Provided for non-commercial research and education use.
Not for reproduction, distribution or commercial use.

This article appeared in a journal published by Elsevier. The attached copy is furnished to the author for internal non-commercial research and education use, including for instruction at the authors institution and sharing with colleagues.

Other uses, including reproduction and distribution, or selling or licensing copies, or posting to personal, institutional or third party websites are prohibited.

In most cases authors are permitted to post their version of the article (e.g. in Word or Tex form) to their personal website or institutional repository. Authors requiring further information regarding Elsevier's archiving and manuscript policies are encouraged to visit:

http://www.elsevier.com/authorsrights

Author's personal copy

Economic Modelling 33 (2013) 367–374

Contents lists available at SciVerse ScienceDirect

Economic Modelling

journal homepage: www.elsevier.com/locate/ecmod

A general to specific approach for constructing composite business cycle indicators ☆

Gianluca Cubadda [a,*], Barbara Guardabascio [b], Alain Hecq [c]

[a] Università di Roma "Tor Vergata", Department of Economics and Finance, Via Columbia 2, 00133 Roma, Italy
[b] ISTAT, DCSC-SER/C, Viale Liegi 13, 00198 Roma, Italy
[c] Maastricht University, Department of Quantitative Economics, P.O. Box 616, 6200 MD Maastricht, The Netherlands

ARTICLE INFO

Article history:
Accepted 3 April 2013
Available online xxxx

JEL classification:
C32

Keywords:
Co-movements
Common cycles
Composite business cycle indicators
Euro area

ABSTRACT

Combining economic time series with the aim to obtain an indicator for business cycle analyses is an important issue for policy makers. In this area, econometric techniques usually rely on systems with either a small number of series, *N*, or, at the other extreme, a very large *N*. In this paper we propose tools to select the relevant business cycle indicators in a "medium" *N* framework, a situation that is likely to be the most frequent in empirical works. An example is provided by our empirical application, in which we study jointly the short-run co-movements of 24 European countries. We show, under not too restrictive conditions, that parsimonious single-equation models can be used to split a set of *N* countries in three groups. The first group comprises countries that share a synchronous common cycle, a non-synchronous common cycle is present among the countries of the second group, and the third group collects countries that exhibit idiosyncratic cycles. Moreover, we offer a method for constructing a composite coincident indicator that explicitly takes into account the existence of these various forms of short-run co-movements among variables.
© 2013 Elsevier B.V. All rights reserved.

1. Introduction

Building a composite business cycle indicator from a set of *N* economic time series is, per se, not too difficult. One can simply either average the relevant individual indicators or combine them using factor models, see inter alia Stock and Watson (1989) and Forni et al. (2001). In practice, it is not obvious that more elaborated methods produce more accurate results. For instance, it is illustrated in Hecq (2005) that randomly generated linear combinations of the four coincident indicators used by The Conference Board provide with composite indicators that deliver very similar turning points of the US economic activity. However, an improvement in forecasting the business cycles is observed for methods that explicitly take into account the existence of short-run co-movements among the individual business cycle indicators (Cubadda, 2007a).

This paper contributes to the literature on the identification of common cycles and the construction of composite business cycle indicators in two ways.

First, we provide methods for selecting the individual cyclical indicators that are based on the detection of common cycles among variables. Indeed, prior to the building of any composite business cycle indicator, we propose to dig deeper into the detection of groups of variables that are homogenous with regard to the presence of short-run co-movements. For instance, let us consider the empirical investigation that we have in this paper, namely the analysis of GDP growth rates of 24 European countries. Several studies have emphasized the existence of business cycle co-movements among European economies, see the survey by De Haan et al. (2008) and the references therein. Our main concern in this paper is to develop a strategy aiming at splitting those *N* countries into three groups of respectively N_1, N_2, and N_3 time series. These three clusters will be obtained thanks to a measure of the degree of cyclical commonality among the various economies. In particular, the first group is such that there is a common synchronous cycle among these N_1 time series (Engle and Kozicki, 1993; Vahid and Engle, 1993); the N_2 variables of the second group share a non-synchronous common cycle (Cubadda and Hecq, 2001), and the last group comprises N_3 series with idiosyncratic short-run dynamics. For small dimensional systems, a VAR analysis with additional reduced rank restrictions can be undertaken to discover these groups (Cubadda, 2007b). However, this strategy is unfeasible when *N* is too large relatively to the number of observations *T*. Hence, we provide some mild assumptions under which each equation of a VAR is endowed with a factor structure. The main attractive features of this approach are twofold: (*i*) the presence of the various kinds of co-movements is determined using

☆ Previous drafts of this paper were presented at the 30th CIRET Conference in New York, the 31st International Symposium on Forecasting in Prague, the 65th European Meeting of the Econometric Society in Oslo, and the CIRET/KOF/HSE Workshop on "National Business Cycles in the Global World" in Moscow. We wish to thank the participants, as well as Gian Luigi Mazzi, Antoni Espasa and two anonymous referees, for helpful comments and suggestions. The usual disclaimers apply.
* Corresponding author.
E-mail addresses: gianluca.cubadda@uniroma2.it (G. Cubadda), guardabascio@istat.it (B. Guardabascio), a.hecq@maastrichtuniversity.nl (A. Hecq).
URL's: http://www.economia.uniroma2.it/nuovo/facolta/docenti/curriculum/GianlucaCubadda.html (G. Cubadda), http://www.personeel.unimaas.nl/a.hecq/ (A. Hecq).

0264-9993/$ – see front matter © 2013 Elsevier B.V. All rights reserved.
http://dx.doi.org/10.1016/j.econmod.2013.04.007

Author's personal copy

only parsimonious single-equation models; (*ii*) these models can be specified according to the general-to-specific methodology; see, inter alia, Campos et al. (2005). In particular, one can rely on automatic selection procedures such as those already implemented in Gretl or in OxMetrics for instance.

Second, after having determined these groups, we offer a method of constructing the "best" composite coincident indicator that explicitly takes into account the existence of these various forms of short-run co-movements among variables. In particular, series from the second group are preliminarily "aligned" in order to display a common synchronous cycle with the variables of the first group. Then we exploit the common cycle property in order to build a unique composite coincident indicator.

The paper is organized as follows. Section 2 presents our new method for investigating the presence of different kind of co-movements in a set of N time series, N being too large to rely on usual multivariate time series tools. Section 3 evaluates our procedure in the light of a Monte Carlo analysis. We compare automatic selection procedures based either on sequential Wald tests or on information criteria. Section 4 is dedicated to analyze co-movements among 24 European countries, build the composite coincident indicator, and compare it with already existing indicators. Section 5 concludes.

2. Identification of business cycle co-movements

2.1. Synchronous and non-synchronous common cycles

Let $Y_t \equiv (y_{1t},\ldots,y_{Nt})'$ denote the N-vector of the time series of interest. We assume that Y_t is generated by the following stationary vector autoregressive model of order p (VAR(p) hereafter):

$$\Phi(L)Y_t = \varepsilon_t, t = 1, 2, \ldots, T, \tag{1}$$

where $\Phi(L) = I_n - \sum_{j=1}^{p} \Phi_j L^j$ and ε_t are i.i.d. innovations with $E(\varepsilon_t) = 0$, $E\left(\varepsilon_t \varepsilon'_t\right) = \Sigma$ (positive definite) and finite fourth moments.

In this framework, serial correlation common feature (SCCF hereafter, see Engle and Kozicki, 1993) holds if there exists a full-rank $(N \times s)$-matrix δ $(s < N)$, whose columns span the cofeature space, such that

$$\delta' Y_t = \delta' \varepsilon_t \tag{2}$$

is a s-dimensional zero mean vector innovation process with respect to the information available at time $t-1$. Consequently, SCCF arises if there exists a matrix δ such that the conditions $\delta'_j \Phi_j = 0_{(s \times N)}, j = 1,\ldots, p$ are jointly satisfied.

Notice that under SCCF, the VAR model (1) can be rewritten as the following reduced-rank regression model

$$Y_t = \delta_\perp \sum_{j=i}^{p} A'_j Y_{t-j} + \varepsilon_t \equiv \delta_\perp X_{t-1} + \varepsilon_t,$$

where $\delta' \delta_\perp = 0$ and A_j is a $(N-s) \times s$ matrix for $j = 1,\ldots, p$. Since all the predictable fluctuations of series Y_t are due to the $(N-s)$ common dynamic factors X_{t-1}, the existence of SCCF is equivalent to the presence of synchronous common cycles among series Y_t.

Moreover, Vahid and Engle (1993) show that if series Y_t are the first differences of I(1) variables, condition (2) is equivalent to the presence of $(N-s)$ common components among the deviations of the series levels from their random walk trends. Hence, the notion of SCCF could in principle be used to construct composite business cycle indicators based on both the traditional notions of "growth cycle" (i.e. fluctuations in the economic activity around a long-run trend) and "growth rate cycle" (i.e. fluctuations of the growth rate of economic activity). Within this paper we follow the growth rate cycle approach mainly because the limited time span of our data prevents a statistically credible analysis of their long-run properties.

In order to allow for adjustment delays, Cubadda and Hecq (2001) propose to look at the presence of non-synchronous common cycles in the context of the polynomial serial correlation common feature (PSCCF hereafter) modeling. In this framework there exists a full-rank $(N \times s)$ matrix δ_0 such that under the null hypothesis that PSCCF of order m $(1 \leq m < p)$ holds if the conditions $\delta'_0 \Phi_h \neq 0_{(s \times N)}$, $h = 1,\ldots, m$, and $\delta'_0 \Phi_j = 0_{(s \times N)}, j = m+1,\ldots, p$ are jointly satisfied. This is equivalent to requiring that there exists a polynomial matrix $\delta(L) = \delta_0 - \sum_{h=1}^{m} \delta_h L^h$ such that

$$\delta'(L) Y_t = \delta'_0 \varepsilon_t,$$

where $\delta_h = \Phi'_h \delta_0$, $i = 1,\ldots, m$.

Under PSCCF, the VAR model can be rewritten as the following partially reduced-rank regression model

$$Y_t - \sum_{h=1}^{m} \Phi_h Y_{t-h} = \delta_{0\perp} \sum_{j=m+1}^{p} A'_{0j} Y_{t-j} + \varepsilon_t \equiv \delta_{0\perp} X_{0t-1} + \varepsilon_t,$$

where A'_{0j} is a $(N-s) \times s$ matrix for $j = m+1,\ldots, p$, which reveals that series Y_t share common dynamics after m periods.

Issler and Vahid (2006) and Cubadda (2007a) discuss how to obtain composite cyclical indicators under, respectively, SCCF and PSCCF. For instance, Issler and Vahid (2006) look at the linear combinations $\delta'_\perp Y_t$. In Section 5 we extend this approach to the case that both SCCF and PSCCF are present in the data.

2.2. A joint determination of the groups

Let us now assume that there exists a partition $Y_t = [Y_{1t}, Y_{2t}, Y_{3t}]$, where the N_1 series Y_{1t} share a synchronous common cycle, the N_2 series Y_{2t} share a non synchronous common cycle, and the remaining series Y_{3t} present idiosyncratic short-run dynamics. According to the definitions provided in the previous section, this is equivalent to assuming that series Y_{1t} are characterized by the presence of $(N_1 - 1)$ SCCF vectors, series Y_{2t} exhibit $(N_2 - 1)$ PSCCF vectors and the remaining series Y_{3t} do not present any SCCF or PSCCF. With the researcher being not aware of this partition, the goal is to find out the series that are co-moving and belong to sets Y_{1t} and Y_{2t}. This new strategy is developed in this section.

Since we have assumed that series Y_t are generated by a stationary VAR(p) model, each series Y_{it} follows the stable dynamic regression model

$$y_{it} = \phi_{ii}(L) y_{it-1} + \sum_{k \neq i}^{N} \phi_{ik}(L) y_{kt-1} + \varepsilon_{it}, \quad i = 1, \ldots, N; t = 1, \ldots, T, \tag{3}$$

where $\phi_{it}(L)$ and $\phi_{ik}(L)$ are scalar polynomials of order $(p-1)$, and ε_{it} is i-th element of the innovation vector ε_t.

A statistical issue arises when the number of regressors, $N \times p$, becomes too large with respect to the sample size T. For instance, in our empirical application we have $N = 24$ and $T = 56$, which implies that it is not even feasible to estimate the unrestricted model (3) by OLS for $p > 2$. Hence, we further assume that

$$\sum_{k \neq i}^{n} \phi_{ik}(L) y_{kt} = \beta_i(L) \sum_{k \neq i}^{n} \omega_k y_{kt}, \tag{4}$$

where $\beta_i(L)$ is a scalar polynomial of order $p-1$ and ω_i is a scalar for $i = 1,\ldots, N$.

Notice that this is equivalent to postulating the following factor-augmented autoregressive (FAAR) structure for each series y_{it}

$$y_{it} = \alpha_i(L) y_{it-1} + \beta_i(L) f_{t-1} + \varepsilon_{it}, \tag{5}$$

where $\alpha_i(L) = [\phi_{ii}(L) - \beta_i(L)\omega_i]$ and the common factor is $f_t = \sum_{k=1}^{n} \omega_k y_{kt}$. In Section 4 we discuss various alternatives for

Author's personal copy

constructing these country weights by using either official statistics or multivariate techniques.

Let us suppose that series i and j share a common cycle with a normalized SCCF vector $(1,-\delta_{ij})'$. This implies that

$$y_{it}-\delta_{ij}y_{jt}=\varepsilon_{it}-\delta_{ij}\varepsilon_{jt},$$

which, in view of equations

$$y_{it}=\phi_{ii}(L)y_{it-1}+\beta_i(L)\sum_{k\neq i}^{N}\omega_k y_{kt}+\varepsilon_{it}, \tag{6}$$

$$y_{jt}=\phi_{jj}(L)y_{jt-1}+\beta_j(L)\sum_{k\neq j}^{N}\omega_k y_{kt}+\varepsilon_{jt}, \tag{7}$$

requires that the following set of conditions holds

$$\beta_i(L)=\delta_{ij}\beta_j(L), \tag{8}$$

$$\phi_{ii}(L)=\delta_{ij}\beta_j(L)\omega_i, \tag{9}$$

$$\beta_i(L)\omega_j=\delta_{ij}\phi_{jj}(L). \tag{10}$$

From Eqs. (8) and (9) we get

$$\phi_{ii}(L)=\omega_i\beta_i(L), \tag{11}$$

whereas from Eqs. (8) and (10) we get

$$\phi_{jj}(L)=\omega_j\beta_j(L). \tag{12}$$

If we put restrictions (11) and (12) in Eqs. (6) and (7) we finally obtain

$$y_{it}=\delta_{ij}\beta_j(L)f_{t-1}+\varepsilon_{it}, \tag{13}$$

$$y_{jt}=\beta_j(L)f_{t-1}+\varepsilon_{jt}. \tag{14}$$

This leads to the following condition for the existence of a common synchronous cycle between variables y_{it} and y_{jt}:

Condition 1. Under assumption (4), it follows from Eqs. (13) and (14) that an SCCF relationship between y_{it} and y_{jt} requires restriction (8) and

$$\alpha_i(L)=\alpha_j(L)=0. \tag{15}$$

It is important to notice, and we will make use of this particular case in our testing strategy, that the above condition is always satisfied when Eq. (15) holds and $p=1$. Indeed, if $p=1$, it turns out that one simply needs to test for the restricted models

$$\begin{aligned}y_{it}&=\beta_i f_{t-1}+\varepsilon_{it},\\ y_{jt}&=\beta_j f_{t-1}+\varepsilon_{jt},\end{aligned} \tag{16}$$

versus the unrestricted models

$$\begin{aligned}y_{it}&=\alpha_i y_{it-1}+\beta_i f_{t-1}+\varepsilon_{it},\\ y_{jt}&=\alpha_j y_{jt-1}+\beta_j f_{t-1}+\varepsilon_{jt},\end{aligned} \tag{17}$$

which include as special cases the idiosyncratic cycle models

$$\begin{aligned}y_{it}&=\alpha_i y_{it-1}+\varepsilon_{it},\\ y_{jt}&=\alpha_j y_{jt-1}+\varepsilon_{jt}.\end{aligned} \tag{18}$$

Indeed, system (16) exhibits an SCCF between y_{it} and y_{jt} when the joint hypothesis H_0: $\alpha_i=\alpha_j=0$ is satisfied. The synchronous common feature coefficient is obtained by the ratio of the factor loading, namely $\delta_{ij}=\beta_i/\beta_j$. Notice that when both β_i and β_j are different from zero, system (17) exhibits a non-synchronous common cycle with a PSCCF vector $[1,-\beta_i/\beta_j]'$.

When $p>1$, we need to test for the reduced-rank regression model

$$\begin{bmatrix}y_{it}\\ y_{jt}\end{bmatrix}=\begin{bmatrix}\delta_{ij}\\ 1\end{bmatrix}\beta(L)f_{t-1}+\begin{bmatrix}\varepsilon_{it}\\ \varepsilon_{jt}\end{bmatrix},$$

where $\beta(L)$ is a scalar polynomial of order $(p-1)$, versus the unrestricted system

$$\begin{aligned}y_{it}&=\alpha_i(L)y_{it-1}+\beta_i(L)f_{t-1}+\varepsilon_{it},\\ y_{jt}&=\alpha_j(L)y_{jt-1}+\beta_j(L)f_{t-1}+\varepsilon_{jt}.\end{aligned} \tag{19}$$

Let us now turn to the condition for the existence of a non-synchronous common cycle between variables y_{it} and y_{jt}. For the sake of interpretation, we only consider PSCCF relationships that solely involve the lags of y_{it} and y_{jt}.

Condition 2. Under assumption (4), in view of Eq. (19), a PSCCF relationship between y_{it} and y_{jt} and their own lags of order m requires the restrictions

$$\beta_i(L)=\delta_{0ij}\beta_j(L),$$

and

$$\alpha_i(L)=\sum_{h=1}^{m}\alpha_{ih}L^h,$$
$$\alpha_j(L)=\sum_{h=1}^{m}\alpha_{jh}L^h.$$

In the particular case with $p=1$ the above condition is trivially satisfied when both β_i and β_j are different from zero and at least one coefficient between α_i and α_j is not null in system (17). When $p>1$, we need to test for the partially reduced-rank regression model

$$\begin{bmatrix}y_{it}\\ y_{jt}\end{bmatrix}=\begin{bmatrix}\sum_{h=1}^{m}\alpha_{ih}y_{it-h}\\ \sum_{h=1}^{m}\alpha_{jh}y_{jt-h}\end{bmatrix}+\begin{bmatrix}\delta_{0ij}\\ 1\end{bmatrix}\beta(L)f_{t-1}+\begin{bmatrix}\varepsilon_{it}\\ \varepsilon_{jt}\end{bmatrix} \tag{20}$$

versus the unrestricted system (19).

Interestingly, as long as the coefficient matrices of f_{t-1} have the same reduced-rank structure as in systems (19) and (20), it is possible to “align” series with PSCCF relationships in order to display a common synchronous cycle with variables having SCCF relationships. Indeed, to simplify matters let us consider the following system for three variables

$$\begin{bmatrix}y_{it}\\ y_{jt}\\ y_{kt}\end{bmatrix}=\begin{bmatrix}0\\ \sum_{h=1}^{m}\alpha_{jh}y_{jt-h}\\ 0\end{bmatrix}+\begin{bmatrix}\delta_{ik}\\ \delta_{0jk}\\ 1\end{bmatrix}\beta(L)f_{t-1}+\begin{bmatrix}\varepsilon_{it}\\ \varepsilon_{jt}\\ \varepsilon_{kt}\end{bmatrix}. \tag{21}$$

It is seen that y_{it} and y_{kt} share a common synchronous cycle with an SCCF vector $[1,-\delta_{ik}]'$, and y_{jt} and y_{kt} share a common non-synchronous cycle with a PSCCF vector $[1,-\delta_{0jk}]'$. We can rewrite system (21) as

$$\begin{bmatrix}y_{it}\\ y^*_{jt}\\ y_{kt}\end{bmatrix}=\tilde{\delta}_{\perp}\beta(L)f_{t-1}+\begin{bmatrix}\varepsilon_{it}\\ \varepsilon_{jt}\\ \varepsilon_{kt}\end{bmatrix},$$

where $y^*_{jt}=y_{jt}-\sum_{h=1}^{m}\alpha_{jh}y_{jt-h}$ and $\tilde{\delta}_{\perp}=\left[\delta_{ik},\delta_{0jk},1\right]'$, from which we notice that y_{it}, y_{kt}, and the aligned series y^*_{jt} share a common synchronous cycle.

Author's personal copy

3. Monte Carlo results

The results of the previous section suggest that, under assumption (4), common synchronous and non-synchronous cycles can be detected simply by applying a general to specific selection procedure to each unrestricted FAAR model (5) and checking if some of the selected models satisfy the (P)SCCF restrictions. Hence, we carry out a Monte Carlo study to evaluate the performances of information criteria as well as Wald tests in selecting restricted models that are compatible with the existence of common cycles.

As in our empirical application, we consider a VAR system with $N = 24$ variables. We simulate 24 time series accordingly to the following stationary VAR(1) model

$$Y_t = \Phi Y_{t-1} + \varepsilon_t,$$
$$\Phi = A + b\omega',$$

where A is a 24×24 matrix that can be partitioned as

$$A = \begin{bmatrix} 0 & 0 & 0 \\ 0 & A^- & 0 \\ 0 & 0 & A^+ \end{bmatrix}.$$

Each block of A is composed by an 8×8 matrix. Moreover, A^- and A^+ are diagonal matrices where the diagonal elements are respectively generated by an $U(0.1,0.9)$ and an $-U(0.1,0.9)$, and b is a 24-vector where the first 16 elements are generated by an $U(0.5,2)$ and the other 8 elements equal zero. Finally, ω is a 24-vector of positive numbers that are drawn from a uniform distribution such that $\sum_{i=1}^{24} w_i = 1$, and the factor is $f_t = \sum_{i=1}^{24} w_i y_{it}$.

This DGP allows us to divide Y_t in three different clusters: (i) the first 8-subvector of Y_t, say Y_{1t}, is a function of its own shocks and the factors f_{t-1} only; (ii) each element of the second 8-subvector of Y_t, say Y_{2t}, is a function of its own shock, its own lag, and the factors f_{t-1}; (iii) each element of the third 8-subvector of Y_t, say Y_{3t}, is a function of its own shock and lag only. Hence, series Y_{1t} have a single common synchronous cycle, series Y_{2t} share the same cycle but in a non-synchronous way, and series Y_{3t} share no cycle among themselves and with the remaining series.

We generate $T + 50$ observations of the vector series Y_t for T = 50, 100, 250 where the first 50 points are used as a burn-in-period. The errors ε_t are Gaussian *i.i.d.* with zero mean and a variance matrix generated from a standard Wishart distribution with 24 degrees of freedom, and 5000 replications are used.

In order to compare alternative methods of identifying series that belong to the three different clusters, we start from the following FAAR model:

$$y_t = \alpha_0 + \alpha_1 y_{t-1} + \alpha_2 y_{t-2} + \beta_1 f_{t-1} + \beta_2 f_{t-2} + \varepsilon_t,$$

where Y_t is the generic element of vector Y_t.[1] First, we apply a sequence of t-tests following a general to specific approach. We recursively delete variables associated with the smallest values of the t-tests until we reject the null at a 10% significance level. Second, we estimate all the possible nested models and keep the one associated with the smallest value of the Bayesian information criterion (BIC).[2]

In the Tables 1 and 2 we report the frequencies with which all the possible nested models are chosen by the considered criteria. We write in bold characters the frequencies with which we correctly select series belonging to the three sub-vectors $[Y_{1t}, Y_{2t}, Y_{3t}]$. For instance, row f_{t-1} reports the percentage with which models that include f_{t-1} as the unique regressor are detected. Notice that this is a sufficient condition for having an SCCF vector because a common cycle may be present also in models retained under either f_{t-2} or $[f_{t-1}, f_{t-2}]$, although in this latter case, we need to check whether the factor loadings share the same left null space.

Overall, we would recommend the use of information criteria instead of Wald type tests because the frequencies with which the groups are detected are much higher. Moreover, when a synchronous cycle is not detected, either $[f_{t-1}, f_{t-1}]$ or $[f_{t-1}, f_{t-2}]$ are more often found than cases associated with no co-movements at all. Similarly, for series from the group without any co-movement, there is a very small percentage to find both synchronous and non-synchronous cycles.

[1] In line with what we do in the empirical application, we treat the factor f_t as observed.

[2] We report only the results relative to the BIC since it outperforms both the Akaike and Hannan–Quinn information criteria. Results are available on request.

Table 1
Selection results: Wald type t-deletion.

Model	$T = 50$			$T = 100$			$T = 250$		
	Y_{1t}	Y_{2t}	Y_{3t}	Y_{1t}	Y_{2t}	Y_{3t}	Y_{1t}	Y_{2t}	Y_{3t}
f_{t-1}	**0.59**	0.11	0.06	**0.63**	0.07	0.03	**0.61**	0.04	0.01
f_{t-2}	0.11	0.03	0.06	0.06	0.01	0.03	0.02	0.00	0.01
y_{t-1}	0.04	0.13	**0.59**	0.01	0.04	**0.66**	0.00	0.01	**0.71**
y_{t-2}	0.04	0.01	0.05	0.01	0.00	0.02	0.00	0.00	0.01
$[f_{t-1}, f_{t-2}]$	0.04	0.01	0.01	0.06	0.01	0.01	0.10	0.00	0.00
$[f_{t-1}, y_{t-1}]$	0.07	**0.48**	0.05	0.09	**0.63**	0.06	0.13	**0.70**	0.06
$[f_{t-1}, y_{t-2}]$	0.06	0.01	0.00	0.07	0.01	0.00	0.07	0.00	0.00
$[f_{t-2}, y_{t-1}]$	0.01	0.09	0.06	0.02	0.05	0.06	0.00	0.02	0.06
$[f_{t-2}, y_{t-2}]$	0.01	0.00	0.00	0.01	0.00	0.00	0.00	0.00	0.00
$[y_{t-1}, y_{t-2}]$	0.00	0.01	0.06	0.00	0.01	0.07	0.00	0.00	0.08
$[f_{t-1}, f_{t-2}, y_{t-1}]$	0.01	0.04	0.03	0.02	0.07	0.04	0.03	0.11	0.04
$[f_{t-1}, f_{t-2}, y_{t-2}]$	0.01	0.00	0.00	0.01	0.00	0.00	0.02	0.00	0.00
$[f_{t-1}, y_{t-1}, y_{t-2}]$	0.01	0.06	0.02	0.01	0.07	0.00	0.01	0.09	0.00
$[f_{t-2}, y_{t-1}, y_{t-2}]$	0.00	0.01	0.01	0.00	0.01	0.01	0.00	0.00	0.01
$[f_{t-1}, f_{t-2}, y_{t-1}, y_{t-2}]$	0.00	0.01	0.00	0.00	0.02	0.01	0.01	0.03	0.01

4. Empirical application

4.1. Clustering European economies on business cycle co-movements

This first subsection applies our new approach for detecting the existence of various kinds of co-movements. We use growth rates of the gross domestic product of 24 EU member states. We select European countries for which at least 50 observations were available at the time we took the data, excluding consequently Greece, Romania and Malta. The quarterly seasonally adjusted series from 1997Q1 to 2011Q1 are taken from the Eurostat short-term indicators database. The series are plotted in Fig. 1. In addition to the presence of similarities that seems obvious from that graph, it also emerges that the growth rates of most countries exhibit very parsimonious individual ARMA structures, e.g. ARMA(1,1). This feature can signify the presence of many co-movements. Indeed, using the final equation

Table 2
Selection results: BIC.

Model	$T = 50$			$T = 100$			$T = 250$		
	Y_{1t}	Y_{2t}	Y_{3t}	Y_{1t}	Y_{2t}	Y_{3t}	Y_{1t}	Y_{2t}	Y_{3t}
f_{t-1}	**0.66**	0.14	0.06	**0.78**	0.12	0.03	**0.85**	0.07	0.01
f_{t-2}	0.12	0.04	0.06	0.07	0.02	0.03	0.03	0.00	0.01
y_{t-1}	0.05	0.17	**0.69**	0.02	0.08	**0.82**	0.00	0.02	**0.92**
y_{t-2}	0.04	0.02	0.05	0.02	0.02	0.03	0.00	0.00	0.01
$[f_{t-1}, f_{t-2}]$	0.03	0.01	0.00	0.03	0.00	0.00	0.03	0.00	0.00
$[f_{t-1}, y_{t-1}]$	0.04	**0.47**	0.04	0.04	**0.65**	0.03	0.13	**0.82**	0.02
$[f_{t-1}, y_{t-2}]$	0.03	0.01	0.00	0.04	0.00	0.00	0.05	0.00	0.00
$[f_{t-2}, y_{t-1}]$	0.02	0.07	0.04	0.00	0.05	0.03	0.04	0.03	0.02
$[f_{t-2}, y_{t-2}]$	0.01	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
$[y_{t-1}, y_{t-2}]$	0.00	0.01	0.04	0.00	0.00	0.03	0.00	0.00	0.01
$[f_{t-1}, f_{t-2}, y_{t-1}]$	0.00	0.02	0.02	0.00	0.02	0.00	0.00	0.03	0.00
$[f_{t-1}, f_{t-2}, y_{t-2}]$	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
$[f_{t-1}, y_{t-1}, y_{t-2}]$	0.00	0.03	0.00	0.00	0.04	0.00	0.00	0.03	0.00
$[f_{t-2}, y_{t-1}, y_{t-2}]$	0.00	0.01	0.00	0.00	0.00	0.00	0.00	0.00	0.00
$[f_{t-1}, f_{t-2}, y_{t-1}, y_{t-2}]$	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00

Author's personal copy

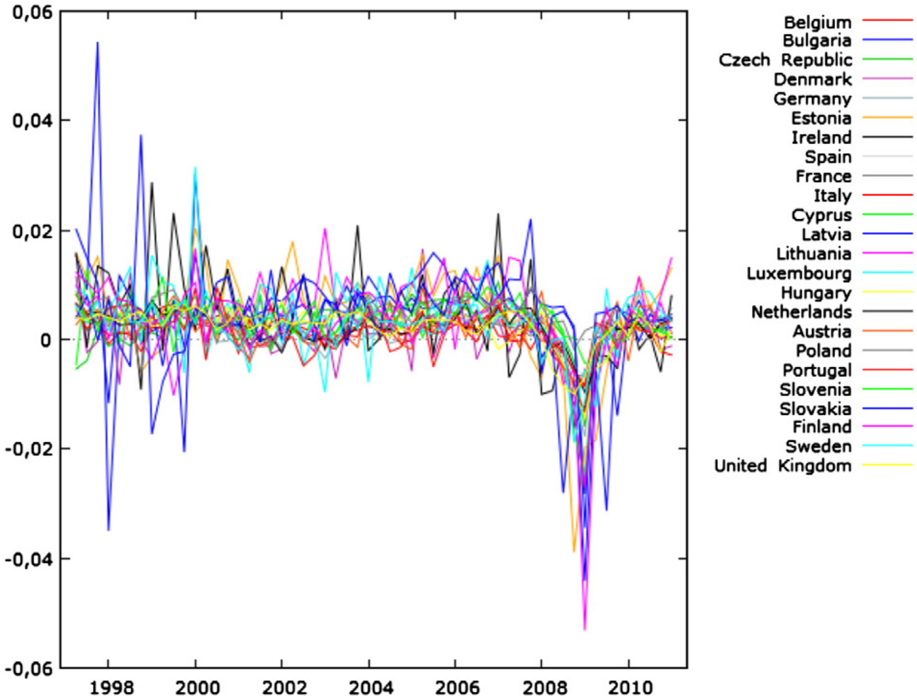

Fig. 1. European country growth rates.

representation of a N-dimensional VAR(p) model, each series can be written as an ARMA (Np, $(N-1)p$) model (Zellner and Palm, 1974). However, Cubadda et al. (2009) show that the presence co-movements implies more parsimonious ARMA structures. As an example, s SCCF vectors yield individual ARMA ($(N-s)$ p, $(N-s)$ p)) models.

In order to implement our new approach, we need to build a factor that represents the aggregate dynamics of the country growth rates. In principle, this factor can be obtained either from official statistics or by means of some statistical models. We consider five potential factors. The first one, denoted *Eur*, is the growth rate of the total GDP of the 24 considered countries. Since *Eur* is, *strictu* sensu, a non-linear aggregate of the individual growth rates, we also consider a linear approximation of *Eur*, denoted as $\widehat{Eur}$, which is obtained by a least-squares projection of *Eur* on the 24 country growth rates. The third factor, denoted *PC*, is the first principal component of the country growth rates. Stock and Watson (2002) provide the conditions under which *PC* is a consistent estimator of the common factor when both the sample size T and the number of series N diverge. The fourth factor, denoted *PLS*, is the first partial least squares factor, whose weights are obtained as the eigenvector corresponding to the largest eigenvalue of $\Gamma'\Gamma$, where Γ is the matrix of the covariances between elements of Y_t and Y_{t-1} after having standardized them to unit variance. Cubadda and Guardabascio (2012) discuss the conditions under which *PLS* is a consistent estimator of the common factor when only the sample size T diverges. Moreover, Cubadda and Hecq (2011) provide evidence that *PLS* are capable to identify a common cycle even when the sample size is small compared to the number of series. However, this approach would be invalid if the common cycle affects some countries in a non-synchronous fashion. Hence, we propose a variant of the method by Cubadda and Hecq (2011) that can handle series generated by models as Eq. (17). In particular, the procedure, which is similar in the spirit to the switching algorithm suggested by Centoni et al. (2007) to jointly test for common trends and common cycles, goes as follows:

I Standardize individual elements of both Y_{t-1} and Y_t and obtain an initial estimate $\hat{\omega}$ of the factor weights $\omega = (\omega_1,\ldots,\omega_N)'$;

II For fixed $\omega = \hat{\omega}$, obtain an estimate $\hat{\alpha}$ of the AR coefficients $\alpha = (\alpha_1,\ldots,\alpha_N)'$ in Eq. (17) by regressing y_{it} on y_{it-1} and $f_{t-1} \equiv \omega' Y_{t-1}$ for $i = 1,2,\ldots, N$;

III For fixed $\alpha = \hat{\alpha}$, obtain $\hat{\omega}$ as the eigenvector corresponding to the largest eigenvalue of $\Gamma'_{*}\Gamma_{*}$, where Γ_{*} is the matrix of the covariances between elements of Y_t^* and Y_{t-1}, and Y_t^* is a N-vector such that its i-th element is equal to $(y_{it} - \alpha_i y_{it-1})$;

IV. Repeat II and III until numerical convergence occurs.

We label the factor obtained by the above procedure as *PLS-AR*. The following matrix

	Eur	$\widehat{Eur}$	*PC*	*PLS*	*PLS−AR*
Eur	1				
$\widehat{Eur}$	0.999	1			
PC	0.970	0.970	1		
PLS	0.974	0.974	0.999	1	
PLS−AR	0.973	0.973	0.999	0.999	1

gives the correlation coefficients of these five potential factors, which are obtained from the growth rates (not the levels) of the 24 variables. It is obvious that all these indicators convey very similar information on the aggregate cycle of the considered countries. Hence, we use the observed variable *Eur* as the factor f_t in the subsequent analysis.

Within the general to specific selection procedure, we also test for the presence of outliers in the variables. In particular, for each country we estimate the FAAR model (5) with 2 lags of both the factor and the country itself. We then test for residual autocorrelation through the application of the Ljung–Box Q test and, if the null is rejected at 5% significance level, additional lags are added. This is the case for Spain, for which three lags of the variables are needed to whiten the residuals. In each FAAR model, outliers are identified as those observations whose residuals are larger than 2.5 times the residual standard deviation. We add an impulse dummy into the model for each of the previously identified outliers and then apply the general to

Author's personal copy

specific approach to remove redundant regressors (both variable lags and impulse dummies). In line with the results of the Monte Carlo analysis, we finally select the model that minimizes the BIC among all those that are nested within the unrestricted model.[3]

In order to validate the selected models, we reply the test for residual autocorrelation. Moreover, for most of countries one cannot reject the hypothesis that errors are Gaussian and homoskedastic. The final model for each country and the associated outliers are reported in Table 3.

From Table 3 we distinguish four different groups of countries:

1. The first lag of the factor, f_{t-1}, is the only relevant explanatory variable for Austria, France, Germany, Italy, Luxembourg, The Netherlands, Poland, Portugal, Slovenia and Sweden. It follows that these countries share a synchronous common cycle;
2. The first lag of the factor as well as some lags of the country are significant regressors for Bulgaria, Denmark, Estonia, Finland, Hungary, Ireland, Latvia, and Slovakia. Hence, these countries share a non-synchronous common cycle among them and with countries from Group 1;
3. Two lags of the factor as well as some lags of the country are included in the final model of Cyprus and Spain. It follows that these countries may share a PSCCF that involves lags of both the country and factor. Since the economic interpretation of such a case is tedious, we will not consider these countries in the subsequent phase of the analysis.
4. No common cycle is found for all remaining countries: Belgium, the Czech Republic, Lithuania and the United Kingdom. Indeed, country lags are the only relevant explanatory variables for these countries.

Overall, these findings seem to corroborate the conclusions of the numerous previous studies (see De Haan et al. (2008) for a recent survey) which document that, starting from the 90s, business cycle synchronization in the Euro area has increased. Indeed, the empirical evidence suggests that 10 countries, including those traditionally seen as the European core, display a synchronous common cycle.

4.2. Building the "best" business cycle indicator

In this subsection we show how to construct a coincident business cycle indicator that takes into account the presence of the various forms of co-movements among countries. The proposed procedure is based on two steps. First, countries from the second group are "aligned" in order to make their common cycle synchronized with the one of the first group. Second, we exploit the common cycle property in order to build the "best" composite coincident indicator.

The procedure goes as follows. For countries from the first group, we can easily compute the normalized SCCF coefficients $-\frac{\beta_i}{\beta_{N_1}}$, for $i = 1,2,\ldots N_1$, where N_1 represents the number of countries from this group. Hence, we obtain a $(N_1 \times N_1 - 1)$ SCCF matrix δ with the following structure:

$$\delta' = \begin{pmatrix} 1 & 0 & 0 & 0 & \ldots & -\frac{\beta_1}{\beta_{N_1}} \\ 0 & 1 & 0 & 0 & \ldots & -\frac{\beta_2}{\beta_{N_1}} \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \ldots & 1 & -\frac{\beta_{N_1-1}}{\beta_{N_1}} \end{pmatrix}. \quad (22)$$

[3] The impulse indicator saturation method (see Hendry et al., 2008) can in principle be used as an alternative to the large outlier approach that we follow here. However, in the specific case of the present empirical application, the large outlier approach provides models that are more parsimonious and economically meaningful than those obtained through impulse dummy saturation. A possible explanation of this outcome is that variability in our sample is dominated by a specific episode, that is the 2008 financial crisis.

Table 3
Selection of single equation models.

Country	BIC selected model	Outliers
Austria	f_{t-1}	[2001:Q2]
Belgium	y_{t-1}	[2008:Q4]
Bulgaria	f_{t-1}, y_{t-1}	[]
Cyprus	$f_{t-1}, f_{t-2}, y_{t-1}, y_{t-2}$	[]
Czech Republic	y_{t-1}	[2009:Q1]
Denmark	$f_{t-1}, y_{t-1}, y_{t-2}$	[2005:Q2]
Estonia	$f_{t-1}, y_{t-1}, y_{t-2}$	[2008:Q4]
Finland	f_{t-1}, y_{t-1}	[2009:Q1]
France	f_{t-1}	[2008:Q4; 2009:Q2]
Germany	f_{t-1}	[2009:Q1; 2009:Q2]
Hungary	f_{t-1}, y_{t-1}	[]
Ireland	f_{t-1}, y_{t-1}	[]
Italy	f_{t-1}	[2009:Q2]
Latvia	$f_{t-1}, y_{t-1}, y_{t-2}$	[1998:Q1; 2008:Q3]
Lithuania	y_{t-1}	[2009:Q1]
Luxemburg	f_{t-1}	[2000:Q1]
Netherlands	f_{t-1}	[1999:Q1]
Poland	f_{t-1}	[]
Portugal	f_{t-1}	[]
Slovakia	f_{t-1}, y_{t-1}	[]
Slovenia	f_{t-1}	[2008:Q4; 2009:Q1]
Spain	$f_{t-1}, f_{t-2}, y_{t-1}, \ldots y_{t-3}$	[]
Sweden	f_{t-1}	[2008:Q4]
United Kingdom	y_{t-1}	[2007:Q2]

Similarly, for countries from the second group, we can construct a $(N_2 \times N_2 - 1)$ matrix δ_0 with the same structure as Eq. (22), Finally, we build the composite coincident indicator as $\tilde{\delta}'_{\perp}\left[Y'_{1t}, Y^{*\prime}_{2t}\right]'$, where

$$\tilde{\delta}' = \begin{bmatrix} \delta' & 0_{(N_1-1)\times N_2} \\ 0_{(N_2-1)\times N_1} & \delta'_0 \end{bmatrix}$$

and Y^*_{2t} is the N_2-vector of the "aligned" time series of the second group, i.e., for each country of this group, the partial effect of the country lags is removed.

Table 4 reports the composite indicator coefficients for countries from the first two groups. In order to facilitate the interpretation, the coefficients refer to variables that have been standardized to unit variance.

Overall, the results suggest that France, Germany, Italy and The Netherlands are responsible for a large portion of the European common cycle.

Table 4
Composite indicator coefficients.

Country	y_t	y_{t-1}	y_{t-2}	Total
Austria	0.0595	0.0000	0.0000	0.0595
Bulgaria	0.0390	− 0.0155	0.0000	0.0235
Denmark	0.0926	− 0.0600	− 0.0250	0.0075
Estonia	0.0194	0.0053	0.0032	0.0279
Finland	0.0875	− 0.0425	0.0000	0.0450
France	0.0891	0.0000	0.0000	0.0891
Germany	0.0710	0.0000	0.0000	0.0710
Hungary	0.0386	0.0177	0.0000	0.0563
Ireland	0.0568	− 0.0212	0.0000	0.0356
Italy	0.1038	0.0000	0.0000	0.1038
Latvia	0.0394	− 0.0028	0.0197	0.0563
Luxemburg	0.0342	0.0000	0.0000	0.0342
Netherlands	0.0806	0.0000	0.0000	0.0806
Poland	0.0327	0.0000	0.0000	0.0327
Portugal	0.0360	0.0000	0.0000	0.0360
Slovakia	0.0478	− 0.0137	0.0000	0.0342
Slovenia	0.0304	0.0000	0.0000	0.0304
Sweden	0.0417	0.0000	0.0000	0.0417

Author's personal copy

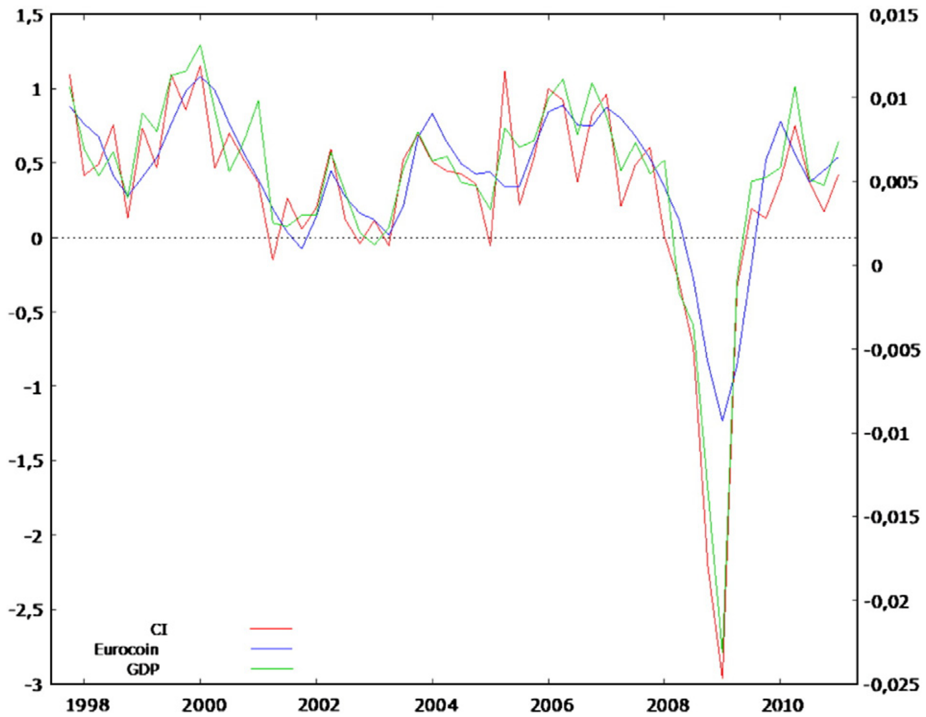

Fig. 2. Growth rates of alternative composite business cycle indicators.

Table 5
Cross-correlation functions of alternative indicators.

Lags	−4	−3	−2	−1	0	1	2	3	4
CI vs GDP	0.0716	0.2146	0.3804	0.6967	0.9621	0.6213	0.3903	0.2295	0.0308
CI vs Eurocoin	0.0515	0.3252	0.6320	0.8282	0.8397	0.6033	0.3392	0.1358	−0.0256
GDP vs Eurocoin	0.0380	0.3126	0.6390	0.8506	0.8664	0.6343	0.3471	0.1268	−0.0210

4.3. Comparison with other business cycle indicators

In this subsection, our business cycle indicator, CI, is compared with two other composite indicators: the quarterly version of Eurocoin (Altissimo et al., 2001) and the growth rate of the European Gross Domestic Product produced by Eurostat, GDP. These series are graphed in Fig. 2. Visual inspection suggests that these indicators provide a similar picture of the business cycle, although CI and GDP more strongly emphasize the deep decrease in economic activity after the 2008 financial crisis.

Table 5 reports the estimated cross-correlation functions between the considered series. These three indicators are highly cross correlated. Moreover, CI and GDP are clearly synchronous, whereas Eurocoin seems to be partially lagging with respect to the other indicators.

Although our sample size includes only a single recession, we also evaluate the performances of the various indicators in detecting the turning points of the euro area business cycle as defined by the CEPR business cycle dating committee. For this purpose, we considered the quarterly version of the Bry–Boschan rule, which states that a peak (trough) occurs when the level of a business cycle indicator attains a local maximum (minimum) relative to 2 quarters on both sides (Harding and Pagan, 2002). We see from Table 6 that the CI and GDP match the CEPR chronology, whereas Eurocoin identifies the turning points with a one quarter delay.

Table 6
Recession periods determined by alternative indicators.

CEPR	CI	GDP	Eurocoin
2008.02–2009.02	2008.02–2009.02	2008.02–2009.02	2008.03–2009.04

5. Conclusions

In this paper we have proposed a simple strategy that allows us to detect the presence and kind of co-movements in a time series system. Under plausible assumptions about the existence of a factor structure, we provide conditions for the existence of a common short-run component among variables. These conditions can easily be checked by means of automatic selection procedures. In particular, a Monte Carlo study has revealed that the BIC works quite well for this purpose. Our strategy allows for clustering time series that are homogeneous with regard to the form of their cyclical co-movements, i.e. a common synchronous cycle, a common but not synchronous cycle, and idiosyncratic short-run dynamics.

Moreover, we have shown how to construct a composite indicator that is based on the co-movements of the individual time series. First, we align those series that are affected by the common cycle in a non-synchronous fashion to construct a new set of variables that are all characterized by synchronous cyclical fluctuations. Second, a composite coincident indicator is constructed as the “most cyclical” linear combination of these variables. Both these steps are entirely performed in the time-domain since they only require the estimation of simple FAAR models.

References

Altissimo, F., Bassanetti, A., Cristadoro, R., Forni, M., Hallin, M., Lippi, M., Reichlin, L., Veronese, G., 2001. A real time coincident indicator for the Euro area business cycle. CEPR Discussion Paper, No. 3108.

Campos, J., Hendry, D.F., Ericsson, N.R., 2005. General-to-specific modeling: an overview and selected bibliography. In: Campos, et al. (Ed.), General-to-specific Modeling, vol. 1. Edward Elgar Publishing, Cheltenham.

Author's personal copy

Centoni, M., Cubadda, G., Hecq, A., 2007. Common shocks, common dynamics, and the international business cycle. Economic Modelling 24, 149–166.
Cubadda, G., 2007a. A reduced rank regression approach to coincident and leading indexes building. Oxford Bulletin of Economics and Statistics 69 (2), 271–292.
Cubadda, G., 2007b. A unifying framework for analyzing common cyclical features in cointegrated time series. Computational Statistics and Data Analysis 52, 896–906.
Cubadda, G., Guardabascio, B., 2012. A medium-*N* approach to macroeconomic forecasting. Economic Modelling 29, 1099–1105.
Cubadda, G., Hecq, A., 2001. On non-contemporaneous short-run comovements. 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, F.C., 2009. Studying interactions without multivariate modeling. Journal of Econometrics 148, 25–35.
De Haan, J., Inklaar, R., Jong-A-Pin, R., 2008. Will business cycles in the Euro area converge? A critical survey of empirical research. Journal of Economic Surveys 22, 234–273.
Engle, R.F., Kozicki, S., 1993. Testing for common features (with comments). Journal of Business and Economic Statistics 11, 369–395.
Forni, M., Hallin, M., Lippi, M., Reichlin, L., 2001. Coincident and leading indicators for the Euro area. The Economic Journal 111, 62–85.
Harding, D., Pagan, A.R., 2002. Dissecting the cycle: a methodological investigation. Journal of Monetary Economics 49, 365–381.
Hecq, A., 2005. Should we really care about building business cycle coincident indexes! Applied Economic Letters 12 (3), 141–144.
Hendry, D.F., Johansen, S., Santos, C., 2008. Automatic selection of indicators in a fully saturated regression. Computational Statistics 23 (2), 317–335 (337–339).
Issler, V., Vahid, F., 2006. The missing link: using the NBER recession indicator to construct coincident and leading indices of economic activity. Journal of Econometrics 132 (1), 281–303.
Stock, J.H., Watson, M.W., 1989. New indexes of coincident and leading economic indicators. NBER Macroeconomics Annual 351–393.
Stock, J.H., Watson, M.W., 2002. Forecasting using principal components from a large number of predictors. Journal of the American Statistical Association 97, 1167–1179.
Vahid, F., Engle, R.F., 1993. Common trends and common cycles. Journal of Applied Econometrics 8, 341–360.
Zellner, A., Palm, F.C., 1974. Time series analysis and simultaneous equation econometric models. Journal of Econometrics 2, 17–54.