zellner and Palm (1974) and Palm (1977). Under a block-triangular structure and a corresponding block-diagonality of Σ , a subset of z_t is exogenous. The second intuition raised by Wallis is based on the observation that coefficients close to but different from zero in the implied ARMA model might give the feeling that models are more parsimonious than they theoretically should be. This explanation is always inherent in most of the papers by Zellner and Palm where economic theories are tested and restrictions are imposed on the system which are compatible with implied individual ARMA models. Also, the strategies using algorithms to systematically delete non-significant coefficients in VAR models (see inter alia Brüggemann et al. (2003)) fall into this setting. The next subsection provides an alternative explanation for solving that paradox thanks to the use of short-run co-movements and reduced rank models.

2.2. Solving the paradox under common features

We introduce a new framework aimed at understanding the gap between implied individual ARMA processes for VARs and the estimated univariate models that we get in empirical studies. More specifically, we investigate the implications of the presence of common features such as common cyclical features and common trends in a multivariate dynamic model for the order of the marginal processes of the individual series.

The strongest form of common cyclical features is the notion of serial correlation common feature (henceforth, SCCF) proposed by Engle and Kozicki (1993) and Vahid and Engle (1993).³ In this framework, series z_t have s SCCF relationships if there exists an

$n \times s$ ($s < n$) matrix δ with full column rank and such that $\delta'\Phi(L)z_t = \delta'z_t = \delta'\varepsilon_t$ in (2). Hence, SCCF implies that impulse response functions of series in z_t are collinear.

Although convenient in terms of parsimony and economic interpretation of shocks' transmission mechanisms (see, inter alia, Centoni et al. (2007)), assumptions underlying SCCF may be too strong. Indeed, the definition of SCCF does not take account of the possibility that common serial correlation is present among non-contemporaneous elements of series z_t (see e.g. Ericsson (1993)). In order to overcome this limitation, several variants of the SCCF were proposed, see, inter alia, Vahid and Engle (1997), Schleicher (2007), Hecq et al. (2006), Cubadda (2007), and Cubadda and Hecq (2001). In particular, these latter authors assume that there exists an $n \times s$ polynomial matrix such that $\delta'(L)z_t \equiv (\delta_0 + \delta_1L)z_t = \delta'_0\varepsilon_t$. This model is a polynomial SCCF model of order one, which is denoted PSCCF(1).⁴ The notion of PSCCF can be generalised to account for adjustment delays of m periods instead of one lag as in the PSCCF(1). In this case the polynomial matrix $\delta(L)$ is of order m , where $m \leq p$, a model denoted PSCCF(m) (see Cubadda and Hecq (2001) for details).

For simplicity reasons, we start with the most parsimonious model, in which there exists a SCCF matrix. We first illustrate the consequences of the presence of one SCCF relationship in a bivariate example with $p = 2$ before generalising to the n dimensional case for any polynomial order p and other forms of non contemporaneous co-movements (e.g. PSCCF).

Example 2. Let us consider the polynomial matrix $\Phi(L)$ of the Example 1 but now there exists a SCCF relationship with a cofeature vector $\delta' = (1 : 1)$ for

$$\Phi(L) = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} - \begin{bmatrix} -0.1 & -0.2 \\ 0.1 & 0.2 \end{bmatrix}L - \begin{bmatrix} 0.2 & 0.6 \\ -0.2 & -0.6 \end{bmatrix}L^2. \quad (4)$$

Because the second and the third coefficient matrices Φ_1 and Φ_2 are of reduced rank with the same left null space, it follows that $\delta'\Phi(L) = \delta'$. The determinant of $\Phi(L)$ in (4) is $0.4L^2 - 0.1L + 1$ and the adjoint matrix is

$$\Phi(L)^{adj} = \begin{bmatrix} -0.2L + 0.6L^2 + 1 & -0.2L + 0.6L^2 \\ 0.1L - 0.2L^2 & 0.1L - 0.2L^2 + 1 \end{bmatrix}.$$

This implies that the two series are **at most** ARMA(2,2) and not ARMA(4,2) as in Example 1.

More generally, Table 1 summarises the reduction of the individual ARMA orders due to common feature restrictions. Basically, Table 1 documents that multivariate systems with additional commonality deliver more parsimonious, and therefore empirically more plausible, ARMA models than those implied by unrestricted VAR models. For instance, in a small VAR system with four variables and two lags, implied models are at most ARMA(8,6) while they would reduce to ARMA(2,2) models in the presence of a unique common cycle, that is, with $s = 3$. This application to Quenouille's results, who already stressed the impact of rank deficiency, is one of the diagnostic tools for data-admissibility that we want to formalise and emphasise. Note also that additional zero coefficient restrictions on common factors and/or loadings may reduce these implied maximal orders. This means that the ARMA(2,2) obtained under short-run co-movements can be more parsimonious with additional exclusion restrictions.

Let us formally prove the results that are reported in Table 1.

² This result generalises to individual ARMA(np , $(n-1)p + q$) for VARMA(p , q) processes (see e.g. Lütkepohl (2005)).

³ This model is also known as a scalar component model of order zero or SCM(0,0) using the terminology of Tiao and Tsay (1989) or white noise direction copendence by Gouriéroux and Peaucelle (1989).

⁴ Notice that a SCCF relationship involving, for instance, z_{1t} and z_{2t-1} , where $z_t = (z'_{1t}, z'_{2t})'$, is a particular case of a PSCCF(1).

Proposition 3 (Stationary VAR with s SCCF). In an n -dimensional stationary VAR(p) with s SCCF, the individual ARMA processes have: (i) AR orders not larger than $(n - s)p$; (ii) MA orders not larger than $(n - s)p$.

Proof. Let us rewrite Eq. (2) as follows

$$Q(L)x_t = e_t,$$

where $x_t = Mz_t$, $e_t = M\varepsilon_t$, $Q(L) = M\Phi(L)M^{-1}$, $M' \equiv [\delta : \delta_\perp]$, and δ_\perp is the orthogonal complement. Given that x_t is a non-singular linear transformation of z_t , the maximum AR and MA orders of the univariate representation of elements of z_t must be the same as those of elements of x_t . Since $M^{-1} = [\bar{\delta} : \bar{\delta}_\perp]$, where $\bar{\delta} = \delta(\delta'\delta)^{-1}$, and $\bar{\delta}_\perp = \delta_\perp(\delta'_\perp\delta_\perp)^{-1}$, we have

$$Q(L) = \begin{bmatrix} I_s & 0_{s \times (n-s)} \\ \delta'_\perp \Phi(L) \bar{\delta} & \delta'_\perp \Phi(L) \bar{\delta}_\perp \end{bmatrix},$$

from which it easily follows that $\det[Q(L)] = \det[\delta'_\perp \Phi(L) \bar{\delta}_\perp]$ is a polynomial of order $(n - s)p$. Hence, the univariate AR order of each element of z_t is, at most, $(n - s)p$. To prove (ii), let $P(L)$ denote a submatrix of $Q(L)$ that is formed by deleting one of the first s rows and one of the first s columns of $Q(L)$. We can partition $P(L)$ as follows

$$P(L) = \begin{bmatrix} P_{11} & P_{12} \\ P_{21}(L) & P_{22}(L) \end{bmatrix}. \tag{5}$$

Now, P_{11} is a $(s - 1) \times (s - 1)$ identity matrix, P_{12} is a $(s - 1) \times (n - s)$ matrix of zeros, $P_{21}(L)$ is a $(n - s) \times (s - 1)$ polynomial matrix of order p , and $P_{22}(L)$ is a $(n - s) \times (n - s)$ polynomial matrix of order p . Hence, $\det[P(L)] = \det[P_{11}] \det[P_{22}(L)]$, which tells us that $\det[P(L)]$ is of order $(n - s)p$. Since cofactors associated with the blocks of $Q(L)$ different from P_{11} are polynomials of degree not larger than $(n - s)p$, we conclude that the univariate MA order of each element of z_t is, at most, $(n - s)p$. ■

Proposition 4 (Stationary VAR with s PSCCF(1)). In an n stationary VAR(p) with s PSCCF(1) the individual ARMA processes have: (i) AR orders not larger than $(n - s)p + s$; (ii) MA orders not larger than $(n - s)p + (s - 1)$.

Proof. Along the same lines of the previous proof, let us consider $R(L) = M_0\Phi(L)M_0^{-1}$, $M'_0 \equiv [\delta_0 : \delta_{0\perp}]$, and $M_0^{-1} \equiv [\bar{\delta}_0 : \bar{\delta}_{0\perp}]$. Then, we have

$$R(L) = \begin{bmatrix} \delta'(L)\bar{\delta}_0 & \delta'_1\bar{\delta}_{0\perp L} \\ \delta'_{0\perp}\Phi(L)\bar{\delta}_0 & \delta'_{0\perp}\Phi(L)\bar{\delta}_{0\perp L} \end{bmatrix}.$$

Hence,

$$\det[R(L)] = \det[\delta'(L)\bar{\delta}_0] \det[\delta'_{0\perp}\Phi(L)\bar{\delta}_{0\perp L} - \delta'_{0\perp}\Phi(L)\bar{\delta}_0(\delta'(L)\bar{\delta}_0)^{-1}\delta'_1\bar{\delta}_{0\perp L}].$$

Writing $(\delta'(L)\bar{\delta}_0)^{-1} = (\delta'(L)\bar{\delta}_0)^{adj} / \det[\delta'(L)\bar{\delta}_0]$ and substituting it above we get

$$\begin{aligned} \det[R(L)] &= \underbrace{\det[\delta'(L)\bar{\delta}_0]^{n-s-1}}_{s(n-s-1)} \\ &= \det \left\{ \underbrace{\delta'_{0\perp}\Phi(L)\bar{\delta}_{0\perp L}}_p \underbrace{\det[\delta'(L)\bar{\delta}_0]}_s \right. \\ &\quad \left. - \underbrace{\delta'_{0\perp}\Phi(L)\bar{\delta}_0}_p \underbrace{(\delta'(L)\bar{\delta}_0)^{adj}}_{s-1} \underbrace{\delta'_1\bar{\delta}_{0\perp L}}_1 \right\} \end{aligned}$$

Table 1

Maximum ARMA orders of univariate series generated by an n -dimensional $I(0)$ VAR(p) with s cofeature restrictions.

Models	AR order	MA order
VAR(p)	np	$(n - 1)p$
SCCF	$(n - s)p$	$(n - s)p$
PSCCF(1)	$(n - s)p + s$	$(n - s)p + (s - 1)$
PSCCF(m)	$(n - s)p + sm$	$(n - s)p + (s - 1)m$

from which follows that the degree of the polynomial of $\det[R(L)]$ is, at most, equal to $(n - s)p + s$. Regarding the order of the MA component, let us denote $G(L)$ a submatrix of $R(L)$ that is formed by deleting one of the first s rows and one of the first s columns of $R(L)$. We can partition $G(L)$ as follows

$$G(L) = \begin{bmatrix} G_{11}(L) & G_{12}L \\ G_{21}(L) & G_{22}(L) \end{bmatrix},$$

where $G_{11}(L)$ is a $(s - 1) \times (s - 1)$ polynomial matrix of order 1, G_{12} is a $(s - 1) \times (n - s)$ matrix, $G_{21}(L)$ is a $(n - s) \times (s - 1)$ polynomial matrix of order p , and $G_{22}(L)$ is a $(n - s) \times (n - s)$ polynomial matrix of order p . Hence, following a similar reasoning as above, we conclude that the individual MA order is, at most, $(n - s)p + (s - 1)$. ■

These results can be easily generalised for the PSCCF(m) case as reported in Table 1. Also note that we do not consider a mixed model that can have both SCCF and PSCCF relationships but results can be trivially deduced from results beneath.

The above propositions can be extended to the case of an $I(1)$ VAR. Let us consider the vector error correction model (VECM henceforth) of variables z_t

$$\Gamma(L)\Delta z_t = \alpha\beta'z_{t-1} + \varepsilon_t, \tag{6}$$

where $\Delta = (1 - L)$, $\Gamma(L) = I_n - \sum_{i=1}^{p-1} \Gamma_i L^i$, $\Gamma_i = -\sum_{j=i+1}^p \Phi_j$ for $i = 1, 2, \dots, p - 1$, α and β are full-column rank $n \times r$ ($r < n$) matrices such that $\Phi(1) = -\alpha\beta'$ and $\alpha'_\perp\Gamma(1)\beta_\perp$ has full rank. The process z_t is cointegrated of order $(1, 1)$, denoted by $CI(1, 1)$, the columns of β span the cointegrating space, the elements of α are the corresponding adjustment coefficients, see e.g. Johansen (1996).

A well known implication of cointegration is that $(n - r - 1)$ unit roots are common to the individual AR and the MA polynomial. Let us formalise this result in order to later extend it to different forms of common cyclical features. To do so, consider again Eq. (2) and apply the transformation

$$B(L)y_t = u_t,$$

where $y_t = Pz_t$, $u_t = P\varepsilon_t$, $B(L) = P\Phi(L)P^{-1}$, and $P' \equiv [\beta : \beta_\perp]$. Again, the maximum AR and MA orders of the univariate representation of elements of z_t are the same as those of elements of y_t . Since $\beta'z_t$ is $I(0)$ and $\beta'_\perp z_t$ is $I(1)$, we can partition $B(L)$ as follows

$$B(L) = \begin{bmatrix} B_{11}(L) & \Delta B_{12}(L) \\ B_{21}(L) & \Delta B_{22}(L) \end{bmatrix}.$$

We have that

$$\begin{aligned} \det[B(L)] &= \Delta^{n-r} \det[B_{22}(L)] \det[B_{11}(L) - B_{12}(L)B_{22}(L)^{-1}B_{21}(L)] \\ &= \Delta^{n-r} \det[B_{22}(L)] \det[B^*(L)]. \end{aligned}$$

Since $B_{22}(1) \neq 0$, and $B^*(1) = B_{11}(1) \neq 0$, there are $(n - r)$ unit roots in $\det[B(L)]$. However, since $\beta'_\perp z_t$ is $I(1)$, there must be $(n - r - 1)$ unit roots in the univariate MA polynomials of elements of $\beta_\perp z_t$. Hence, the AR and MA orders of elements of z_t respectively are, at most, $n(p - 1) + r + 1$ and $(n - 1)(p - 1) + r$.

Table 2

Maximum ARMA orders of univariate series generated by an n -dimensional $CI(1, 1)$ $VAR(p)$ with s cofeature restrictions.

Models	AR order	Integration order	MA order
$VAR(p)$	$n(p - 1) + r$	1	$(n - 1)(p - 1) + r$
SCCF	$(n - s)(p - 1) + r$	1	$(n - s)(p - 1) + r$
PSCCF(1)	$(n - s)(p - 1) + r + s$	1	$(n - s)(p - 1) + r + s - 1$
WF	$(n - s)(p - 1) + r$	1	$(n - s)(p - 1) + r$

For the VECM in (6), the existence of s SCCF's implies that there exists an $n \times s$ full-rank matrix δ such that $\delta' \Delta z_t = \delta' \varepsilon_t$ or equivalently

$$\delta' \Phi(L) = \delta' \Delta.$$

As shown by Vahid and Engle (1993), the presence of s SCCF's is equivalent to the existence of $(n - s)$ common and synchronous cycles.⁵

The existence of s PSCCF(1) relationships requires, instead, that there exists an $n \times s$ full-rank matrix δ_0 such that $\delta_0' \Delta z_t + \delta_1' \Delta z_{t-1} = \delta_0' \varepsilon_t$ where $\delta_1 = -\Gamma_1' \delta_0$ or equivalently

$$\delta(L)' \Phi(L) = \delta_0' \Delta.$$

The presence of the PSCCF(1) has an interesting implication for the cycles of series z_t . Indeed, Cubadda and Hecq (2001) proved that the same PSCCF(1) relationships cancel the dependence from the past of both the first differences and cycles of series z_t .

Finally, the existence of s weak form SCCF (WF henceforth) implies that there exists an $n \times s$ full-rank matrix $\tilde{\delta}$ such that $\tilde{\delta}' \Delta z_t - \tilde{\delta}' \alpha \beta' z_{t-1} = \tilde{\delta}' \varepsilon_t$ (Hecq et al., 2006) or equivalently

$$\tilde{\delta}(L)' \Phi(L) = \tilde{\delta}',$$

where $\tilde{\delta}(L) = \tilde{\delta} + \tilde{\delta}_1 L$, $\tilde{\delta}_1 = -(\beta \alpha' + I_n) \tilde{\delta}$. Interestingly, the same WF relationships cancel the dependence from the past of both the levels and cycles of series z_t (Cubadda, 2007). Hence, we can interpret the WF as an analogous property to the PSCCF(1) that applies to the levels rather than to the differences of series z_t .

Table 2 reports the largest ARIMA orders of individual series that are generated by a VECM with common feature restrictions. Let us formally prove these results.

Proposition 5 (Cointegrated VAR with s SCCF). *In a $CI(1, 1)$ $VAR(p)$ with s SCCF the individual ARMA processes of the first differences have: (i) AR orders not larger than $(n - s)(p - 1) + r$; (ii) MA orders not larger than $(n - s)(p - 1) + r$.*

Proof. By construction of M , and keeping in mind that $\delta' \Phi(L) = \delta' \Delta$, we note that

$$\det[Q(L)] = \det \begin{bmatrix} \Delta I_s & 0_{s \times (n-s)} \\ \delta'_\perp \Phi(L) \tilde{\delta} & \delta'_\perp \Phi(L) \tilde{\delta}_\perp \end{bmatrix} = \Delta^s \det[\delta'_\perp \Phi(L) \tilde{\delta}_\perp]$$

is a polynomial of order $(n - s)p + s$. Since $\det[Q(L)]$ is a polynomial of the same order as $\det[\Phi(L)]$, we conclude that the univariate AR order of each element of z_t is, at most, $(n - s)p + s$. Note that this result means that the AR part is of order $(n - s)(p - 1) + n$. The polynomial degree of the moving average component is obtained along the same line as in Proposition 3 but with $P_{11} \Delta$ instead of P_{11} in (5). Hence, $\det[P(L)] = \det[P_{11} \Delta] \det[P_{22}(L)]$, which tells us that $\det[P(L)]$ is of order $(n - s)p + (s - 1)$. Since the presence of cointegration implies that $(n - r - 1)$ common unit roots cancel out from the individual AR and the MA polynomials, the result follows. ■

⁵ In this framework, cycles are the remainder of a vector $I(1)$ process after subtracting the (possibly common) random-walk components.

Proposition 6 (Cointegrated VAR with s PSCCF(1)). *In a $CI(1, 1)$ $VAR(p)$ with s PSCCF (1) the individual ARMA processes of the first differences have: (i) AR orders not larger than $(n - s)(p - 1) + s + r$; (ii) MA orders not larger than $(n - s)(p - 1) + r + s - 1$.*

Proof. Since it is similar to that of Proposition 4, it is not reported to save space. ■

Proposition 7 (Cointegrated VAR with s WF). *In a $CI(1, 1)$ $VAR(p)$ with s WF the ARMA processes of the first differences have: (i) AR orders not larger than $(n - s)(p - 1) + r$; (ii) MA orders not larger than $(n - s)(p - 1) + r$.*

Proof. Let us consider $H(L) = \tilde{M} \Phi(L) \tilde{M}^{-1}$, $\tilde{M}' \equiv [\tilde{\delta} : \tilde{\delta}_\perp]$, and

$$H(L) = \begin{bmatrix} \tilde{\delta}'(L) \tilde{\delta} (\tilde{\delta} \tilde{\delta}')^{-1} & \tilde{\delta}'_1 \tilde{\delta}_\perp (\tilde{\delta}'_1 \tilde{\delta}'_\perp)^{-1} L \\ \tilde{\delta}'_\perp \Phi(L) \tilde{\delta} (\tilde{\delta} \tilde{\delta}')^{-1} & \tilde{\delta}'_\perp \Phi(L) \tilde{\delta}_\perp (\tilde{\delta}'_\perp \tilde{\delta}'_\perp)^{-1} \end{bmatrix}.$$

Hence, the proof is entirely analogous to that of Proposition 4 with the $H(L)$ matrix in place of $R(L)$. ■

Note that the results for the VAR in the presence of common features can be extended in a fairly trivial way for the VARMA model. In the presence of common features in a VARMA(p, q) model, the orders reported in Tables 1 and 2 hold true as well for the AR part whereas the order of the MA parts has to be augmented by q (at most). If common feature restrictions affect the VAR part only, the MA part of the univariate processes is expected to be increased by q . The proofs are similar to those reported for the VAR(p) model except for the presence of the q extra lags in the MA parts. If the common features restrictions affect the MA part of the VARMA model, the increase of the order of the MA part of the implied univariate models might be smaller than q .

2.3. Block-diagonal and block-triangular VARs and the separation hypothesis

From the results of the previous subsection, there naturally arises the question about block-diagonal and block-triangular systems, namely about a situation in which we can disentangle groups of variables having many within co-movements but being roughly independent of other sets of variables. Indeed, intuition suggests that additional restrictions due to block-diagonality or block-triangularity should induce further simplifications of the individual ARMA structures. A block-triangular structure is required for the absence of Granger-causation between subvectors of z_t accordingly partitioned and for strong exogeneity.

For the sake of simplicity, we consider two groups of variables and SCCF restrictions only. The first model under analysis is a block-diagonal stationary VAR(p) with SCCF within each block of variables n_1 and n_2 respectively, with $n_1 + n_2 = n$ such that

$$\Phi(L) = \begin{bmatrix} \Phi_{11}(L) & 0_{n_1 \times n_2} \\ 0_{n_2 \times n_1} & \Phi_{22}(L) \end{bmatrix},$$

where there exist two full column rank matrices with ranks s_1 and s_2 respectively, with $s_1 + s_2 = s$ such that $\delta'_1 \Phi_{11}(L) = \delta'_1$ and $\delta'_2 \Phi_{22}(L) = \delta'_2$. This is a system with separability in common features (Hecq et al., 2002) with a block-diagonal co-feature matrix such that

$$\delta = \begin{bmatrix} \delta_1 & 0_{n_1 \times s_2} \\ 0_{n_2 \times s_1} & \delta_2 \end{bmatrix}.$$

In view of Table 1, the maximum orders for the implied univariate ARMA processes are $(n_1 + n_2 - s_1 - s_2)p$ for both the AR and MA components, assuming the same polynomial order p for each group. With different polynomial orders p_1 and p_2 , these orders become $(n_1 - s_1)p_1 + (n_2 - s_2)p_2$ respectively. However, it is clear that these orders can be further reduced due to the presence of

blocks of zeros. Because this leads to the presence of common roots between the AR and MA parts, the maximum orders of both AR and MA components are $(n_1 - s_1)p_1$ for the first block and $(n_2 - s_2)p_2$ for the second group of series.

Due to separation in the form of block-diagonality of $\Phi(L)$, the implied univariate ARMA-processes need not to have identical AR polynomials. In fact, the model allows for inter-block AR parameter heterogeneity and intra-block AR parameter homogeneity. Note that this form of separation is often used when studying data for a set of countries using a panel-data framework. In some panel-data studies, a single variable is analysed for a set of countries which can be clustered in groups with intergroup block-diagonality of $\Phi(L)$ combined with intragroup homogeneity so that the implied univariate ARMA models for a given group have identical AR polynomials.

The second model is a block-triangular VAR such that

$$\Phi(L) = \begin{bmatrix} \Phi_{11}(L) & \Phi_{12}(L) \\ 0_{n_2 \times n_1} & \Phi_{22}(L) \end{bmatrix},$$

with $\delta'_1[\Phi_{11}(L) : \Phi_{12}(L)] = \delta'_1$ and $\delta'_2\Phi_{22}(L) = \delta'_2$. The situation is similar as before because δ_\perp is still block-diagonal. However fewer cancellations of common roots are observed and the implied models are of AR degree $(n_1 - s_1)p_1 + (n_2 - s_2)p_2$ for block 1 and $(n_2 - s_2)p_2$ in the block 2.

In this case, the AR polynomials of the two blocks could have factors in common which are identical to the AR polynomial of the second block. Notice also that under the additional assumption of appropriate block-diagonality of the contemporaneous covariance matrix Σ , the $n_2 \times 1$ subvector x_t of $z_t = (y'_t, x'_t)'$ is strongly exogenous. Multiplying the subsystem for y_t by the adjoint matrix of $\Phi_{11}(L)$, one obtains the set of transfer functions (TF henceforth) for y_t , that is a set of dynamic equations with the same scalar AR polynomial, and vector moving averages in x_t and in ε_{1t} , with ε_{it} being the subvector of ε_t corresponding to the partitioning of z_t as $(y'_t, x'_t)'$. These TF equations are conditional models which can be used to study the dynamics of y_t or each of its components given x_t and its past and the past of y_t .

Obviously, similar results can be obtained for the general case of block-diagonal or block-triangular systems with k blocks.

2.4. Interpretation in terms of a VMA index model

This subsection further investigates the consequences of the presence of short-run co-movements for the VMA part of FEs. We consequently look at the adjoint matrix of $\Phi(L)$ and emphasise that the VMA representation follows a sort of multivariate index model (see Reinsel (1983)). In order to introduce the problem, let us look more closely at the adjoint matrix $\Phi(L)^{adj}$, we have computed in Example 2 for $s = 1$ SCCF relationship, i.e.

$$\Phi(L)^{adj}\varepsilon_t = \begin{bmatrix} 1 - 0.2L + 0.6L^2 & -.2L + .6L^2 \\ 0.1L - 0.2L^2 & 1 + 0.1L - 0.2L^2 \end{bmatrix} \begin{bmatrix} \varepsilon_{1t} \\ \varepsilon_{2t} \end{bmatrix}.$$

We can also write the previous expression in terms of coefficient matrices and we observe the presence of a common right factor, namely of common right null spaces such that

$$\Phi(L)^{adj}\varepsilon_t = \begin{bmatrix} \varepsilon_{1t} \\ \varepsilon_{2t} \end{bmatrix} + \begin{bmatrix} -.2 \\ 0.1 \end{bmatrix} \begin{bmatrix} 1 & 1 \end{bmatrix} \begin{bmatrix} \varepsilon_{1t-1} \\ \varepsilon_{2t-1} \end{bmatrix} + \begin{bmatrix} 0.6 \\ -.2 \end{bmatrix} \begin{bmatrix} 1 & 1 \end{bmatrix} \begin{bmatrix} \varepsilon_{1t-2} \\ \varepsilon_{2t-2} \end{bmatrix}$$

with the obvious numerical observation that $\Phi(L)^{adj}\delta_\perp = \delta_\perp$, and $\delta'_\perp = (1 : -1)$. We now show that the VMA component of (3) has a structure that is analogous to the multivariate index model of Reinsel (1983), and this leads to the following general proposition:

Proposition 8. In a stationary VAR(p), the existence of s SCCF vectors implies that in the FE representation the VMA coefficient matrices associated with degrees strictly larger than $(n - s - 1)p$ have a common right null space that is spanned by δ_\perp . Hence, post-multiplying the adjoint matrix $\Phi(L)^{adj}$ by δ_\perp reduces the order of the VMA component to a degree of at most $(n - s - 1)p$ instead of $(n - s)p$.

Proof. We know from Proposition 3 that $\det[\Phi(L)]$ and $\Phi(L)^{adj}$ are both polynomials of order $(n - s)p$. Moreover, under SCCF we have that

$$\Phi(L) = I_n - \sum_{j=1}^p \delta_\perp \Psi'_j L^j,$$

where $\Phi_j = \delta_\perp \Psi'_j$. Hence we can write

$$\det[\Phi(L)]I_n = \Phi(L)^{adj}\Phi(L) = \Phi(L)^{adj} - [\Phi(L)^{adj}\delta_\perp] \sum_{j=1}^p \Psi'_j L^j,$$

from which it follows that $\Phi(L)^{adj}\delta_\perp$ is a polynomial of order $(n - s - 1)p$. ■

Corollary 9. In the particular case with $n - 1 = s$, the VMA component of the FEs follows an index model as in Reinsel (1983).

The last result is more a mathematical curiosity than a device to be used in an empirical analysis. We might think to use it for fully efficient estimation of the FEs for instance. But this shows that there exists a factor structure in the VMA component. In empirical investigations, this result can partially explain why the MA order is often found to be smaller than what it should be according to the theoretical implied models in Tables 1 and 2. Consider, for instance, the case with $w_t = \varepsilon_t + H_1\varepsilon_{t-1}$, where $w_t = \Phi(L)^{adj}\varepsilon_t$. The first autocovariance of $w_t = \det[\Phi(L)]z_t$ is then given by $E(w_{t-1}w'_t) = H_1\Sigma$ and, since δ_\perp spans the right null space of H_1 , we have $E(w_{t-1}w'_t) = 0$ when $\Sigma = \delta_\perp\Lambda\delta'_\perp$, where Λ is a symmetric semi-positive definite matrix. Of course, under the condition $\Sigma = \delta_\perp\Lambda\delta'_\perp$ the VAR innovations would be perfectly collinear and this cannot occur in practice. However, it might well be that $\Sigma \approx \delta_\perp\Lambda\delta'_\perp$, and we label this case as a near coincidental situation.

2.5. An example: Industrial output growth in Canada and the US

Engle and Kozicki (1993) find that there exists a SCCF relationship between the Canadian and the US quarterly growth rates of output (seasonally adjusted series). They have considered a sample from the late 1950s to the late 1990s. We select the same countries and we use the seasonally adjusted industrial production indexes from OECD main indicators over the period 1960:Q1-2004:Q3, namely we have 175 observations. Fig. 1 plots these series. The model selection criteria LR, AIC, HQ and SBC lead to selecting $p = 2$ for the log-levels of the bivariate processes. We reject the presence of cointegration at usual significance levels using Johansen's trace test. Consequently, the analysis is performed in first differences, namely with quarterly growth rates, with one lag only.

The estimation by OLS of the VAR(1) in first differences delivers (standard errors in brackets)

$$\begin{bmatrix} \Delta \ln USA_t \\ \Delta \ln CA_t \end{bmatrix} = \begin{bmatrix} 0.003 \\ (0.001) \\ 0.003 \\ (0.001) \end{bmatrix} + \begin{bmatrix} 0.333 & 0.273 \\ (0.088) & (0.079) \\ 0.265 & 0.360 \\ (0.102) & (0.092) \end{bmatrix} \begin{bmatrix} \Delta \ln USA_{t-1} \\ \Delta \ln CA_{t-1} \end{bmatrix}.$$

In order to check whether the coefficient matrix is of reduced rank we compute a SCCF test statistics using a canonical correlations approach (e.g. Vahid and Engle (1993)) between $\Delta z_t = (\Delta \ln USA_t : \Delta \ln CA_t)'$ and Δz_{t-1} . The results are as

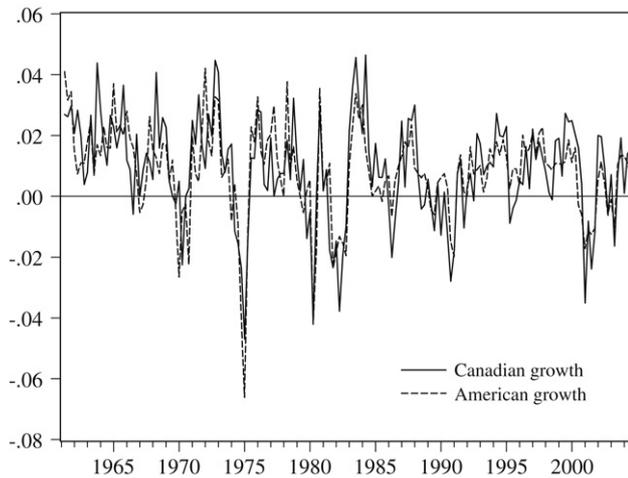


Fig. 1. Quarterly growth rates of industrial production indexes (industrial sector).

follows: p -value = 0.31 is the p -value associated with the null hypothesis ($\chi^2_{(1)}$) that a linear combination of Δz_t is orthogonal to the past of Δz_t . Information criteria also lead to select $s = 1$ (Hecq, 2006). The estimated common cyclical feature relationship is $\Delta \ln USA_t - 1.05 \Delta \ln CA_t$. Without any constraints on the short-run, both implied series should follow ARMA(2,1) processes. However, the sample ACF and PACF indicate (figures not reported in the paper) that the orders are probably shorter and that AR(1) models for industrial production in both countries are more appropriate. For univariate AR(1) models, the estimated equations have quite similar AR roots

$$\Delta \widehat{\ln USA}_t = 0.003 + 0.554 \Delta \ln USA_{t-1},$$

(0.001) (0.062)

$$\Delta \widehat{\ln CA}_t = 0.004 + 0.533 \Delta \ln CA_{t-1},$$

(0.001) (0.064)

where for both equations, the null of no disturbance autocorrelation is not rejected using LM tests.

One part of our story is confirmed, i.e., the implied individual ARMA models have smaller AR orders than those that we should get in absence of common feature restrictions and the two series exhibit one significant SCCF relationship. Moreover, the AR roots are similar for both countries. What might look as evidence against our prediction is the absence of a MA component because both growth rates should be ARMA(1,1).⁶ However, since δ is close to $(1 : -1)'$, the variances of VAR residuals are similar, and the correlation of VAR residuals is around 0.65, the results presented in Section 2.4 may explain why the MA(1) components are almost negligible.

3. Estimation procedures for implied univariate models and modelling considerations

As mentioned above, the finding that a set of series have identical autoregressive polynomials is an indication that these series have been generated by a VAR with non-(block)-diagonal and non-(block)-triangular matrix $\Phi(L)$ and that beyond having this common univariate autoregressive polynomial, they share other common features.⁷ This indication for the existence of

common features calls for testing whether the moving average part of the set of variables z_{it} with identical autoregressive polynomials exhibits the multivariate index structure implied by the presence of common features (see Section 2.4). A test of the multivariate index structure could be carried out using the residuals of $\det[\Phi(L)]z_{it}$. The properties of such a residual-based test would be affected by the properties of the estimator of the univariate autoregressive polynomials $\phi_i(L)$. It is therefore important to estimate these polynomials as accurately as possible.

We address the problem of estimation and testing of univariate ARMA models implied by a VAR(p) model with cofeature restrictions. Univariate ARMA models can be estimated in a straightforward way by the maximum likelihood method, that is we identify and estimate by ML for each series individually the parsimonious empirical ARMA(p_i, q_i) such that

$$z_{it} = \hat{\alpha}_i + \sum_{j=1}^{p_i} \hat{\phi}_{ij} z_{it-j} + \sum_{k=1}^{q_i} \hat{\theta}_{ik} \hat{u}_{it-k} + \hat{u}_{it},$$

$$i = 1 \dots n, t = 1 \dots T, \quad (7)$$

where $\hat{\alpha}_i$, $\hat{\phi}_{ij}$ and $\hat{\theta}_{ik}$ are estimated scalar parameters for series i , $i \in \{1, 2, \dots, n\}$; p_i and q_i are the lag orders of the ARMA model for the i th series and they might empirically differ from series to series. This first stage of the analysis is helpful because we can get a first idea about a maximum AR order p as well as of the number of series that might share co-movements.⁸ In doing so, we obtain a rough indication, a sort of upper bound say, regarding the possible number of common feature vectors, with $\max(p_i) = (n-s)p$. Also, if for a series the AR order differs much from that for other series, then probably it does not share a common cycle with these other variables. Indeed, in absence of cancellation of common roots in the ARMA representations, the ϕ_{ij} 's should be the same for all z_{it} 's.

For sets of univariate ARMA models derived from a stationary multivariate ARMA model, Wallis (1977) has considered maximum likelihood estimation, whereas Palm and Zellner (1980) have considered both maximum likelihood and efficient two-step estimation. For (difference)-stationary series, the asymptotic properties of univariate and multivariate estimation and test procedures are standard and known. For integrated processes with or without cointegration present, when the unit root is not imposed, the asymptotic procedures generally are non-standard. Whether standard asymptotics hold or not, likelihood and efficient two-step methods for systems may be cumbersome to implement when the set of variables is large. Therefore, for empirical work, it will be useful to have tools at the disposal which can be easily implemented and lead to reliable inference. Satisfying this need is the objective of this section.

3.1. Estimation of the common autoregressive polynomial using aggregates of individual series

It has been emphasised in Section 2 that the series implied by the VAR must have the same AR coefficients. Consequently, we investigate whether imposing this restriction using an estimator of the common AR part⁹ such that

$$z_{it} = \hat{\alpha}_i + \sum_{j=1}^{p_i} \hat{\phi}_{ij} z_{it-j} + \sum_{k=1}^{q_i} \hat{\theta}_{ik} \hat{u}_{it-k} + \hat{u}_{it}, \quad (8)$$

⁶ Using univariate maximum likelihood estimation, the MA(1) component has p -values of 0.16 and 0.32 for respectively the US and Canada.

⁷ Of course variables z_{it} with different univariate autoregressive polynomials could be generated by different sub-systems of a multivariate VAR with a corresponding block-diagonal matrix $\Phi(L)$.

⁸ Note that even though we do not identify the VAR order, we may nonetheless have an idea about p . For instance, with annual data p is typically at most 4 but $p = 2$ is the most common choice in empirical applications, whereas for quarterly data p is rarely found to be larger than 9.

⁹ Note that the common AR operator can be obtained by factorising the AR-polynomials for the individual series into the product of the individual operators and then cancelling common factors (Lütkepohl, 2005, 496).

where $\hat{\phi}_j$ is the j th lag order common coefficient to all n that should be preferred to estimating ARMA models for the individual series. In order to avoid new notation, we use the same symbols as in (7) for the estimated intercept and moving average parameters.

The intuition underlying the use of (8) is twofold. First, under the null hypothesis, this is a correct way to impose the commonality observed in different series. Secondly, under the alternative namely when including a set of n_1 variables that does not belong to the same multivariate process then, erroneously imposing a common AR root for the $n - n_1$ variables induces autocorrelation in the error terms. As a result, the detected order of the MA part can be higher than those observed for individual components. But before tackling this issue, we need to get estimates of the common AR component.

After having identified and estimated the parsimonious empirical ARMA(p_i, q_i) by ML for each series individually, one solution is to compute the average of n AR parameters such that $\hat{\phi}_j^{mg} = n^{-1} \sum_{i=1}^n \hat{\phi}_{ij}, j = 1, \dots, \max(p_i)$. However proceeding this way the homogeneity we have under the null is ignored. Moreover the estimation of n equations by ML induces much variability in the estimation of that average.

Instead, because the implied model for $\det[\Phi(L)]\bar{z}_t$ is at most a MA[($n - s$) p] model in the presence of SCCF vectors for instance, we further use this observation and estimate an ARMA model for the average of the n series in the ML estimation of

$$\bar{z}_t = \hat{\alpha} - \sum_{j=1}^p \hat{\phi}_j^{av} \bar{z}_{t-j} + \sum_{k=1}^q \hat{\theta}_k \hat{v}_{t-k} + \hat{v}_t, \tag{9}$$

with $\bar{z}_t = n^{-1} \sum_{i=1}^n z_{it}$ the simple average of n series, v_t being the innovation of the univariate ARMA(p, q) for \bar{z}_t and θ_k being the k -th lag parameter of the moving average part of \bar{z}_t . In this second setting, we might use a LM test for no autocorrelation or graphical tools (i.e. ACF, PACF) on \hat{v}_t to check the white noise hypothesis. A rejection of that null is a sign of misspecification, namely that we have likely included in the analysis a series that is not implied by the same system and consequently does not have the same final equation or does not have features in common with the other series in the system and therefore has higher degree AR and MA polynomials.

Cubadda et al. (2008) have shown by simulations that estimating a parsimonious ARMA model on aggregates is the preferred strategy for obtaining a common AR coefficient. Indeed, this procedure not only imposes the common AR parameters but also reduces the MA parts that might be annihilated by linearly combining the series (see Section 3). For instance, using the averages for Canada and the US with $\Delta \bar{z}_t = \frac{1}{2}(\Delta \ln USA_t + \Delta \ln CA_t)$, the best model (using SBC) for the aggregates is the AR(1)

$$\Delta \bar{z}_t = \underset{(3.55)}{0.016} + \underset{(10.40)}{0.617} \Delta \bar{z}_{t-1} + \hat{\epsilon}_t,$$

where t -ratios are in brackets. LM autocorrelation tests on the residuals do not reject the null at any common significance level.

Moreover, aggregation can also yield an additional “virtuous” property for such models. Indeed, cross-sectional aggregation implies a left multiplication of the adjoint by a vector, and this operation can reduce, in some circumstances, the order of the MA component of the aggregate. From Eq. (3), we see that the ARMA model for the aggregate is the following

$$\det[\Phi(L)]\bar{z}_t = \kappa' \Phi(L)^{adj} \epsilon_t,$$

where $\kappa = \frac{1}{n} \iota$, and ι is an n vector of ones. Moreover, we know from Proposition 8 that the coefficient matrices of $\Phi(L)^{adj}$ associated with degrees strictly larger than $(n - s - 1)p$ have a reduced-rank structure with a common right null space that is spanned by δ_\perp . Hence, it is possible that the polynomial matrix $\kappa' \Phi(L)^{adj}$ is, for

instance, of order $(n - s)p - 1$ instead of $(n - s)p$. This would require that the vector κ lies in the left null space of the coefficient matrix of $\Phi(L)^{adj}$ associated with degree $(n - s)p$, which is equal to $\Psi_{p\perp} G \delta'$, where G is a full rank $s \times s$ matrix, because of the relation

$$\underbrace{\det[\Phi(L)]}_{(n-s)p} I_n = \Phi(L) \Phi(L)^{adj} = \left(I_n - \sum_{j=1}^p \delta_\perp \Psi_j' L^j \right) \underbrace{\Phi(L)^{adj}}_{(n-s)p}.$$

By similar reasoning, the order of the MA component of model (9) can be reduced till $(n - s - 1)p$. This result explains why in empirical work it might occur that the MA order of the aggregate is even smaller than those of the individual series. Thus, we may interpret this occurrence as an evidence of the presence of co-movements, although the opposite conclusion is not valid.

3.2. A Monte Carlo experiment on the estimator using aggregates

Using the cross-sectional average is fine under the null that the FEs are from the same initial VAR model. In this subsection, we evaluate by simulation whether this is also true when we “erroneously substitute” to the group of variables having common cycles, additional series from another group. Next, we look at alternative weighted cross-sectional averages of series having different standard deviations.

In both situations, we look at the FEs of a VAR(1) with a reduced rank structure

$$z_t = \alpha + \Phi_1 z_{t-1} + \epsilon_t = \alpha + \delta_\perp C_1' z_{t-1} + \epsilon_t. \tag{10}$$

In order to keep fixed the AR and the MA orders in the n implied models, we follow Cubadda et al. (2008) and impose that $s = n - 1$ SCCF relationships, leading to n implied ARMA(1, 1) models for any n . The associated common feature matrix assumes full short-run convergence between economies (or variables). Indeed, one of the issues we want to address is to select a core of countries (or variables) with the maximum number of co-movements. The cofeature matrix has thus the following shape, for instance for $n = 5$,

$$\delta = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ -1 & -1 & -1 & -1 \end{pmatrix},$$

which leads to (up to a normalization) $\delta_\perp = (1 : 1 : 1 : 1 : 1)'$. α is a n dimensional vector of intercepts that we generate from a uniform distribution on $(0, 1)$. In order to have the same AR coefficient whatever the odd number of individuals we choose $C_1' = (0.5 : -0.5 : 0.5 : -0.5 : \dots : -0.5, 0.5)$. This guarantees both the stationarity of the multivariate process and the common value of the AR root ($\phi_{i1} = 0.5, i = 1, \dots, n$) for every series for any n . To show this we can simply use the property of partitioned matrices and compute $\det[I - \Phi_1 L]$ when considering the $(n - 1) \times (n - 1)$ upper-left block of $(I - \Phi_1 L)$. We use $M = 2000$ replications of $T + 50$ observations before dropping the first 50 points to initialize the random sequence.

3.2.1. Experiment 1: Impact of adding series not generated from a VAR

In this first experiment, we successively consider $n = 5, 11, 23$ series generated by the VAR (10) for $T = 50, 100, 250$. However, for the last $n_2 = (n - 1)/2$ series, we substitute series generated by the following diagonal VAR:

$$y_t = \Gamma_1 y_t + \epsilon_t, \quad t = 1, \dots, T, \tag{11}$$

where $\Gamma_1 = \text{diag}(\gamma) + U(-0.125, 0.125)$, and $\gamma = 0, 0.25, 0.5$. These are heterogeneous AR(1) processes such that only when $\gamma =$

Table 3Estimation of a common autoregressive parameter $\phi_{11} = 0.5$ in a mixed system.

n_1	n_2	γ	Bias			RMSE		
			$T = 50$	$T = 100$	$T = 250$	$T = 50$	$T = 100$	$T = 250$
3	2	0.50	-0.074	-0.034	-0.010	0.196	0.124	0.068
		0.25	-0.100	-0.060	-0.036	0.214	0.139	0.078
		0.00	-0.122	-0.078	-0.056	0.225	0.152	0.093
6	5	0.50	-0.088	-0.048	-0.032	0.207	0.124	0.075
		0.25	-0.098	-0.058	-0.042	0.213	0.131	0.082
		0.00	-0.106	-0.067	-0.048	0.219	0.137	0.087
12	11	0.50	-0.080	-0.041	-0.015	0.207	0.137	0.067
		0.25	-0.083	-0.043	-0.018	0.212	0.136	0.068
		0.00	-0.085	-0.044	-0.020	0.213	0.137	0.069

Note: A parsimonious empirical ARMA is chosen for the aggregates using SBC. The DGP is such that n_1 series are ARMA(1, 1) with a common AR parameter equal to 0.5 and n_2 series are AR(1) with a common coefficient equal to $\gamma + U(-0.125, 0.125)$.

Table 4

Various estimators of a common autoregressive parameter.

n		Bias			RMSE		
		$T = 50$	$T = 100$	$T = 250$	$T = 50$	$T = 100$	$T = 250$
5	Av_1	-0.072	-0.028	-0.007	0.205	0.128	0.069
	Av_2	-0.054	-0.012	0.011	0.199	0.127	0.069
	Av_3	-0.046	-0.006	0.018	0.190	0.128	0.073
	Av_4	-0.040	-0.001	0.019	0.183	0.122	0.073
11	Av_1	-0.065	-0.033	-0.011	0.185	0.120	0.068
	Av_2	-0.054	-0.022	-0.001	0.182	0.115	0.067
	Av_3	-0.051	-0.017	0.003	0.186	0.115	0.067
	Av_4	-0.046	-0.015	0.005	0.177	0.113	0.067
23	Av_1	-0.079	-0.033	-0.010	0.211	0.124	0.067
	Av_2	-0.077	-0.029	-0.006	0.216	0.123	0.066
	Av_3	-0.073	-0.027	-0.003	0.214	0.122	0.066
	Av_4	-0.067	-0.025	-0.003	0.201	0.121	0.065

Note: A parsimonious empirical ARMA model is chosen for the aggregates using SBC.

0.5 the coefficient comes from an uniform distribution centered around the same AR parameter of the FEs associated with (10). For the covariance matrix of the VAR errors we use in this first experiment $\Sigma_\varepsilon = I$. This implies cross-correlated errors $w_t = \Phi(L)^{adj} \varepsilon_t$ in the FEs because their contemporaneous covariance matrix is $E(w_t w_t') = I + \Phi_1 \Phi_1'$. This corresponds to a correlation between the disturbances of the implied equations of $\rho = 0.55$ for all pairs.

We estimate the common AR parameter using the parsimonious ARMA model on aggregates. Indeed, Cubadda et al. (2008) have found that this estimator has the best properties among various methods when the DGP is given by (10). Table 3 reports both the empirical bias and RMSE of this estimator when we generate n_1 series from system (10) and n_2 dimensional series from system (11). It emerges that the bias significantly increases if we include series with AR coefficients more distant from the final equation common parameters. However, the differences in terms of both bias and RMSE decrease as n and T increase.

3.2.2. Experiment 2: Simple vs. weighted averages

In this second experiment, we extend the Monte Carlo analysis of Cubadda et al. (2008) to allow for heterogeneous variances of the individual series. In particular, we now use in DGP (10) a covariance matrix of the VAR Σ_ε such that its diagonal elements are $(1^2, 2^2, \dots, n^2)$ and the correlation is equal to 0.7 for any pair of errors. This implies a contemporaneous covariance matrix $E(w_t w_t') = \Sigma_\varepsilon + \Phi_1 \Sigma_\varepsilon \Phi_1'$. We use $M = 2000$ replications of $n = 5, 11, 23$ individuals with $T = 50, 100, 250$ observations after having discarded 50 points as starting values.

Since weighted averages are more appropriate than simple averages when series have different variances, we compare the performances of four types of aggregates:

1. Av_1 is a simple average of the series;
2. Av_2 is a weighted average where the weights are proportional to the inverse of the unconditional standard deviations of the series;
3. Av_3 is a weighted average where the weights are proportional to the inverse of the conditional standard deviations of the series σ_{u_i} obtained through the estimated ARMA(1,1) models;
4. Av_4 is a weighted average where the weights are proportional to the inverse of the standard deviations $\sqrt{\sigma_{u_i}^2 (1 + \theta_1^2)}$ of the MA(1) innovations of the series obtained through the estimated ARMA(1,1) models.

Table 4 reports the empirical bias and RMSE of the estimators of the common AR parameter for the above aggregates. It emerges that the estimator based on Av_4 has, overall, the best performance. The intuition behind this finding is that, since the MA component is noise in this estimation problem, the Av_4 weighted average gives more weight to less noisy series. However, the differences in terms of performance of the various estimators are not large, and they tend to disappear as both n and T become large.

4. Analysis of per capita real GDP in Latin America

We study the first differences of the log levels of the per capita real gross domestic product (namely the growth rates) of nine Latin American economies. We have used annual data from 1950 to 2002, i.e. 53 observations. The series are derived from the Total Economy Database.¹⁰ We have used, on purpose, the

¹⁰ University of Groningen and The Conference Board, GGDC Total Economy Database www.eco.rug.nl/ggdc. The variables are expressed in 1990 US dollars and converted at "Geary-Khamis" purchasing power parities.

Table 5
Identified ARMA(p, q) models.

Identified components:	AR	MA	Outliers
Brazil	1 (0.51)	0	1981
Chile	0	1 (0.33)	1982, 1975
Columbia	1 (0.39)	0	1999
Mexico	0	1 (0.44)	1995
Peru	0	1 (0.58)	1983
Venezuela	0	0	–
Argentina	0	1 (0.23)	–
Ecuador	0	0	1987, 1999
Guatemala	1 (0.57)	0	–

same series as in Hecq et al. (2006) because it emerged there that it was not obvious to determine a priori the variables to be included in a common trend, common cycle analysis. Table 5 reports the list of these countries as well as the univariate ARMA models and outliers we have identified using the software program Tramo/Seats implemented in Eviews 6 (see Gómez and Maravall (1996)). This illustrates the use of a well known alternative to ACF and PACF analyses with the advantage to determine the presence of aberrant values. We also report the estimates of the significant coefficients at a 10% level.

As far as short-run co-movements are concerned, we can first form a group of countries that includes Brazil, Columbia and Guatemala for which an AR(1) was found.¹¹ With these three countries, $p = 2$ for the log levels was sufficient to capture the dynamics using a VAR. Moreover we do not conclude to the existence of long-run cointegration relationships between the log of per capita real GDP using Johansen's trace tests. Consequently, the VAR(1) for first differences implies ARMA(3,2) models for growth rates. This is not compatible with the parsimonious AR(1) we identify from the data. We interpret this as a sign of the presence of short-run co-movements. Indeed, ARMA(1,1) and AR(1) are compatible with $s = 2$ short-run cofeature relationships. If we identify an ARMA model for the aggregate of the three countries using Tramo, the most parsimonious model is the AR(1) without outliers

$$\Delta \bar{z}_t^* = \underset{(3.00)}{0.64} + \underset{(5.21)}{0.597} \Delta \bar{z}_{t-1}^* + \hat{\epsilon}_t,$$

where $\Delta \bar{z}_t^*$ is the weighted average Av_3 with estimated weights, which is obviously equal to Av_4 in this case, for which the residuals $\hat{\epsilon}_t$ show no sign of autocorrelation.

In Table 5, we report the orders of the selected processes, the point estimates for the coefficients (between parentheses) and the years for which outliers have been detected and accounted for.

From these results, we conjecture that there exist strong short-run co-movements characterised by a unique common cycle for that group of three countries. Both the individual ARMA orders and the absence of any autocorrelation in the residuals of the aggregated estimated equation support this conclusion. To have an alternative look at this issue (it is possible in this case because n is small), we use SCCF test statistics for the number of cofeatures using a canonical correlation approach as well as the value of information criteria. We estimate the VAR after correcting for the outliers detected at the univariate level. The p -value associated with $s = 2$ is 0.2 and information criteria also select $s = 2$ (see Hecq et al. (2006) and Hecq (2006)).

¹¹ Note that when looking for a set of countries sharing common cycles, white noise processes do not affect the order of the univariate ARMA processes of the remaining variables. This is because having one more white noise series included leads to consider the degree $(n+1 - (s+1))p = (n-s)p$ for the implied AR and MA parts for instance. But this simplifies the search procedure.

5. Conclusion

This paper has derived the implications for the univariate processes of the presence of common factors or trends resulting from common cyclical features or cointegration in multivariate time series. Exploiting the information about common dynamics contained in individual series and exhibited in individual ARIMA estimations, a strategy has been proposed to study the presence of co-movements in large sets of series without the need to develop and jointly estimate a large complex multivariate model for these series. This strategy is based on the theoretical results for the orders of the FEs for VAR models with co-features. The strategy is shown to yield sensible results in an application involving GDP data for nine countries.

Many further developments are currently under investigation such as, inter alia, the development of panel unit root tests and the extension to seasonal models. Moreover, the small sample and asymptotic properties of the proposed methods must be more deeply evaluated.

The tools we introduce can be extended in several directions. For instance, they allow to study co-movements between variables or convergence among economies, forecast series, and build business cycle indices without requiring a full parametric system with many variables. The advantages of our approach are: (1) its feasibility when it is not possible to jointly analyse a complete system or when we prefer to work using a sub-system, (2) its usefulness to detect sets of variables that are likely to be generated by seemingly unrelated subsystems for subsets of variables that have some features in common, (3) the accuracy of forecasts, (4) the ease of its implementation when a large number of jointly dependent variables has to be studied in complex situations, (5) the potential empirical applications in many fields.

Finally, the insights obtained from the analyses of single series and subsets of series with identical AR parts, and from testing for short-run and long-run co-feature restrictions can be incorporated in the multivariate model for all series to be modelled.

References

- Ahn, S.K., 1997. Inference of vector autoregressive models with cointegration and scalar components. *Journal of the American Statistical Association* 92, 350–356.
- Bai, J., Ng, S., 2004. A PANIC attack on unit roots and cointegration. *Econometrica* 72, 1127–1177.
- Brüggemann, R., Krolzig, H.M., Lütkepohl, H., 2003. Comparison of model reduction methods for VAR processes, *Economics Papers #2003-W13*, Economics Group, Nuffield College, University of Oxford.
- Centoni, M., Cubadda, G., Hecq, A., 2007. Common shocks, common dynamics, and the international business cycle. *Economic Modelling* 24, 149–166.
- Cubadda, G., Hecq, A., 2001. On non-contemporaneous short-run comovements. *Economics Letters* 73, 389–397.
- Cubadda, G., 2007. A unifying framework for analyzing common cyclical features in cointegrated time series. *Computational Statistics and Data Analysis* 52, 896–906.
- Cubadda, G., Hecq, A., Palm, F.C., 2008. Macro-panels and reality. *Economics Letters* 99, 537–540.
- Engle, R.F., Kozicki, S., 1993. Testing for common features (with comments). *Journal of Business and Economic Statistics* 11, 369–395.
- Ericsson, N., 1993. Comment (on Engle and Kozicki (1993)). *Journal of Business and Economic Statistics* 11, 380–383.
- Forni, M., Hallin, M., Lippi, M., Reichlin, L., 2000. The generalized dynamic factor model: Identification and estimation. *Review of Economics and Statistics* 82, 540–554.
- Franses, P.H., 1998. *Time Series Models for Business and Economic Forecasting*. Cambridge University Press.
- Gómez, V., Maravall, A., 1996. Programs TRAMO (Time series regression with arima noise, missing observations, and outliers) and SEATS (Signal extraction in arima time series). Instructions for the User. Working Paper 9628, Research Department, Banco de España.
- Gouriéroux, Ch., Peaucelle, I., 1989. Detecting a long-run relationship. CEPREMAP Discussion Paper, 8902.
- Hecq, A., 2006. Cointegration and common cyclical features in VAR models: Comparing small sample performances of the 2-step and iterative approaches. University of Maastricht Research Memorandum.

- Hecq, A., Palm, F.C., Urbain, J.P., 2002. Separation, weak exogeneity and P-T decomposition in cointegrated VAR systems with common features. *Econometric Reviews* 21, 273–307.
- Hecq, A., Palm, F.C., Urbain, J.P., 2006. Common cyclical features analysis in VAR models with cointegration. *Journal of Econometrics* 132, 117–141.
- Johansen, S., 1996. *Likelihood-Based Inference in Cointegrated Vector Autoregressive Models*. Oxford University Press, Oxford.
- Lütkepohl, H., 2005. *New Introduction to Multiple Time Series Analysis*. Springer Verlag, Berlin.
- Maravall, A., Mathis, A., 1994. Encompassing univariate models in multivariate time series: A case study. *Journal of Econometrics* 61, 197–233.
- Palm, F.C., 1977. On univariate time series methods and simultaneous equation econometric models. *Journal of Econometrics* 5, 379–388.
- Palm, F.C., Zellner, A., 1980. Large-sample estimation and testing procedures for dynamic equation systems. *Journal of Econometrics* 12, 251–283.
- Quenouille, M.H., 1957. *The Analysis of Multiple Time-Series*. Hafner Publishing Co., New-York. 2nd ed. 1968.
- Reinsel, G., 1983. Some results on multivariate autoregressive index models. *Biometrika* 70, 145–156.
- Schleicher, Ch., 2007. Codependence in cointegrated autoregressive models. *Journal of Applied Econometrics* 22, 137–159.
- Stock, J.H., Watson, M.W., 2002. Macroeconomic forecasting using diffusion indexes. *Journal of Business and Economic Statistics* 20 (2), 147–162.
- Tiao, G.C., Tsay, R.S., 1989. Model specification in multivariate time series (with comments). *Journal of Royal Statistical Society, Series B* 51, 157–213.
- Vahid, F., Engle, R.F., 1993. Common trends and common cycles. *Journal of Applied Econometrics* 8, 341–360.
- Vahid, F., Engle, R.F., 1997. Codependent cycles. *Journal of Econometrics* 80, 199–221.
- Wallis, K.F., 1977. Multiple time series analysis and the final form of econometric models. *Econometrica* 45, 1481–1497.
- Zellner, A., Palm, F.C., 1974. Time series analysis and simultaneous equation econometric models. *Journal of Econometrics* 2, 17–54.
- Zellner, A., Palm, F.C., 1975. Time series and structural analysis of monetary models of the US economy. *Sanhya: The Indian Journal of Statistics, Series C* 37, 12–56.
- Zellner, A., Palm, F.C., 2004. *The Structural Econometric Time Series Analysis Approach*. Cambridge University Press.