

MODELLING COMOVEMENTS OF ECONOMIC TIME SERIES: A SELECTIVE SURVEY

M. Centoni, G. Cubadda

1. INTRODUCTION

A large body of literature in time series analysis is dedicated to modeling comovements amongst multiple economic variables. The word comovement has been used in different meaning; even the spelling is uncertain, in most cases it can be found co-movement. In the present paper, comovement can be defined as “move together”, that is as movement that several series have in common. Comovements are often studied using correlation (e.g. Forbes and Rigobon, 2002; Imbs, 2004; Baxter and Kouparitsas, 2005): some extension are, among others, dynamic correlation of Croux et al. (2001) and the concordance index of Harding and Pagan (2006). However, economic time series have many characteristics such as trends, cycles, seasonality, serial correlation, and so on. Then it is important to consider these properties when analyzing comovement. To this respect (Vahid and Engle, 1993, pag. 341):

comovement among time series indicates existence of common components.

A well known example of comovement is cointegration (Engle and Granger, 1987; Johansen, 1988; Stock and Watson, 1988): a group of series that possess stochastic trends is cointegrated when its element share some common stochastic trends. Due to the importance of cointegration for empirical research, most of the intuitions behind that were extended to the analysis of comovements among stationary series. Gouriéroux and Peaucelle (1988) first extended the concept of cointegration to stationary time series and called this codependence: a group of variables are codependent if there exists a linear combination of them that possesses shorter memory than individual series. Later, Vahid and Engle (1997) show that codependence implies common short-run components. Another example of linear combinations of stationary variables that have lower order dynamics than the individual series is the scalar component model (SCM) of Tiao and Tsay (1989), that studied this type of structure in a multivariate ARMA model. These ideas were further developed and formalized in a series of paper by Engle and Kozicki (1993) and Vahid and Engle (1993). Engle and Kozicki (1993) considered features that satisfy the following axioms:

1. The vector series X_t has (does not have) the feature if any non-singular linear transformation of X_t still has (does not have) such feature;
2. If two n -vector time series Y_{1t} and Y_{2t} do not have the feature then $(Y_{1t} + Y_{2t})$ does not have the feature;
3. If Y_t does not have the feature but X_t has the feature then $(Y_t + X_t)$ has the feature.

Then, any dynamic property of the data could be viewed as a special case of feature: for example, series with stochastic trends satisfies all axioms. As pointed out by Engle and Kozicki (1993) a linear combination of two series that both have the feature does not necessarily possess the feature. This is the most interesting case, and to this issue Engle and Kozicki gave particular attention by the following definition:

Definition 1: A feature, which is present in each of a set of series, is said to be common to those series when there exists a nonzero linear combination of these series that does not have the feature.

It follows that cointegration is a special case of common features. Nowadays, there is a huge collection of special cases of common features. A comprehensive, although still partial, list includes: serial correlation (Engle and Kozicki, 1993; Vahid and Engle, 1993), when a linear combination of serially correlated series is white noise; cointrending (Chapman and Ogaki, 1993), when a linear combination of trend-stationary time series no longer display deterministic trend; common volatility (Engle and Kozicki, 1993; Engle and Susmel, 1993), when a linear combination of ARCH series eliminates the ARCH effects; common seasonality (Engle and Hylleberg, 1996), when a linear combination of seasonally integrated series is nonseasonal; co-breaking (Hendry, 1996), when a set of series appear subject to structural breaks but a linear combination of variables does not display the breaks; codependent cycles (Vahid and Engle, 1997), when a linear combination of a group of variables has shorter memory than the individual series; nonlinearity (Anderson and Vahid, 1998), when the conditional expectation of each element of a vector time series is nonlinear w.r.t. the conditional vector but there exist a linear combination of them whose conditional expectation is linear w.r.t. the conditional vector; common seasonal cycle (Cubadda, 1999), when there exists a linear combination of seasonal differenced series which follows an MA process of low order; common panel structure (Hecq et al., 2000), when there is a linear combination of the variables in a panel data which is white noise for all individuals of the panel; nonlinear cointrending (Bierens, 2000), when a linear combination of the components of a set of stationary time series around nonlinear deterministic time trends is stationary around a linear trend or a constant; polynomial serial correlation (Cubadda and Hecq, 2001), when there exists a polynomial combination of serially correlated time series that is white noise; long-run pure variance (Engle and Marcucci, 2006), when the conditional variances of a collec-

tion of assets all depend upon a small number of variance factors; non-innovation (Paruolo, 2006), when a linear combination of a demeaned autocorrelated process is an innovation process, i.e. non predictable; weak serial correlation (Hecq et al., 2006), when a linear combination of serially correlated series adjusted for long-run effects is white noise; periodic correlation (Haldrup et al., 2007), when a linear combination of periodic series is not periodic.

Instead of analyzing separately different properties of a time series, it could be useful to develop a unifying framework in order to contemporaneously analyzing several properties. Cubadda (2007) made this with respect to different forms of common cyclical features, introducing the notion of weak polynomial serial correlation which encompasses most of the existing formulation.

From a statistical point of view, comovements imply a reduction to a more parsimonious structure such as common factor representation (see, *inter alia*, Gonzalo and Granger, 1995), that can be estimated by reduced rank regression techniques (Anderson, 1984, 1999) with considerable efficiency gain (Vahid and Issler, 2002). Let Y_t be an n -vector time series such that

$$Y_t = Bf_t + e_t$$

where $(n-s)$ common factors f_t have the feature and e_t does not have the feature. Consider a s -vector δ such that $\delta'B = 0$, then $\delta'Y_t$ does not have the feature.

Comovements among time series are often predicted by economic theory. For example, in King et al. (1988) the solution of their macro model implies that output, consumption and investment have a common trend and a common cycle. The common stochastic trend is generated by an integrated productivity shock, while the deviation of capital stock from its steady state value determines the transitional dynamics of output, consumption and investment. Other example is Campbell (1987) where the saving path implies that disposable income and consumption cointegrate (Issler and Vahid, 2001). However, it could happen that observed comovements are in contrast with the theoretical model, and some efforts are toward reconciling the model with the data: as an example, in the DSGE model of Justiniano et al. (2010), an extension of the well known model of Christiano et al. (2005), investment shocks are the driven forces of fluctuations over the business cycle. In this model, however, consumption typically falls after an investment shock. This finding is in contrast with the observed business cycle comovement between consumption and investment. Khan and Tsoukalas (2011) try to solve this “comovement problem” introducing some modifications to traditional DSGE model.

The large body of literature on comovement obviously forces one to contain the discussion to some arguments: see Urga (2007) for a recent survey. Then, the plan of the rest of the paper is as follows. After introducing the general concept in Section 2, Section 3 discusses the implications of the presence of common trends in I(1) processes. Section 4 presents as different types of common season-

ality as many seasonal patterns. Section 5 deals with alternative notions of common cycles. Section 6 takes into account the consequences of the presence of common features with respect to univariate representation of time series models. Finally, Section 7 concludes.

2. COMMON FEATURES IN ECONOMIC TIME SERIES

The main idea is that a small number of unobserved components possess a given feature and transmit it to a larger set of economic time series. It is then possible to combine such time series in order to cancel the influence of these unobserved components, thus removing the common feature from the data. Reduced-Rank Regression (RRR) (Anderson, 1984) is often the statistical solution to the inferential problem.

Let Y_t be a stationary n -vector time series such that

$$Y_t = \beta'X_t + \gamma'Z_t + \varepsilon_t$$

where X_t and Z_t are, respectively, vectors of k and m variables, $\beta \neq 0$, and ε_t are i.i.d. $N_n(0, \Omega)$ errors that are independent from both X_t and Z_t . Both X_t and Z_t may contain (linear functions of) lags of Y_t .

Let us assume that:

1. Variables X_t possess the feature of interest whereas variables Z_t don't.
2. There exists a $n \times s$ ($s < \min\{n, k\}$) full-rank matrix δ such that $\delta'Y_t$ do not possess the feature. Then, it is said that variables Y_t have s Common Features.

In view of the above assumptions we get

$$\delta'Y_t = \delta'\gamma'Z_t + \delta'\varepsilon_t$$

that is equivalent to

$$\delta'\beta = 0 \Leftrightarrow \beta' = \delta_{\perp}'\psi'$$

where δ_{\perp} is a $n \times (n-s)$ full-rank matrix such that $\delta'\delta_{\perp} = 0$, and ψ is a $k \times (n-s)$ matrix.

If $n > k$, the matrix β' has not full column-rank, then there exist $(n-k)$ "trivial" CF. Hence, let us also assume that $s < n \leq k$, so the matrix ψ has full-rank as well.

2.1 Common Features & Reduced Rank Regression

The cofeature matrix δ and the coefficient matrix ψ are usually obtained by means of RRR. The procedure can be summarized as follows:

1. Regress both Y_t and X_t on Z_t and obtain the partial regression model

$$y_t = \beta' x_t + \varepsilon_t$$

where $y_t = Y_t - E(Y_t | Z_t)$, $x_t = X_t - E(X_t | Z_t)$, $\beta = \Sigma_{xx}^{-1} \Sigma_{xy}$, and

$$\text{var} \begin{bmatrix} y_t \\ x_t \end{bmatrix} = \begin{bmatrix} \Sigma_{yy} & \Sigma_{yx} \\ \Sigma_{xy} & \Sigma_{xx} \end{bmatrix}$$

2. Solve the following maximization problem

$$\alpha_1 = \arg \max_{\alpha \in \mathfrak{R}^n} \left\{ \frac{\overbrace{\alpha' \Sigma_{yx} \Sigma_{xx}^{-1} \Sigma_{xy} \alpha}^{\beta' \Sigma_{xx} \beta}}{\alpha' \Sigma_{yy} \alpha} \right\} \quad \text{s.t. } \alpha' \Sigma_{yy} \alpha = 1$$

Then we get the solution $\alpha_1 = \Sigma_{yy}^{-1/2} v_1$, where v_j ($j=1, \dots, n$) is the eigenvector associated to the j -th largest eigenvalue λ_j ($\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_n$) of the symmetric, semi-positive definite matrix

$$\Sigma_{yy}^{-1/2} \Sigma_{yx} \Sigma_{xx}^{-1} \Sigma_{xy} \Sigma_{yy}^{-1/2}$$

Note that $R^2(\alpha' y_t | \gamma' x_t)$, where $\gamma = \beta \alpha$, is maximized for $\alpha = \alpha_1$ and $\gamma = \gamma_1 \equiv \Sigma_{xx}^{-1} \Sigma_{xy} \alpha_1$ since

$$\alpha_1' \Sigma_{yy} \alpha_1 = 1$$

$$\gamma_1' \Sigma_{xx} \gamma_1 = v_1' \Sigma_{yy}^{-1/2} \Sigma_{yx} \Sigma_{xx}^{-1} \Sigma_{xy} \Sigma_{yy}^{-1/2} v_1 = \lambda_1$$

3. Solve the following maximization problem

$$\alpha_2 = \arg \max_{\alpha \in \mathfrak{R}^n} \left\{ \frac{\alpha' \Sigma_{yx} \Sigma_{xx}^{-1} \Sigma_{xy} \alpha}{\alpha' \Sigma_{yy} \alpha} \right\} \quad \text{s.t. } \begin{aligned} \alpha' \Sigma_{yy} \alpha &= 1 \\ \alpha_1' \Sigma_{yy} \alpha &= 0 \end{aligned}$$

We get the solution $\alpha_2 = \Sigma_{yy}^{-1/2} v_2$

Note that $R^2(\alpha_2' y_t | \gamma_2' x_t) = \lambda_2$ where $\gamma_2 = \Sigma_{xx}^{-1} \Sigma_{xy} \alpha_2$. Moreover, $R^2(\alpha_1' y_t | \gamma_2' x_t) = 0$ since

$$\gamma_2' \Sigma_{xy} \alpha_1 = v_2' \Sigma_{yy}^{-1/2} \Sigma_{yx} \Sigma_{xx}^{-1} \Sigma_{xy} \Sigma_{yy}^{-1/2} v_1 = 0$$

and v_1 and v_2 are eigenvectors corresponding to different eigenvalues.

4. In general, we have that $\alpha_j = \Sigma_{yy}^{-1/2} \nu_j$ is the solution of:

$$\alpha_j = \arg \max_{\alpha \in \mathbb{R}^n} \left\{ \frac{\alpha' \Sigma_{yx} \Sigma_{xx}^{-1} \Sigma_{xy} \alpha}{\alpha' \Sigma_{yy} \alpha} \right\} \quad \text{s.t.} \quad \begin{aligned} \alpha_j' \Sigma_{yy} \alpha_j &= 1 \\ \alpha_j' \Sigma_{yy} \alpha_j &= 0 \end{aligned}$$

for $i \leq j = 2, \dots, n$. Moreover,

$$R^2(\alpha_i' y_t | \gamma_j' x_t) = \begin{cases} \lambda_j & \text{for } i = j \\ 0 & \text{for } i \neq j \end{cases},$$

where $\gamma_j = \Sigma_{xx}^{-1} \Sigma_{xy} \alpha_j$.

5. Since $\delta' \beta' = 0$ then so is $\delta' \beta' \Sigma_{xx} \beta \delta$, that implies that $\lambda_{n-s+1} = \dots = \lambda_n = 0$, then the cofeatures matrix δ is obtained (up to an identification matrix) as follows

$$\delta = (\alpha_{n-s+1}, \dots, \alpha_n)$$

6. Since $\lambda_1 \geq \dots \geq \lambda_{n-s} \geq 0$, the coefficient matrix ψ is obtained (up to an identification matrix) as follows

$$\psi = (\gamma_1, \dots, \gamma_{n-s})$$

7. Finally, the loading matrix δ_{\perp} is given by the regression coefficients of y_t on $\psi' x_t$.

It is worth remarking that the eigenvalue problem

$$\Sigma_{yy}^{-1/2} \Sigma_{yx} \Sigma_{xx}^{-1} \Sigma_{xy} \Sigma_{yy}^{-1/2} \nu = \nu \lambda \quad (1)$$

is equivalent to finding the roots of the equation:

$$\Sigma_{yy}^{-1} \Sigma_{yx} \Sigma_{xx}^{-1} \Sigma_{xy} \alpha = \alpha \lambda \quad (2)$$

with the normalization $\alpha' \Sigma_{yy} \alpha = 1$.

The solution of the above problem (equivalent to RRR) is known in statistics as partial canonical correlation analysis between Y_t and X_t conditional to Z_t , and it is denoted by

$$\text{CanCor}(Y_t, X_t | Z_t)$$

2.2 Common Features & Reduced Rank Regression: Statistical Inference

Under the assumption that $(y'_t, x'_t)'$ are normally distributed, RRR provides ML inference on δ and ψ (Anderson, 1984). The LR test statistic on the existence of s Common Features is:

$$LR(s) = -T \sum_{i=n-s+1}^n \ln(1 - \hat{\lambda}_i) \tag{3}$$

where $\hat{\lambda}_i$ is the i -th largest eigenvalue of the sample matrix $\hat{\Sigma}_{yy}^{-1/2} \hat{\Sigma}_{yx} \hat{\Sigma}_{xx}^{-1} \hat{\Sigma}_{xy} \hat{\Sigma}_{yy}^{-1/2}$. Under proper regularity conditions on the variables X_t and Z_t the test statistic (3) is distributed

$$LR(s) \xrightarrow{d} \chi^2[s(k - n + s)].$$

The ML estimator of the cofeatures matrix δ is obtained (up to an identification matrix) as follows

$$\hat{\delta} = (\hat{\alpha}_{n-s+1}, \dots, \hat{\alpha}_n),$$

where $\hat{\alpha}_j = \hat{\Sigma}_{yy}^{-1/2} \hat{v}_j$ and \hat{v}_j ($j=1, \dots, n$) is the eigenvector associated to the j -th largest sample eigenvalue $\hat{\lambda}_j$.

The ML estimator of the coefficient matrix ψ is obtained (up to an identification matrix) as follows

$$\hat{\psi} = (\hat{\gamma}_1, \dots, \hat{\gamma}_{n-s}),$$

where $\hat{\gamma}_j \equiv \hat{\Sigma}_{xx}^{-1} \hat{\Sigma}_{xy} \hat{\alpha}_j$.

The ML estimator of the loading matrix δ_{\perp} is finally obtained by the regression coefficients of y_t on $\hat{\psi}'x_t$.

The degrees of freedom of the test statistic (3) are obtained as follows. Under H_1 the matrix β is composed by $n \times k$ elements. Under H_0 we have that

$$\beta = \underbrace{\psi}_{k \times (n-s)} \left(\underbrace{\delta_{\perp,1}}_{(n-s) \times (n-s)} : \underbrace{\delta_{\perp,2}}_{(n-s) \times s} \right).$$

Without loss of generality, we assume that $\delta_{\perp,1}$ has full-rank. Then we can write

$$\beta = \psi \delta_{\perp,1} \delta_{\perp,1}^{-1} \delta'_{\perp} = \tilde{\psi} \tilde{\delta}'_{\perp}$$

where $\tilde{\delta}'_{\perp} = (I_s, \tilde{\delta}'_{\perp,2})$. Hence, the matrix β is composed by $(n-s) \times (s+k)$ elements. It follows that under $H_{0,s}(k-n+s)$ restrictions are imposed.

3. COMMON TRENDS

Macroeconomic time series often display a clear trending behavior, and standard unit-root tests suggest that series need to be treated as processes integrated of order 1 (I(1) for short) rather than trend stationary. However, linear combinations of such series do not show similar patterns, instead appears to be stationary: this is what is intended by cointegration. Cointegration introduced by Granger (1981) has become one of the most widely used concept in time series econometrics, due to the straightforward economic interpretation of the cointegrating relations as long-run equilibrium relationships. Comovement arises by the presence of common stochastic trends in the series, that are removed by the cointegrating relations. It is worth noting that if the non stationary behavior is due to presence of deterministic rather than stochastic trend, we refer to a stationary linear combination of these series as cotrending relations (Chapman and Ogaki, 1993) or nonlinear cotrending relations, if non stationarity is due to nonlinear deterministic time trends (Bierens, 2000). The cointegrated vector autoregressive (VAR) model pioneered by Johansen (1988, 1996) is the most commonly used econometric tool: see Juselius (2006) for the use of cointegrated VAR models in macroeconomic empirical research. In what follows we briefly review this model and describe the procedure to determine the number of cointegrating relations.

3.1 Cointegration: Assumptions & Definition

A n -vector of time series $\{y_t, t = 1, \dots, T\}$ is generated by the VAR(p) model:

$$A(L)y_t = \Phi D_t + \varepsilon_t \quad (4)$$

where $A(L) = I_n - \sum_{i=1}^p A_i L^i$, D_t is a vector of deterministic terms, and ε_t are i.i.d. $N_n(0, \Omega)$ errors. Assume that elements of y_t are, at most, I(1), denoted by $y_t \sim I(1)$. By expanding $A(L)$ on 0 and 1, we rewrite (4) as

$$\Gamma(L)\Delta y_t = \Phi D_t - A(1)y_{t-1} + \varepsilon_t \quad (5)$$

where $\Gamma(L) = I_n - \sum_{i=1}^{p-1} \Gamma_i L^i$, and $\Gamma_i = -\sum_{j=i+1}^p A_j$. Since $\Delta y_t \sim I(0)$, $A(1)y_t$ must be stationary as well.

Three possible cases can arise:

1. $A(1) = 0$, means that all the elements of y_t are I(1) and do not cointegrate, then a VAR in first differences is appropriate;

2. $\text{rank}[A(1)] = n$, then all the elements of y_t are $I(0)$;
3. $\text{rank}[A(1)] < n$, then $A(1) = -\alpha\beta'$, where α and β are full-rank $n \times r$ -matrices $r = 1, \dots, n-1$ and it follows $\beta' y_t \sim I(0)$. Series y_t are cointegrated of order $(1,1)$, $y_t \sim CI(1,1)$, the space spanned by β is the cointegration space.

If variables cointegrate, we can rewrite the model (5) in his Vector Error Correction Model (VECM) representation:

$$\Gamma(L)\Delta y_t = \Phi D_t + \alpha\beta' y_{t-1} + \varepsilon_t \tag{6}$$

where series Δy_t react to previous stationary deviations from the equilibria. Series y_t admit the following Wold representation:

$$\Delta y_t = \Theta D_t + C(L)\varepsilon_t \tag{7}$$

where $\Theta D_t = C(L)\Phi D_t$, and $C(L) = I_n + \sum_{i=1}^{\infty} C_i L^i$ is such that $\sum_{j=1}^{\infty} j |C_j| < \infty$.

By expanding $C(L)$ on 1 and integrating both sides of the above equation, we get the multivariate Beveridge-Nelson representation (Beveridge and Nelson, 1981):

$$y_t = y_0 + \tilde{\Phi} \tilde{D}_t + \sum_{j=1}^t C(1)\varepsilon_j + \tilde{C}(L)\varepsilon_t, \tag{8}$$

where $\Delta \tilde{\Theta} \tilde{D}_t = \Theta D_t$, $\tilde{C}(L) = \sum_{i=0}^{\infty} \tilde{C}_i L^i$, and $\tilde{C}_i = -\sum_{j>i} C_j$ for all i .

Since $C(1) = \beta_{\perp}(\alpha'_{\perp} \Gamma \beta_{\perp})^{-1} \alpha'_{\perp}$, series y_t share the following $(n-r)$ common stochastic trends:

$$\tau_t = \sum_{j=1}^t \alpha'_{\perp} \varepsilon_j.$$

Note the analogy between (8) and the common factor representation of the introduction: $\tilde{C}(L)\varepsilon_t$ does not have the feature, while τ_t has the feature, and the $(n-r)$ common stochastic trends are removed by the cointegrating vectors, since $\beta' \beta_{\perp} (\alpha'_{\perp} \Gamma \beta_{\perp})^{-1} = 0$. For alternative identifications of the common trends see Gonzalo and Granger (1995) and Johansen (1996) among others.

3.2 Cointegration: ML Inference

In order to determine the number of cointegrating relations and consistently estimate parameters of the VECM (6), Johansen (1988, 1991, 1996) proposed to resort to RRR in a Gaussian ML framework. In particular, the number of cointegrating relations is based on the solution of the following problem:

$$\text{CanCor} \left\{ y_{t-1}, \Delta y_t \begin{pmatrix} D_t \\ \Delta y_{t-1} \\ \dots \\ \Delta y_{t-p+1} \end{pmatrix} \right\},$$

and then conducts a LR test for $H_0 \text{ rank}[A(1)]=r$ vs $H_1 \text{ rank}[A(1)]=n$ using the test statistic:

$$-T \sum_{j=r+1}^n \ln(1 - \hat{\lambda}_j), \quad r = 1, \dots, n,$$

where $\hat{\lambda}_j$ is the j -th largest squared canonical correlation. If the null hypothesis is not rejected, the ML estimator of β are the r eigenvectors associated with the r largest eigenvalues $\hat{\lambda}_1, \dots, \hat{\lambda}_r$.

Note that when D_t is a constant, the test statistic $Q(r|n)$ converges weakly in distribution to

$$\text{tr} \left\{ \int_0^1 d B(u) F'(u) \left[\int_0^1 F(u) F'(u) du \right]^{-1} \int_0^1 F(u) dB'(u) \right\}$$

where $\text{tr}\{\cdot\}$ denotes trace, $B(u)$ is the standard Brownian motion of dimension $(n-r)$, and $F(u) = B(u) - \int_0^1 B(v) dv$.

Moreover, when D_t is not constant (linear trend, step dummy, ...) in the estimated model, $F(u)$ has a different formula. However, the limit distribution of the test statistics does not depend on Φ (Johansen, 1991).

The ML estimator of β , $\hat{\beta}_{ML}$, is T -consistent and its limit distribution is mixed-Gaussian, so tests on β are asymptotically χ^2 but the exact distribution of (normalized) $\hat{\beta}_{ML}$ has Cauchy-like tails (Phillips, 1994).

Finally, the remaining parameters in VECM (6) are estimated by OLS after fixing the matrix β to its estimated value, see Johansen (1996) for further details.

4. COMMON STOCHASTIC SEASONALITY

Most of economic time series display a seasonal pattern that evolves in a stochastic fashion (Hylleberg et al., 1993). Focusing, for simplicity, on quarterly time series, variables y_t are seasonally integrated of order 1 (Hylleberg et al., 1990) if $\Delta_4 y_t$ is stationary, where $\Delta_4 = (1 - L^4)$. Notice that

$$\Delta_4 = \underbrace{(1-L)}_{\text{frequency } 0} \underbrace{(1+L)}_{\text{frequency } \pi} \underbrace{(1+L^2)}_{\text{frequency } \pi/2},$$

which allows for three diverse kinds of cointegration relations:

1. Cointegration at frequency 0, if a linear combination of $y_{1,t} \equiv (1+L)(1+L^2)y_t$ is stationary.
2. Cointegration at frequency π , if a linear combination of $y_{2,t} \equiv (1-L)(1+L^2)y_t$ is stationary.
3. Polynomial cointegration at frequency $\pi/2$, if a linear combination of $y_{3,t} \equiv (1-L)(1+L)y_t$ and its lags is stationary.

4.1 Seasonal Cointegration: Assumptions & Definition

Let $\{y_t, t = 1, \dots, T\}$ be an n -vector of time series generated by the following VAR(p) model:

$$\Pi(L)y_t = \Phi D_t + \varepsilon_t \quad (9)$$

where $\Pi(L)$ is a p -order polynomial matrix, ε_t are i.i.d. $N_n(0, \Omega)$, and D_t is a deterministic kernel.

Assume that elements of y_t are, at most, I(1) at frequencies 0, π , and $\pi/2$. By expanding $\Pi(L)$ on 0, ± 1 and $\pm i$, we rewrite the VAR (9) as

$$\tilde{\Gamma}(L)\Delta_4 y_t = \Phi D_t + \Pi_1 y_{1,t-1} + \Pi_2 y_{2,t-1} + \Pi_* y_{*,t-1} + \bar{\Pi}_* \bar{y}_{*,t-1} + \varepsilon_t$$

where $\tilde{\Gamma}(L) = I_n - \sum_{j=1}^{p-4} \Gamma_j L^j$, $\tilde{\Gamma}_i = -\sum_{l=1}^{\lfloor (p-j)/4 \rfloor} \Pi_{j+4l}$, $\Pi_1 = -\frac{1}{4}\Pi(1)$, $\Pi_2 = \frac{1}{4}\Pi(-1)$, $\Pi_* = -\frac{1}{4}\Pi(i)$ and \bar{C} denotes the complex conjugate of a complex matrix C .

Since $\Delta_4 y_t \sim I(0)$, then $\Pi_1 y_{1,t}$, $\Pi_2 y_{2,t}$, and $\Pi_* y_{*,t}$ must be stationary as well.

If the series are cointegrated of order (1,1) at frequencies 0, π , and $\pi/2$, the VAR(p) model may be rewritten in the following Complex VECM (CVECM)

$$\tilde{\Gamma}(L)\Delta_4 y_t = \Phi D_t + \alpha_1 \beta'_1 y_{1,t-1} + \alpha_2 \beta'_2 y_{2,t-1} + \alpha_* \bar{\beta}'_* y_{*,t-1} + \bar{\alpha}_* \beta'_* \bar{y}_{*,t-1} + \varepsilon_t$$

where α_j and β_j are $n \times r_j$ -matrices with rank equal to r_j for $j=1,2$, and α_* and β_* are complex $n \times r_3$ -matrices with rank equal to r_3 , see Cubadda (2001) and Cubadda and Omtzigt (2005).

Notice that four cointegrating relationships are present in the CVECM. Indeed, β_1 and β_2 are, respectively, the cointegration matrices at frequencies 0 and π , whereas the conjugate complex cointegration matrices $-\bar{\beta}_*$ and β_* are respectively associated with frequencies $\frac{\pi}{2}$ and $\frac{3\pi}{2}$.

Since complex valued cointegration vectors are not amenable to economic interpretation, we rewrite the complex VECM in the following Seasonal VECM (SVECM)

$$\tilde{\Gamma}(L)\Delta_4 y_t = \Phi D_t + \alpha_1 \beta'_1 y_{1,t-1} + \alpha_2 \beta'_2 y_{2,t-1} + (\alpha_4 - \alpha_3 L)(\beta'_3 - \beta'_4 L) y_{3,t-1} + \varepsilon_t$$

where $\alpha_* \bar{\beta}'_* \equiv \frac{1}{2}(\alpha_3 + \alpha_4 i)(\beta'_3 - \beta'_4 i)$, see Lee (1992) and Ahn and Reinsel (1994).

It is worth remarking that the SVECM exhibits a real but polynomial cointegration vector $\beta(L) = \beta_3 - \beta_4 L$ for the annual frequency. Further, the CVECM is more convenient than the SVECM for statistical inference because the cointegration restrictions at the various frequencies imply the usual reduced-rank structure of the Π 's matrices.

4.2 Seasonal Cointegration: ML Inference

In order to determine the number of seasonal cointegrating relations, Hylleberg et al. (1990) suggest a two-step procedure after having appropriately filtered the series. Since this procedure need pretesting for seasonal root, Cubadda (2001) proposed a more general procedure that do not require any prior knowledge about seasonal integration and allows to test for seasonal cointegration at any collection of frequencies. The procedure can be summarized as follows. Regress $\Delta_4 y_t$, $y_{1,t-1}$, $y_{2,t-1}$, and $y_{*,t-1}$ on $(D_t, \Delta_4 y_{t-1}, \dots, \Delta_4 y_{t-p+4})$ and take respectively the residuals $R_{0,t}$, $R_{1,t}$, $R_{2,t}$, and $R_{*,t}$. These residuals asymptotically satisfy the following equation

$$R_{0,t} = \alpha_1 \beta'_1 R_{1,t-1} + \alpha_2 \beta'_2 R_{2,t-1} + \alpha_* \bar{\beta}'_* R_{*,t} + \bar{\alpha}_* \beta'_* \bar{R}_{*,t} + \varepsilon_t$$

Since the processes $R_{1,t}$, $R_{2,t}$, $R_{*,t}$, and $\bar{R}_{*,t}$ are mutually asymptotically uncorrelated, we can ignore reduced rank restrictions at the frequencies different from the one of interest.

Hence, the (concentrated) ML solution is obtained by solving

$$\text{CanCor}\{R_{0,t}, R_{1,t} | R_{2,t}, R_{*,t}, \bar{R}_{*,t}\}$$

for cointegration analysis at frequency 0, and

$$\text{CanCor}\{R_{0,t}, R_{2,t} \mid R_{1,t}, R_{*,t}, \bar{R}_{*,t}\}$$

for cointegration analysis at frequency π .

Johansen and Schaumburg (1998) showed that the (concentrated) ML solution for cointegration analysis at frequencies $\frac{\pi}{2}$ and $\frac{3\pi}{2}$ cannot be obtained by RRR.

However, the solution of

$$\text{CanCor}\{R_{0,t}, \bar{R}_{*,t} \mid R_{1,t}, R_{2,t}, R_{*,t}\}$$

provides asymptotically optimal inference for cointegration at frequency $\frac{\pi}{2}$ (Cubadda, 2001).

It is worth remarking that small sample improvements may be obtained by estimating the cointegration vectors jointly at the zero and seasonal frequencies (Cubadda and Omtzigt, 2005).

5. COMMON CYCLES

As we have seen in Section 3, a set of I(1) cointegrated time series share some common stochastic trends. However, detrended economic time series often display clear evidence of comovements (Lucas, 1977), which cannot be due to cointegration, thus suggesting the presence of common cycles. If this is the case, we expect that there exists linear combination of cyclical series that are not cyclical. This *Common Cyclical Features* [CCF] (Engle and Kozicki, 1993; Vahid and Engle, 1993) can be interpreted as short-run equilibrium relationships, similarly to the interpretation of the cointegrating relations as long-run equilibrium. Indeed, common cycles appear in many theoretical economic models. For examples, in models of aggregate consumption behavior as discussed in Vahid and Engle (1993) and Issler and Vahid (2001), the proportion of total income that accrues to “myopic” individuals (Campbell and Mankiw, 1990), that is, individuals that consume their income entirely in every period, or excess sensitivity of consumption to current income (Flavin, 1993), provide that first differences of I(1) consumption and income share a common cycle. Note that in the consumption theory of Hall (1978) consumption and income share only a common stochastic trend. In real business cycle models as in King et al. (1991) consumption, investment and income have a common cycle due to the fact that transitional dynamics of the system is a function of the deviation of the capital stock from its stationary value: see Issler and Vahid (2001). As another example, in the real business cycle model for sectoral output *à la* Long and Plosser (1983), as in Engle and Issler (1995), common cycles depend on the propagation mechanisms through the restrictions

on the production function, i.e. technological constraints. Hence, CCF implies restrictions on the VAR parameters that have meaningful implications for the short-run components of an I(1) vector time series, that can be estimated by the RRR techniques of Section 2.

5.1 *Alternative notions of Common Cyclical Features*

Let us start with the multivariate Beveridge and Nelson (1981, BN) cycles of series y_t :

$$\kappa_t = \tilde{C}(L)\varepsilon_t,$$

where $\tau_t = C(1)\sum_0^{t-1}\varepsilon_{t-1}$, and $\kappa_t = \tilde{C}(L)\varepsilon_t$, and $\tilde{C}_i = -\sum_{j>i}C_j$.

As showed in Section 3, from the Granger representation theorem (Engle and Granger, 1987) we know that the presence of cointegration is equivalent to the existence of $(n-r)$ common stochastic trends since $\beta'\tau_t = 0$. The analysis of CCF's is instead concerned with reduced-rank restrictions on the VECM parameters that have interesting implications on the cycles κ_t .

However, differently from cointegration, there is not a unique notion of common short-run components. Indeed, also the degree of synchronicity of the common cycle plays a role in the definitions. Alternative notions of CCF impose differing reduced-rank structures to the VAR. Let us briefly review various form of CCF starting with the seminal notion proposed by Engle and Kozicki (1993).

5.1.1 *Serial Correlation Common Feature [SCCF]*

Series Δy_t have s ($s < n$) SCCF's iff there exists an $n \times s$ matrix δ_s with full column rank such that the VECM (6) can be rewritten into the following RRR model

$$\Delta y_t = \Phi D_t + \delta_{s\perp}' \psi_s' (y_{t-1}'\beta, \Delta y_{t-1}', \dots, \Delta y_{t-p+1}')' + \varepsilon_t \quad (10)$$

where ψ_s is an $(np-n+r) \times (n-s)$ matrix with full column rank (Engle and Kozicki, 1993). Since $\delta_s'\Delta y_t = \delta_s'\Phi D_t + \delta_s'\varepsilon_t$, there exist s linear combinations of series Δy_t that are unpredictable from the past, i.e. they are innovation processes. Since $\delta_s'\kappa_t = 0$, there exist $(n-s)$ common cycles in series y_t (Vahid and Engle, 1993) that are perfectly synchronized, as a results the impulse responses functions are exactly collinear. However, one observes that due to technical reasons (e.g., seasonal adjustment) as well as economic reasons (e.g., adjustment costs, labor market rigidities), the hypothesis underlying SCCF is sometimes too strong, since SCCF is not able to detect the existence of non-contemporaneous cyclical comovements (Ericsson, 1993). Then, some less restrictive variants of the SCCF have been introduced in the literature, which we will review in turn.

5.1.2 Polynomial Serial Correlation Common Feature [PSCCF]

Cubadda and Hecq (2001) propose the notion of polynomial serial correlation common features as a measure of non-contemporaneous cyclical comovements. Non-synchronous common cycles arises, for example, in economic model of consumption with several types of consumer goods as, among others, in Vahid and Engle (1997) and Schleicher (2007): it is shown that the maximization problem of the representative agent implies a unsynchronized common cycle (codependent cycle) in the consumer goods vector.

By definition, series Δy_t have s PSCCF's iff there exists an $n \times s$ matrix δ_p with full column rank such that $\delta_p' \Gamma_1 \neq 0$, and the VECM (6) can be rewritten into the following partial RRR model

$$\Delta y_t = \Phi D_t + \Gamma_1 \Delta y_{t-1} + \delta_{p\perp} \psi_p' (\Delta y_{t-2}, \dots, \Delta y_{t-p+1}, y_{t-1}' \beta)' + \varepsilon_t, \quad (11)$$

where ψ_p is an $(np-2n+r) \times (n-s)$ matrix with full column rank (Cubadda and Hecq, 2001).

In order to interpret the notion of PSCCF, Cubadda and Hecq (2001) show that there exists a first-order polynomial matrix $\delta(L) = \delta_p - \Gamma' \delta_p L$ such that

$$\delta(L)' \Delta y_t = \delta_p' \Phi D_t + \delta_p' \varepsilon_t.$$

Hence, PSCCF requires that there exists a first-order polynomial matrix $\delta(L)$ such that $\delta(L)' \Delta y_t$ is white noise. The presence of PSCCF has an interesting implication for the BN cycles of series y_t : indeed, since $\delta(L)' \kappa_t = \delta_p' C(1) \varepsilon_t$, the same PSCCF relationships cancel the dependence from the past of both the first differences and cycles of series y_t .

5.1.3 Weak Form of Serial Correlation Common Feature [WF]

In the above definitions of CCF, the number of SCCFs or PSCCFs, s , cannot exceed the number of common trends $(n-r)$. In order to remove this restriction, Hecq et al. (2006) proposed the notion of weak form of SCCF: series Δy_t have s WF's iff there exists an $n \times s$ matrix δ_W with full column rank such that $\delta_W' \alpha \neq 0$, and the VECM in (6) can be rewritten into the following partial RRR model

$$\Delta y_t = \Phi D_t + \alpha \beta' y_{t-1} + \delta_{W\perp} \psi_W' (\Delta y_{t-1}, \dots, \Delta y_{t-p+1})' + \varepsilon_t$$

where ψ_W is an $(np-n) \times (n-s)$ matrix with full column rank (Hecq et al., 2006).

In order to uncover interesting implications of the WF for the BN cycle, Cubadda (2007) shows that there exists a first-order polynomial matrix $\delta_W(L) \equiv \delta_W - (\beta \alpha' + I_n) \delta_W L$ such that

$$\delta_W(L)' y_t = \delta_W' \Phi D_t + \delta_W' \varepsilon_t.$$

As consequence, since $\delta_W(L)' \kappa_t = \delta_W'(I_n - C(1)) \varepsilon_t$, the same WF relationships cancel the dependence from the past of both the cycles and linearly detrended levels of series \mathcal{Y}_t .

5.1.4 Weak Form of Polynomial Serial Correlation Common Feature [WFP]

A limitations of the above methods for cyclical features analysis is that they cannot handle the possible coexistence of differing types of reduced-rank restrictions in the same vector. In order to overcome this limitation, Cubadda (2007) introduced the notion of weak form of PSCCF, which encompasses most of the existing formulations: series Δy_t have s WFP's iff there exists an $n \times s$ matrix δ_F with full column rank such that $\delta_F' \alpha \neq 0$, $\delta_F' \Gamma_1 \neq 0$, and the VECM in (6) can be rewritten into the following partial RRR model

$$\Delta y_t = \Phi D_t + \alpha \beta' y_{t-1} + \Gamma_1 \Delta y_{t-1} + \delta_{F\perp} \psi_F' (\Delta y'_{t-2}, \dots, \Delta y'_{t-p+1})' + \varepsilon_t$$

where ψ_F is an $(np-2n) \times (n-s)$ matrix with full column rank.

The WFP requires the existence of a second-order polynomial matrix $\delta_F(L) \equiv \delta_F - (\beta \alpha' + I_n + \Gamma_1') \delta_F L + \Gamma_1' \delta_F L^2$ such that

$$\delta_F(L)' y_t = \delta_F' \Phi D_t + \delta_F' \varepsilon_t.$$

An important implication of the WFP is that the polynomial matrix $\delta_F(L)$ transforms the BN cycles κ_t into a process with shorter memory, since $\delta_F(L)' \kappa_t \sim \text{VMA}(1)$.

5.2 Common Cyclical Features: ML Inference

In order to determine the number of CCF's and consistently estimate parameters of the VECM (6), it is standard practice to resort to RRR (i.e., Engle and Kozicki, 1993; Vahid and Engle, 1993; Cubadda and Hecq, 2001). In particular, let $\text{CanCor}\{\Delta y_t, x_t | z_t\}$ denote the partial canonical correlations between series Δy_t and x_t conditional on z_t . ML inference on the various forms of common features is obtained by solving $\text{CanCor}\{\Delta y_t, x_t | z_t\}$ for proper choices of the variables x_t and z_t , as in Table 1A.

TABLE 1A
Canonical correlations and CCF's

Model	x_t'	z_t'
SCCF	$(\Delta y'_{t-1}, \dots, \Delta y'_{t-p+1}, y'_{t-1} \beta)$	D_t'
WF	$(\Delta y'_{t-1}, \dots, \Delta y'_{t-p+1})$	$(D_t', y'_{t-1} \beta)$
PSCCF	$(\Delta y'_{t-2}, \dots, \Delta y'_{t-p+1}, y'_{t-1} \beta)$	$(D_t', \Delta y_{t-1})$
WFP	$(\Delta y'_{t-2}, \dots, \Delta y'_{t-p+1})$	$(D_t', \Delta y'_{t-1}, y'_{t-1} \beta)$

Let $\hat{\lambda}_i$ denote the i -th smallest squared partial canonical correlation for $i = 1, \dots, n$. Under the null that s common features of a given form exist, the test statistic

$$LR_1 = -T \sum_{i=1}^s \ln(1 - \hat{\lambda}_i), \quad s = 1, \dots, n,$$

is asymptotically distributed as a $\chi^2(d_1)$ as detailed in Table 1B (see, *inter alia*, Velu et al., 1986; Anderson, 2002).

TABLE 1B
Tests for common features

Model	d_1
SCCF	$s \times (n(p-2) + r + j)$
WF	$s \times (n(p-2) + j)$
PSCCF	$s \times (n(p-3) + r + j)$
WFP	$s \times (n(p-3) + j)$

Moreover, let $\hat{\phi}_i^{\Delta y}$ and $\hat{\phi}_i^x$ respectively denote the partial canonical coefficients of Δy_i and x_i associated with $\hat{\lambda}_i$. Optimal estimates of both the common features vectors and (partial) RRR coefficients are then obtained as described in Table 2 (see, *inter alia*, Velu et al., 1986; Ahn and Reinsel, 1988).

TABLE 2
Estimators of the common features vectors and RRR coefficients

Model	$(\hat{\phi}_1^{\Delta y}, \dots, \hat{\phi}_s^{\Delta y})$	$(\hat{\phi}_{j+1}^x, \dots, \hat{\phi}_n^x)$
SCCF	$\hat{\delta}_S$	$\hat{\psi}_S$
WF	$\hat{\delta}_W$	$\hat{\psi}_W$
PSCCF	$\hat{\delta}_P$	$\hat{\psi}_P$
WFP	$\hat{\delta}_F$	$\hat{\psi}_F$

Finally, the remaining parameters of the various RRR models are estimated by OLS after fixing the matrices ψ 's to their estimated values.

5.2.1 Serial Correlation Common Feature when n is large

ML inference discussed so far may perform poorly with high-dimensional systems, because inversions of large variance-covariance matrices are required. Hence, there exist problems of size and power in usual tests when the number of series, n , and the number of lags in the VAR, p , are large relative to the sample size, T . Cubadda and Hecq (2011) propose an alternative strategy to detect the presence of common cyclical features, SCCF in particular: they show that a Partial Least Squares (PLS) approach can be used to test and impose a RRR structure to

“medium” n systems VAR’s even in cases when canonical correlation analysis is not feasible due to a lack of degrees of freedom. The idea is to replace the condition that a linear combination of variables must be orthogonal to the past with the one of absence of autocorrelation (see, e.g., Lucke, 1994). To illustrate the idea, consider that equation (10) implies

$$E(\delta'_s \Delta y_t | w_{t-1}) = 0 \quad (12)$$

where

$$w_{t-1} \equiv (y'_{t-1} \beta, \Delta y'_{t-1}, \dots, \Delta y'_{t-p+1})' \quad (13)$$

However, when n or p are large relatively to the sample size T , it is convenient to resort to a weaker orthogonality condition, namely

$$E(\delta'_{s_i} \Delta y_t | [\Delta y'_{t-1} \delta_{s_i}, \dots, \Delta y'_{t-p} \delta_{s_i}]) = 0, \quad i = 1, 2, \dots, s \quad (14)$$

where $\delta_s = [\delta_{s1}, \delta_{s2}, \dots, \delta_{ss}]$. This drastically decreases the number of restrictions to be imposed under the null hypothesis, thus making a test for (an implication of) SCCF feasible even when canonical correlation analysis [CCA] is not. In order to detect SCCF in high-dimensional systems, Cubadda and Hecq (2011) propose to test for this condition by means of univariate tests for no serial correlation of each $\delta'_{s_i} \Delta y_t$ having fixed δ_s to a consistent estimate of a base for the SCCF space. However, it remains the problem of estimating common feature vectors. Cubadda and Hecq (2011) use a PLS approach instead of CCA that works well even in cases when CCA is not feasible due to a lack of degrees of freedom (both in simulations and empirical applications). PLS, introduced by Wold (1985), are a family of multivariate techniques with the aim of maximizing the covariance between linear combinations of two variable sets, see, e.g., Rosipal and Kraemer (2006) for a recent survey.

From Section 2.1 and 5.2, we know that the CCA estimator of SCCF is

$$\hat{\delta}_s^{CCA} = [\hat{v}_1^{CCA}, \dots, \hat{v}_s^{CCA}],$$

where \hat{v}_i^{CCA} is the eigenvector associated with the i -th smallest eigenvalue of the matrix $\hat{\Sigma}_{yy}^{-1} \hat{\Sigma}_{yw} \hat{\Sigma}_{ww}^{-1} \hat{\Sigma}_{wy}$, where $\hat{\Sigma}_{yy} = E(\Delta y_t \Delta y_t')$, $\hat{\Sigma}_{ww} = E(w_{t-1} w'_{t-1})$ and $\hat{\Sigma}_{yw} = E(\Delta y_t w'_{t-1})$.

The PLS estimator of the SCCF matrix δ_s is

$$\hat{\delta}_s^{PLS} = [\hat{v}_1^{PLS}, \dots, \hat{v}_s^{PLS}] \quad (15)$$

where \hat{v}_i^{PLS} ($i = 1, 2, \dots, s$) is the eigenvector associated with the i -th smallest ei-

genvalue of the matrix $\hat{D}_{yy}^{-1}\hat{\Sigma}_{yy}\hat{D}_{yy}^{-1}\hat{\Sigma}_{yy}$, with D_{yy} and D_{ww} are diagonal matrices having the diagonal elements of, respectively, Σ_{yy} and Σ_{ww} .

Note that, since PLS require to invert diagonal matrices only, this method can provide estimates of δ_s (up to an identification matrix) that are less disperse and more numerically stable than the CCA estimator when the dimension of w_t approaches the sample size T .

To consistently estimate the factor weights ψ_s , let's assume that the columns of the matrix ψ_s are equal to $(n-s)$ distinct eigenvectors of the matrix $D_{ww}^{-1}\Sigma_{ww}$, then we get

$$D_{ww}^{-1}\Sigma_{wy}D_{yy}^{-1}\Sigma_{yw} = \psi_s V_\psi \delta_{s\perp}' D_{yy}^{-1}\Sigma_{yw}$$

where V_ψ is the diagonal matrix of the $(n-s)$ eigenvalues of the matrix $D_{ww}^{-1}\Sigma_{ww}$ that are associated with the eigenvectors ψ_s , which implies that the matrix ψ_s lies in the space generated by the eigenvectors associated with the positive eigenvalues of the matrix $D_{ww}^{-1}\Sigma_{wy}D_{yy}^{-1}\Sigma_{yw}$.

In order to test for the presence of SCCF in systems where $n \times p$ is large, Cubadda and Hecq (2011) consider tests for the null hypothesis that $\delta_{s_i}' \Delta y_t$ is a white-noise for $i = 1, 2, \dots, s$ as in (14). In particular, we look at the Box-Pierce test statistic

$$Q_i^j = T \sum_{l=1}^k \hat{r}_{i,j,l}^2$$

where

$$\hat{r}_{i,j,l}^2 = \left(\sum_{t=l+1}^T \frac{e_{t,i}^j e_{t-l,i}^j}{T-l} \right) / \left(\frac{1}{T} \sum_{t=1}^T e_{t,i}^j{}^2 \right)$$

and $e_{t,i}^j = \delta_{s_i}^j' \Delta y_t$ for $j = CCA, PLS$ and $i = 1, 2, \dots, s$. The test statistic Q_i^j follows asymptotically a $\chi_{(k)}^2$ distribution under the null. However, one should control the overall size of the test when applied for different values of s . Cubadda and Hecq (2011) propose two different strategies to solve this problem: using a Bonferroni correction to fix the overall significance level of the s tests computing

$$B_s^j = \max(Q_i^j), \quad j = CCA, PLS$$

and confronting it with the critical level at $\frac{\alpha}{s}\%$ in the $\chi_{(k)}^2$; or, based on Cubadda et al. (2008), testing for the null of no autocorrelation on the aggregate $a_{t,s}^j = \sum_{i=1}^s e_{t,i}^j$ for $j = CCA, PLS$. Under the null, $a_{t,s}^j$ is a white noise in large sample whereas $a_{t,s}^j$ does not converge to an innovation process when the SCCF rank is less than s .

6. COMMON FEATURES AND UNIVARIATE TIME SERIES MODELS

ARIMA models are still a popular tool for forecasting economic time series, since they often outperform large macro models. Indeed, in macroeconometrics we are typically interested in the interactions between entities, and in order to account for the links between variables we usually resort to a multivariate framework (VAR, VARMA, VECM...). However, univariate ARIMA models are also the implied final equations (FEs) of systems such as VARs or VARMA (Zellner and Palm, 1974). But here a paradox comes: FEs are theoretically non-parsimonious, whereas low order ARIMA models are empirically appropriate. Cubadda et al. (2009) show that VAR models with a reduced-rank structure arising from the presence of common features can give an answer to this paradox.

Let us start by referring to a recent solution to the well-known “autoregressivity paradox”, i.e., multivariate time series models imply highly non parsimonious ARIMA models for individual time-series.

6.1 The Final Equations (FEs) of the VAR

Let us consider the n dimensional VAR(p) (4)

$$A(L)y_t = \varepsilon_t \quad (16)$$

where deterministic terms are omitted for simplicity.

Following Zellner and Palm (1974), the univariate representation of elements of y_t can be obtained by premultiplying both sides of (16) by $A(L)^{adj}$, the adjoint matrix (or the adjugate) associated with $A(L)$, such that $A(L)^{adj} = \det[A(L)]A(L)^{-1}$, in order to obtain the Final Equations (FEs)

$$\det[A(L)]y_t = A(L)^{adj} \varepsilon_t, \quad (17)$$

where $\det[A(L)]$ is a polynomial of order np and $A(L)^{adj}$ is matrix polynomial of order $(n-1)p$.

6.2 Implications of the Final Equations

Equation (17) implies that every FE is an univariate ARMA(p^*, q^*) although it is derived from the finite order VAR(p). They have identical autoregressive parameters because $\det[\mathcal{A}(L)]$ is the same of the n equations. Here the paradox: an n -dimensional VAR(p) would imply ARMA($np, (n-1)p$) processes, an implication that is rejected when tested on economic data where one usually finds quite parsimonious ARIMA models.

To illustrate this point, let us assume that $n=3$ series are generated by the following VAR(1):

$$y_t = \begin{bmatrix} 0.5 & -0.5 & 0.5 \\ 0.25 & -0.25 + \omega & 0.25 \\ 0.5 & -0.5 & 0.5 + \omega \end{bmatrix} y_{t-1} + \varepsilon_t$$

where $y_t = (y_{1t}, y_{2t}, y_{3t})'$.

For the above VAR, the FEs are:

$$\det[\mathcal{A}(L)] = 1 - 0.5L^3\omega^2 - (0.75 + 2\omega)L + (1.25\omega + \omega^2)L^2,$$

such that individual series follow ARMA(3,2) models.

However, if $\omega = 0$, the VAR has reduced-rank structure

$$y_t = \begin{bmatrix} 0.5 \\ 0.25 \\ 0.5 \end{bmatrix} [1 \quad -1 \quad 1] y_{t-1} + \varepsilon_t$$

which produces the FEs:

$$(1 - 0.75L)y_t = \varepsilon_t + \begin{bmatrix} -0.25 & -0.5 \\ 0.25 & -1 \\ 0.5 & -0.5 \end{bmatrix} \begin{bmatrix} 1 & 0 & -1 \\ 0 & 1 & -0.5 \end{bmatrix} \varepsilon_{t-1}.$$

This implies that the univariate models are parsimonious ARMA(1,1) models with the same autoregressive parameter and cross-correlated errors and a VMA component of FEs with a factor structure.

More generally, Table 3 summaries the reduction of the individual ARMA orders due to common features restrictions. As one can see, the presence of common short-run components significantly reduces the order of the univariate ARMA representations. Hence, the existence of CCF's is a possible, economically meaningful, solution of the paradox.

TABLE 3
*Maximum ARIMA orders of univariate series generated by an n-dimensional VAR(p)
 with s cofeature restrictions*

Models	AR order	I(d)	MA order
I(0)	np	0	$(n-1)p$
SCCF	$(n-s)p$	0	$(n-s)p$
PSCCF	$(n-s)p+s$	0	$(n-s)p+(s-1)$
C(1,1)	$n(p-1)+r$	1	$(n-1)(p-1)+r$
SCCF	$(n-s)(p-1)+r$	1	$(n-s)(p-1)+r$
PSCCF	$(n-s)(p-1)+r+s$	1	$(n-s)(p-1)+r+s-1$
WF	$(n-s)(p-1)+r$	1	$(n-s)(p-1)+r$

It is worth noting that Table 3 provides the *maxima* ARIMA orders under CCF's restrictions. However, the orders can be even smaller due to additional restrictions on the VAR parameters, such as block-diagonal or block-triangular structures.

6.3 Implications of Common Cyclical Features for the VMA component

The presence of short-run comovements has also consequences for the VMA part of FEs. Cubadda et al. (2009) show that in a stationary VAR(p), the existence of s SCCFs implies that in the FEs 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, it is possible to reduce the order of the VMA component to a degree of at most $(n-s-1)p$ instead of $(n-s)p$. In particular, when $n-1=s$, the FEs follow a model that is popular in the macro-panel literature: an homogeneous AR component and cross-correlated VMA errors having a factor structure. However, in the CF's framework this model is not *a priori* assumed but it is an implication of the presence of a (testable) time series property of the data.

In order to clarify this point, it may be appropriate to use an empirical example. Cubadda et al. (2009) considered the IPIs of Canada and the US. Since they found no cointegration in log levels, a VAR(1) in first differences seemed appropriate. The estimation by OLS delivers (standard errors in brackets)

$$\begin{bmatrix} \widehat{\Delta \ln US_t} \\ \widehat{\Delta \ln CA_t} \end{bmatrix} = \begin{bmatrix} 0.333 & 0.273 \\ (0.088) & (0.079) \\ 0.265 & 0.360 \\ (0.102) & (0.092) \end{bmatrix} \begin{bmatrix} \Delta \ln US_{t-1} \\ \Delta \ln CA_{t-1} \end{bmatrix}$$

Theoretical FEs orders implies an ARMA(2,1) processes. However, SCCF test statistics is in favor of $s=1$ (p -value=0.31), with the estimated CCF relationship $(\Delta \ln US_t - 1.05 \Delta \ln CA_t)$, that is an innovation with respect to the past. Empirical estimation of univariate models gives the following ARIMA(1,1,0)

$$\begin{aligned} \widehat{\Delta \ln US_t} &= 0.003 + 0.554 \Delta \ln US_{t-1} \\ &\quad (0.001) \quad (0.062) \\ \widehat{\Delta \ln CA_t} &= 0.004 + 0.533 \Delta \ln CA_{t-1} \\ &\quad (0.001) \quad (0.064) \end{aligned}$$

As expected, AR coefficients are very similar. Moreover, since the estimated cofeature vector $\hat{\delta}'$ is close to (1: -1), the variances of VAR residuals are similar, and the correlation of VAR residuals is around 0.65, the previous discussion may explain why the MA(1) components are almost negligible.

6.4 Estimation of the Common AR parameters

As illustrated above, the finding that a set of series have identical autoregressive polynomials calls for the estimation of the common AR component that should be preferred to estimating ARMA models for the individual series. Indeed, the univariate parsimonious empirical ARMA(p_i, q_i) models can be estimated by ML for each series individually such that

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

Instead, imposing the restrictions implied by the VAR is equivalent to using an estimator of the common AR part such that

$$y_{it} = \hat{\alpha}_i + \sum_{j=1}^{p_i} \hat{\phi}_j y_{it-j} + \sum_{k=1}^{q_i} \hat{\theta}_{ik} \hat{\varepsilon}_{it-k} + \hat{\varepsilon}_{it}$$

where $\hat{\phi}_j$ is the j th lag order common coefficient to all n which should be preferred to $\hat{\phi}_{ij}$ for individual series.

In order to do so, we have various alternative at hand, such as the mean group estimator (the average of n individual estimators)

$$\hat{\phi}_j^{mg} = n^{-1} \sum_{i=1}^n \hat{\phi}_{ij}, \quad j = 1, \dots, \max(p_i)$$

and the estimation on aggregates (univariate ARMA for the average of the n series)

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

where $\bar{y}_t = n^{-1} \sum_{i=1}^n y_{it}$, v_t is the innovation of the univariate ARMA(p, q) for \bar{y}_t

and θ_k is the k th lag parameter of the MA part of \bar{y}_t .

Cubadda et al. (2008) have shown by simulation that the best estimator is simply to fit the parsimonious model on aggregates. It is worth noting that the use of this estimator is fine under the null that the FEs are from the same initial VAR

model. However, one could evaluate whether this is also true when we “erroneously substitute” to the group of variables having common cycles additional series from another group. Cubadda et al. (2009) conducted a Monte Carlo experiment and found that the empirical bias increases if one includes series with AR coefficients more distant from the FEs. However, the differences in terms of bias and RMSE decrease as n and T increase.

When we work on the aggregate series, we impose the common autoregressive roots whose estimation can be volatile in individual series. It is worth remarking that due to the factor structure of the VMA component of the FEs, aggregation can reduce the order of the univariate MA components, which are noise in this estimation problem. For instance, in the FEs

$$(1 - 0.75L)y_t = \varepsilon_t + \begin{bmatrix} -0.25 & -0.5 \\ 0.25 & -1 \\ 0.5 & -0.5 \end{bmatrix} \begin{bmatrix} 1 & 0 & -1 \\ 0 & 1 & -0.5 \end{bmatrix} \varepsilon_{t-1}$$

pre-multiplication of both sides by $(1, 1, 1)'$ annihilates the VMA component.

7. FINAL REMARKS

The main idea behind this survey was to create a “common thread” between various topics, related to each other by trying to model different kind of comovements typically present in economic time series, and in many cases also predicted by economic theory. Indeed, from a statistical point of view, comovements imply a reduction to a more parsimonious structure such as common factor representation: a small number of unobserved components possess a given feature and transmit it to a larger set of economic time series. As we have seen, RRR is often the solution to the inferential problem.

The large amount of literature on comovement has forced us to keep the discussion on a limited number of issues: specifically, trends, cycles and seasonality. However, we believe we have taken account of important extensions: the implications of common features for univariate time series models and how to address efficiency problems when the number of the variables is fairly numerous.

The main drawback of the methods discussed in this paper is that different features are evaluated separately: one important exception is the unifying framework for analyzing common cyclical features as discussed in Section (5.1.4). Then, a major challenge ahead is to develop estimation and testing procedures that allow for joint identification of several common features using an integrated approach. Further, the analysis of various forms of reduced rank structure in large dimensional systems appears to be promising.

LUMSA Università, Roma

*Dipartimento SEFEMEQ
Università di Roma “Tor Vergata”*

MARCO CENTONI

GIANLUCA CUBADDA

REFERENCES

- AHN S, REINSEL G. 1988. Nested reduced-rank autoregressive models for multiple time-series. *Journal of the American Statistical Association* **83**: 849-856.
- AHN SK, REINSEL GC. 1994. Estimation of partially nonstationary vector autoregressive models with seasonal behavior. *Journal of Econometrics* **62**: 317-350.
- ANDERSON HM, VAHID F. 1998. Testing multiple equation systems for common nonlinear components. *Journal of Econometrics* **84**: 1-36.
- ANDERSON T. 1984. *An Introduction to Multivariate Statistical Analysis*. Wiley, second edition. New York.
- ANDERSON T. 1999. Asymptotic theory for canonical correlation analysis. *Journal of Multivariate Analysis* **70**: 1-29.
- ANDERSON T. 2002. Canonical correlation analysis and reduced rank regression in autoregressive models. *Annals of Statistics* **30**: 1134-1154.
- BAXTER M, KOUPARITSAS MA. 2005. Determinants of business cycle comovement: A robust analysis. *Journal of Monetary Economics* **52**: 113-157.
- BEVERIDGE S, NELSON C. 1981. A new approach to decomposition of economic time series into permanent and transitory component with particular attention to measurement of the 'business cycle'. *Journal of Monetary Economics* **7**: 151-174.
- BIERENS H. 2000. Nonparametric nonlinear cotrending analysis, with an application to interest and inflation in the United States. *Journal of Business & Economic Statistics* **18**: 323-337.
- CAMPBELL J. 1987. Does saving anticipate declining labor income? An alternative test of the permanent income hypothesis. *Econometrica* **55**: 1249-1273.
- CAMPBELL JY, MANKIW NG. 1990. Permanent income, current income, and consumption. *Journal of Business & Economic Statistics* **8**: 265-279.
- CHAPMAN DA, OGAKI M. 1993. Cotrending and the stationarity of the real interest rate. *Economics Letters* **42**: 133-138.
- CHRISTIANO LJ, EICHENBAUM M, EVANS CL. 2005. Nominal rigidities and the dynamic effects of a shock to monetary policy. *Journal of Political Economy* **113**: 1-45.
- CROUX C, FORNI M, REICHLIN L. 2001. A measure of comovement for economic variables: Theory and empirics. *Review of Economics and Statistics* **83**: 232-241.
- CUBADDA G. 1999. Common cycles in seasonal non-stationary time series. *Journal of Applied Econometrics* **14**: 273-291.
- CUBADDA G. 2001. Complex reduced rank models for seasonally cointegrated time series. *Oxford Bulletin of Economics and Statistics* **63**: 497-511.
- CUBADDA G. 2007. A unifying framework for analysing common cyclical features in cointegrated time series. *Computational Statistics & Data Analysis* **52**: 896-906.
- CUBADDA G, HECQ A. 2001. On non-contemporaneous short-run co-movements. *Economics Letters* **73**: 389-397.
- CUBADDA G, HECQ A. 2011. Testing for common autocorrelation in data-rich environments. *Journal of Forecasting* **30**: 325-335.
- CUBADDA G, HECQ A, PALM FC. 2008. Macro-panels and reality. *Economics Letters* **99**: 537-540.
- CUBADDA G, HECQ A, PALM FC. 2009. Studying co-movements in large multivariate data prior to multivariate modelling. *Journal of Econometrics* **148**: 25-35.
- CUBADDA G, OMTZIGT P. 2005. Small-sample improvements in the statistical analysis of seasonally cointegrated systems. *Computational Statistics and Data Analysis* **49**: 333-348.
- ENGLE RF, GRANGER CW. 1987. Cointegration and error correction: Representation, estimation and testing. *Econometrica* **55**: 251-276.

- ENGLE RF, HYLLEBERG S. 1996. Common seasonal features: Global unemployment. *Oxford Bulletin of Economics and Statistics* **58**: 615-630.
- ENGLE RF, ISSLER JV. 1995. Estimating common sectoral cycles. *Journal of Monetary Economics* **35**: 83-113.
- ENGLE RF, KOZICKI S. 1993. Testing for common features. *Journal of Business & Economic Statistics* **11**: 369-380.
- ENGLE RF, MARCUCCI J. 2006. A long-run pure variance common features model for the common volatilities of the Dow Jones. *Journal of Econometrics* **132**: 7-42.
- ENGLE RF, SUSMEL R. 1993. Common volatility in international equity markets. *Journal of Business & Economic Statistics* **11**: 167-176.
- ERICSSON NR. 1993. [Testing for common features]: Comment. *Journal of Business & Economic Statistics* **11**: 380-383.
- FLAVIN M. 1993. The excess smoothness of consumption: Identification and interpretation. *Review of Economic Studies* **60**: 651-666.
- FORBES K, RIGOBON R. 2002. No contagion, only interdependence: Measuring stock market comovements. *Journal of Finance* **57**: 2223-2261.
- GONZALO J, GRANGER CW. 1995. Estimation of common long-memory components in cointegrated systems. *Journal of Business & Economic Statistics* **13**: 27-35.
- GOURIEROUX C, PEAUCELLE I. 1988. Detecting a long run relationship (with an application to the P.P.P. hypothesis). Working Paper 8902, CEPREMAP.
- GRANGER CW. 1981. Some properties of time series data and their use in econometric model specification. *Journal of Econometrics* **16**: 121-130.
- HALDRUP N, HYLLEBERG S, PONS G, SANZO A. 2007. Common periodic correlation features and the interaction of stocks and flows in daily airport data. *Journal of Business & Economic Statistics* **25**: 21-32.
- HALL RE. 1978. Stochastic implications of the life cycle-permanent income hypothesis: Theory and evidence. *The Journal of Political Economy* **86**: 971-987.
- HARDING D, PAGAN A. 2006. Synchronization of cycles. *Journal of Econometrics* **132**: 59-79.
- HECQ A, PALM F, URBAIN J. 2000. Testing for common cyclical features in nonstationary panel data models. In *Advances Econometrics, VOL 15, 2000*, Volume 15 of *Advances in Econometrics: a Research Annual*. JAI-Elsevier Science Inc, 131-160.
- HECQ A, PALM F, URBAIN J. 2006. Common cyclical features analysis in VAR models with cointegration. *Journal of Econometrics* **132**: 117-141.
- HENDRY DF. 1996. A theory of co-breaking. mimeo, Nuffield College, University of Oxford.
- HYLLEBERG S, ENGLE R, GRANGER C, YOO B. 1990. Seasonal integration and cointegration. *Journal of Econometrics* **44**: 215-238.
- HYLLEBERG S, JØRGENSEN C, SØRENSEN NK. 1993. Seasonality in macroeconomic time series. *Empirical Economics* **18**: 321-335.
- IMBS J. 2004. Trade, finance, specialization, and synchronization. *Review of Economics and Statistics* **86**: 723-734.
- ISSLER JV, VAHID F. 2001. Common cycles and the importance of transitory shocks to macroeconomic aggregates. *Journal of Monetary Economics* **47**: 449-475.
- JOHANSEN S. 1988. Statistical analysis of cointegrating vectors. *Journal of Economic Dynamics and Control* **12**: 231-254.
- JOHANSEN S. 1991. Estimation and hypothesis testing of cointegrating vectors in gaussian vector autoregressive models. *Econometrica* **59**: 1551-1580.
- JOHANSEN S. 1996. *Likelihood-based Inference in Cointegrated Vector Autoregressive Models*. Oxford University Press, second edition.

- JOHANSEN S, SCHAUMBURG E. 1998. Likelihood analysis of seasonal cointegration. *Journal of Econometrics* **88**: 301-339.
- JUSELIUS K. 2006. *The Cointegrated VAR Model*. Oxford University Press.
- JUSTINIANO A, PRIMICERI GE, TAMBALOTTI A. 2010. Investment shocks and business cycles. *Journal of Monetary Economics* **57**: 132-145.
- KHAN H, TSOUKALAS J. 2011. Investment shocks and the comovement problem. *Journal of Economic Dynamics and Control* **35**: 115-130.
- KING R, PLOSSER C, REBELO S. 1988. Production, growth and business cycles. 2. New directions. *Journal of Monetary Economics* **21**: 309-341.
- KING RG, PLOSSER CI, STOCK JH, WATSON MW. 1991. Stochastic trends and economic fluctuations. *The American Economic Review* **81**: 819-840.
- LEE H. 1992. Maximum likelihood inference on cointegration and seasonal cointegration. *Journal of Econometrics* **54**: 1-47.
- LONG JB, PLOSSER CI. 1983. Real business cycles. *The Journal of Political Economy* **91**: 39-69.
- LUCAS R. 1977. Understanding business cycles. *Carnegie-Rochester Conference Series on Public Policy* **5**: 7-29.
- LUCKE B. 1994. Analysis of West Germany macroeconomic data using common trends and common cycles. Discussion Paper 3, Free University of Berlin.
- PARUOLO P. 2006. Common trends and cycles in I(2) VAR systems. *Journal of Econometrics* **132**: 143-168.
- PHILLIPS PC. 1994. Some exact distribution-theory for maximum-likelihood estimators of cointegrating coefficients in error-correction models. *Econometrica* **62**: 73-93.
- ROSIPAL R, KRAEMER N. 2006. Overview and recent advances in partial least squares. In Saunders G, Grobelenik M, Gunn S, ShaweTaylor J (eds.) *Subspace, Latent Structure and Feature Selection*, volume 3940 of *Lecture Notes in Computer Science*. Springer-Verlag Berlin, 34-51.
- SCHLEICHER C. 2007. Codependence in cointegrated autoregressive models. *Journal of Applied Econometrics* **22**: 137-159.
- STOCK JH, WATSON MW. 1988. Testing for common trends. *Journal of the American Statistical Association* **83**: 1097-1107.
- TIAO GC, TSAY RS. 1989. Model specification in multivariate time series. *Journal of the Royal Statistical Society. Series B (Methodological)* **51**: 157-213.
- URGA G. 2007. Common features in economics and finance: An overview of recent developments. *Journal of Business & Economic Statistics* **25**: 2-11.
- VAHID F, ENGLE RF. 1993. Common trends and common cycles. *Journal of Applied Econometrics* **8**: 341-360.
- VAHID F, ENGLE RF. 1997. Codependent cycles. *Journal of Econometrics* **80**: 199-221.
- VAHID F, ISSLER JV. 2002. The importance of common cyclical features in VAR analysis: a Monte-Carlo study. *Journal of Econometrics* **109**: 341-363.
- VELU R, REINSEL G, WICHERN D. 1986. Reduced rank models for multiple time series. *Biometrika* **73**: 105-118.
- WOLD H. 1985. Partial least squares. In Kotz S, Johnson NL (eds.) *Encyclopedia of Statistical Sciences*, volume 6. New York: Wiley, 581-591.
- ZELLNER A, PALM F. 1974. Time series analysis and simultaneous equation econometric models. *Journal of Econometrics* **2**: 17-54.

SUMMARY

Modelling comovements of economic time series: a selective survey

Modelling comovements amongst multiple economic variables takes up a relevant part of the literature in time series econometrics. Comovement can be defined as “move together”, that is as movement that several series have in common. The pattern of the series could be of different nature, such as trend, cycles, seasonality, being the results of different driving forces. As a results, series that comove share some common features. Common trends, common cycles, common seasonality are terms that are often found in the literature, different in scope but all aimed at modeling common behavior of the series. However, modeling comovements is not only a statistical matter, since in many cases common features are predicted by economic theory, resulting from the optimizing behavior of economic agents.