ISTITUTO DI STUDI E ANALISI ECONOMICA

Building smooth indicators nearly free of end-of-sample revisions

by

Claudia Cicconi

ISAE, piazza dell'Indipendenza, 4, 00185 Rome

Working paper n. 49

April 2005

The Series "*Documenti di Lavoro" of the Istituto di Studi e Analisi Economica* – Institute for Studies and Economic Analyses (ISAE) hosts the preliminary results of the research projects carried out within ISAE. The diffusion of the papers is subject to the favourable opinion of an anonymous referee, whom we would like to thank. The opinions expressed are merely the Authors' own and in no way involve the ISAE responsability.

The series is meant for experts and policy-makers with the aim of submitting proposals and raising suggestions and criticism.

La serie "Documenti di Lavoro" dell'Istituto di Studi e Analisi Economica ospita i risultati preliminari di ricerche predisposte all'interno dell'ISAE: La diffusione delle ricerche è autorizzata previo il parere favorevole di un anonimo esperto della materia che qui si ringrazia. Le opinioni espresse nei "Documenti di Lavoro" riflettono esclusivamente il pensiero degli autori e non impegnano la responsabilità dell'Ente.

La serie è destinata agli esperti e agli operatori di politica economica, al fine di formulare proposte e suscitare suggerimenti o critiche.

Stampato presso la sede dell'Istituto

ISAE - Piazza dell'Indipendenza, 4 – 00185 Roma.
Tel. +39-06444821; www.isae.it

ABSTRACT

Aim of this paper is the construction of smooth indicator of the Italian industrial production index providing reliable end-of-sample information. Traditional smooth indicators are obtained using univariate filtering procedures based on symmetric or asymmetric filters inducing serious revisions. Here, the smoothing is obtained by exploiting the information embedded in the cross-sectional dimension which allows to use a very narrow window, reducing the need for revisions at the end of the sample. As a by-product, we also obtained a smooth composite leading indicator of the industrial sector, based on eleven selected leading sectors.

Key Words: dynamic factor models, multivariate filtering, cyclical indicators; end-of-sample revisions.

JEL Classification: C30, E32, L60.

NON-TECHNICAL SUMMARY

This work presents a smooth indicator of the Italian industry. This indicator can be used for the month-per-month analysis of the industrial activity, since it has two main properties: *(i)* it is sufficiently smooth, hence free of high frequencies dynamics, and *(ii)* provides reliable information at the very end of the sample, with respect to smooth indicators obtained by unilateral filtering.

The goal has been achieved by using monthly four-digit industrial production data to implement the multivariate filtering procedure initially proposed by Altissimo et al. (2001) as an alternative to the traditional univariate filters.

The indicator is obtained by estimating a dynamic factors model. Factor models had already been used to build a composite leading indicator of the Italian manufacturing sector. Compared to that methodology, consisting in applying the model to pre-filtered data, the approach we use has, at least in principle, three main advantages: *(i)* the estimation of the common component is one-sided, hence more accurate at the end of the sample; *(ii)* it is not distorted by the preliminary filtering procedure and, finally, *(iii)* the smoothing is obtained exploiting the information embedded in the cross-sectional dimension which only requires a very narrow window, reducing the need for revisions at the end of the sample.

In addition, the extracted signal is perfectly comparable with the one obtained trough two-sided univariate filtering in the middle-of the sample and can therefore be considered as reliable.

LA COSTRUZIONE DI UN INDICATORE CICLICO DELLA PRODUZIONE INDUSTRIALE PER L'ANALISI REAL-TIME

SINTESI

Questo lavoro presenta un indicatore del ciclo dell'attività industriale italiana che, per le sue proprietà, si presta meglio dei tradizionali indicatori ciclici ad essere utilizzato nell'analisi congiunturale. Mentre infatti questi ultimi sono caratterizzati da importanti revisioni alla fine del periodo di stima (quello di maggiore interesse per chi è interessato all'analisi di brevissimo periodo), dovuti all'utilizzo dei tradizionali filtri univariati, l'indicatore qui proposto mostra una stabilità alla fine del campione del tutto trascurabile per entità e caratteristiche.

L'indicatore, ispirato ad un precedente lavoro di Altissimo et al. (2001), è ottenuto applicando un modello dinamico a fattori al data-set costituito dalle serie degli indici della produzione di 178 settori industriali (4 cifre nella classificazione ATECO). In letteratura ci sono indicatori dell'attività industriale ottenuti dall'applicazione di modelli dinamici a fattori, ma a differenza di quelli, stimati su dati preliminarmente trattati con filtri bilaterali e quindi affetti dal consueto *end-of-sample problem*, l'indicatore qui proposto è ottenuto applicando il modello ai dati grezzi. L'estrazione della componente ciclica dell'indicatore grezzo è ottenuta sfruttando l'elevata multicollinearità che caratterizza le serie considerate. La multicollinearità riduce drasticamente l'ampiezza della finestra necessaria per estrarre la componente ciclica riducendo quindi al minimo il problema della revisione alla fine del periodo. La sostanziale coincidenza riscontrata tra i punti di svolta dell'indicatore proposto e quelli di un indicatore analogo ma ottenuto con un filtro univariato nella parte centrale del campione sembrerebbe avallare la qualità del segnale estratto con il metodo multivariato.

Parole chiave: modello dinamico a fattori, filtri multivariati, indicatori ciclici, revisioni alla fine del campione.

Classificazione JEL: C30, E32, L60.

CONTENTS

1 INTRODUCTION

This work presents a smooth indicator of the Italian industry. To use it for the month-per-month analysis of the industrial activity, this indicator should have two main properties: (i) being sufficiently smooth, hence free of high frequencies dynamics, and (ii) providing reliable information at the very end of the sample, i.e. for the period a short term analyst is interested in.

Unfortunately, matching the two goals is very difficult because obtaining a smooth indicator requires filtering the series through a sufficiently large window, which would imply a rather long period of revision of the indicator; moreover, since the theoretical filter is two-sided, the estimation is very poor at the beginning and, most importantly, at the end of the sample.

Industrial production data at a high level of disaggregation (four-digit) are monthly released and easily available. This circumstance suggested the idea of using this large data-set to implement the multivariate filtering procedure initially proposed by Altissimo et al. (2001) as an alternative to the traditional univariate filters. The dynamic factors modelling framework had already been applied to build a composite leading indicator of the Italian industrial sector. That approach, however, is different from the one proposed here: there, the factors structure is only exploited to estimate the common component of the industrial production indices (and, consequently the composite indicator) whose cyclical components are obtained, in a preliminary phase, using the traditional bilateral filters (Baxter and King or Hodrick and Prescott). Compared to that methodology, the approach we use has, at least in principle, three main advantages: (i) the estimation of the common component is one-sided, hence more accurate at the end of the sample; (ii) it is not distorted by the preliminary filtering procedure and, finally, (iii) the smoothing is obtained exploiting the information embedded in the cross-sectional dimension which requires a very narrow window, reducing the need for revisions at the end of the sample. To understand why good results are obtained here in spite of a narrow window, consider the extreme case where the length of the window is zero. Admittedly, no smoothing is obtained with the univariate procedures; conversely, in a multivariate framework, even the static projection on the common factors can provide smooth linear combinations.

The paper is organized as follows. Section 2 a basic description of the theory underlying the modelling framework adopted. Section 3 defines the composite leading indicator and describe the procedure through which smooth cyclical indicators, with the desired properties, can be obtained in this modelling framework; Section 4 briefly lists the characteristics of the data-set used in the empirical analysis, while Section 5 presents the main empirical results. Finally, Section 6 provides some concluding remarks and points out the lines along which further research on this topic is being developed.

2 THE MODELLING FRAMEWORK

2.1 The theory

The j-th time series, suitably transformed, is here assumed to be a realization from a zero-mean, wide-sense stationary process x_{jt}; moreover, all the x's are co-stationary, in the sense that stationarity holds for the n-dimensional vector process $(x_{1t},...,x_{nt})'$ for any n.

As in the traditional dynamic factors model, each variable is represented as the sum of two mutually-orthogonal unobservable components: the 'common component' and the 'idiosyncratic component'. The former is driven by a small number, say q, of 'factors' (or 'shocks') common to all variables in the system, but possibly loaded with different lag structures. By contrast, the idiosyncratic component is driven by unit-specific shocks. In traditional factor models, such component is assumed to be orthogonal to all other idiosyncratic components in the cross-section, while here, as in Forni *et al.* (2000), a limited amount of correlation is allowed for.

More formally, our model is

$$\begin{aligned} x_{jt} &= \chi_{jt} + \xi_{jt} \\ &= \sum_{h=1}^{q} b_{jh}(L)u_{ht} + \xi_{jt} \\ &= \mathbf{b}_j(L)\mathbf{u}_t + \xi_{jt} \qquad (1) \end{aligned}$$

where χ_{jt} is the common component, $\mathbf{u}_t = (u_{1t},...,u_{qt})'$ is the vector of the common shocks, i.e. a covariance-stationary process that can be assumed, with no loss of generality, to be an orthonormal white noise; $\mathbf{b}_j(L) = (b_{j1}(L)...b_{jq}(L))$ is a row vector of possibly bilateral polynomials in the lag operator L and ξ_{jt}, the idiosyncratic component, is orthogonal to $\mathbf{u}_{t-k}$ for any k. The model developed in Forni *et al.* (2000) is valid under very general dynamic loadings $b_{jh}(L)$. The problem with that model formulation is that, since $\mathbf{b}_j(L)$ is two-sided, the common component might be badly estimated at the end of the sample and should not therefore be used to forecast or as a real-time indicator.

By imposing mild additional assumptions, the common component can be obtained through one-sided filters, thus improving the quality of the estimation at the end of the sample. In particular, Forni *et al.* (2003) assume the common component is a finite-order $VARMA$ process:[1]

$$\chi_t = \mathbf{B}(L)\mathbf{A}(L)^{-1}\mathbf{u}_t \qquad (2)$$

where $\mathbf{B}(L) := \mathbf{B}_0 + \mathbf{B}_1 L + ... + \mathbf{B}_s L^s$ is a $n \times q$ polynomial of order s and $\mathbf{A}(L) := \mathbf{I} - \mathbf{A}_1 L - ... - \mathbf{A}_S L^S$ a $q \times q$ polynomial of order S with all solutions of $\det\left[\mathbf{A}(z)\right] = 0$, $z \in C$, lying outside the unit circle.

This means assuming the factor loadings $\mathbf{B}(L)$ to be of finite order s and the factor $\mathbf{f}_t$ itself to be a VAR process of order S. Note also that, writing $\mathbf{f}_t := (f_{1t}...f_{qt})'$ for $[\mathbf{A}(L)]^{-1}\mathbf{u}_t$ and $\mathbf{F}_t$ for $(\mathbf{f}_t'\mathbf{f}_{t-1}'...\mathbf{f}_{t-s}')'$, the model becomes a static factor model, in the sense that

$$\mathbf{x}_t = \mathbf{B}(L)\mathbf{f}_t + \xi_t = \mathbf{B}\mathbf{F}_t + \xi_t \qquad (3)$$

where $\mathbf{B} := (\mathbf{B}_0 \quad \mathbf{B}_1 ... \mathbf{B}_s)$ and $\mathbf{F}_t$, the vector of static factors, has dimension $r = q(s+1)$.

2.2. One-sided estimation and forecasting

Following Forni *et al*. (2003), the estimation method we propose consists of two steps[2].

Step1: COVARIANCE STRUCTURE (*Dynamic principal components analysis*).

The first step consists of the estimation of the spectral-density matrix of the common components. To begin with, we estimate the spectral density

1 See Forni et al. (2003) for a detailed description of the theoretical background.

2 The basic steps required for in-sample and out of-sample estimation of the common component are provided here without a detailed formal description of the theoretical foundation and the statistical properties of the estimators, which can be found in Forni *et al.* (2003).

matrix of $\mathbf{x}_t = (x_{1t} ... x_{nt})'$ and denote it $\hat{\Sigma}(\theta)$. Then, a dynamic principal components decomposition (see Brillinger, 1981) is performed: the eigenvalues and eigenvectors of $\hat{\Sigma}(\theta)$ are computed for each frequency θ in the grid of frequencies and the (sample) eigenvalue and eigenvector functions $\lambda_j(\theta)$ and $U_j(\theta)$ are then obtained by ordering, for each frequency, the eigenvalues, and correspondingly the associated eigenvectors, in descending order. If q is the number of common factors, according to the common-idiosyncratic decomposition outlined in Forni *et al.* (2000), a consistent estimate (as both n and T go to infinity) of the spectral density matrix of the vector of common components $\chi_t = (\chi_{1t} ... \chi_{nt})'$ is given by

$$\hat{\Sigma}_{\chi}(\theta) = \mathbf{U}_q(\theta) \Lambda_q(\theta) \tilde{\mathbf{U}}_q(\theta) \qquad (4)$$

where $\Lambda_q(\theta)$ is the diagonal matrix having $\lambda_1(\theta), ..., \lambda_q(\theta)$ on the diagonal, $\mathbf{U}_q(\theta)$ is the $n \times q$ matrix $(U_1(\theta) ... U_q(\theta))$ and □ denotes conjugation and transposition[3]. Finally, an estimate of the covariance matrix of χ_t at lag k can be obtained through the inverse discrete Fourier transform of the estimated spectral-density matrix, i.e.

$$\hat{\Gamma}_{\chi}^{k} = \frac{2\pi}{2w+1} \sum_{h=-w}^{w} \hat{\Sigma}_{\chi}(\theta_h) e^{i\theta_h k} \qquad k = \pm 1, \quad \pm 2, ... \qquad (5)$$

where w is the number of points in $(0, \pi]$ on which the spectral density matrix is estimated.

Moreover, estimates of the covariance matrices of the cyclical component χ_t^{C} can be easily obtained by applying the inverse transform to the frequency band of interest $[-2\pi/\tau, 2\pi/\tau]$. More precisely, letting $\Gamma_{\chi^C}^{k} = E(\chi_t^C \chi_{t-k}^{C'})$, its estimate will be

$$\hat{\Gamma}_{\chi^C}^{k} = \frac{2\pi}{2H+1} \sum_{h=-H}^{h=H} \hat{\Sigma}_{\chi}(\theta_h) e^{i\theta_h k} \qquad (6)$$

where H is defined by the conditions $H/(2w+1) > \tau$ and $(H+1)/(2w+1) < \tau$.

---

[3] An estimate of the spectral density matrix of the vector of the idiosyncratic components $\xi_t = (\xi_{1t} ... \xi_{nt})'$ can then be obtained as the difference $\hat{\Sigma}_{\xi}(\theta) = \hat{\Sigma}(\theta) - \hat{\Sigma}_{\chi}(\theta)$.

Step2: ESTIMATION OF THE COMMON COMPONENT (*Static principal components analysis*).

In the second step, the static factors are estimated[4]. Such estimates are obtained as the first r generalized principal components of $\hat{\Gamma}_{\chi}^{0}$ with respect to $\hat{\Gamma}_{\xi}^{0}$.

First, the generalized eigenvalues λ_j, i.e. the n complex numbers solving $\det(\hat{\Gamma}_{\chi}^{0} - \lambda\hat{\Gamma}_{\xi}^{0}) = 0$, along with the corresponding generalized eigenvectors V_j, i.e. the vectors satisfying

$$V_j \hat{\Gamma}_{\chi}^{0} = \lambda_j V_j \hat{\Gamma}_{\xi}^{0} \tag{7}$$

subject to $V_j \hat{\Gamma}_{\xi}^{0} V_i'$ equal 1 for *j* = *i* and equal 0 for $j \neq i$ are computed. Then, the eigenvalues are sorted in descending order and the eigenvectors corresponding to the largest r eigenvalues are taken. The estimated static factors are the generalized principal components[5]

$$\hat{F}_{jt} = V_j \mathbf{x}_t \qquad j = 1, 2, ..., r \tag{8}$$

Finally, χ_t, the common component, is estimated and forecast through the orthogonal projection on the static factors. Letting $\mathbf{V} = (V_1 ... V_r)$ and $\hat{\mathbf{F}}_t = (\hat{F}_{1t} \cdots \hat{F}_{rt})' = \mathbf{V}'\mathbf{x}_t$, the estimate of χ_{T+h}, given the information available at time T, is

$$\hat{\chi}_{T+h} = \hat{\Gamma}_{\chi}^{h} \mathbf{V} (\mathbf{V}' \hat{\Gamma}_0 \mathbf{V})^{-1} \hat{\mathbf{F}}_T \tag{9}$$

Within the sample, the described method yields the projection

---

4 The static factors cannot be identified in the model unless additional assumptions are introduced. What we are looking for, indeed, is (an estimate of) a basis of the vector space spanned by the static factors, i.e. by the u_{ht}'s and their lags; therefore, we shall estimate a vector of linear combinations of such factors.

5 The rationale, and the motivation, behind this strategy is that, given the estimated variance-covariance matrices $\hat{\Gamma}_{\chi}^{0}$ and $\hat{\Gamma}_{\xi}^{0}$, the V_j's maximize the common-to-idiosyncratic variance ratio of the estimated factors $\hat{F}_j$'s.

$$\hat{\chi}_t = \hat{\Gamma}_{\chi}^{0}\mathbf{V}(\mathbf{V}'\hat{\Gamma}_0\mathbf{V})^{-1}\hat{\mathbf{F}}_t \qquad (10)$$

which is a one-sided estimate of the common component that, for fixed T, avoids the end-of-sample inconsistency problems proper of two-sided estimation. Both (1) and (2) are consistent estimators of the corresponding population quantities $\chi_{T+h/T} = E(\chi_{T+h} \setminus I_T)$ and χ_t [6].

3 THE CONSTRUCTION OF A SMOOTH LEADING INDICATOR

3.1 Definition and properties

In the estimation of cyclical indicators of economic activity, the use of large panel of time series has, at least in principle, a great advantage with respect to univariate or small multivariate models, since it allows to obtain an indicator fulfilling two very important requirements for an indicator to be useful and economically meaningful:

(i) CROSS-SECTIONAL SMOOTHING. The index is cleaned from the idiosyncratic component of the industrial production indexes mainly capturing sectoral specific shocks and measurement error which would add misleading information;

(ii) INTERTEMPORAL SMOOTHING. To construct a cyclical indicator the common component of the industrial production indexes should also be cleaned from temporary changes and from the trending component to unveil the underlying cyclical dynamics of the economy.

This can be done by band-passing the series, i.e. by applying to the series a filter which only passes the components lying in a certain frequency band, Ω^C. Actually,

$$y_t^{\ c} = d(L)y_t = \sum_{h=-\infty}^{+\infty} d_h y_{t-h} \qquad (11)$$

6 See Forni et al. (2003).

where $d^C(L)$ is a two-sided, symmetric, infinite-order, square-summable filter whose h-th coefficient is such that

$$d_h = \int_{-\pi}^{\pi} \beta(\omega) e^{ih\omega} d\omega \qquad (12)$$

where $\beta(\omega) = 1$ for $\omega \in \Omega^c$ and 0 otherwise.

The sample counterpart of the filter is normally obtained either by truncating the filter $d^c(L)$, as in Baxter and King (1999), or by implementing the data-dependent filter suggested by Christiano and Fitzgerald (2001).

If that indicator is built for real-time short-term analysis, it should also be well estimated at the very end of the sample. Unfortunately, as stressed before, applying bilateral filters ends up in quite large and long lasting revision of the filtered series at the end of the sample.

Factor models can exploit the superior information embedded in the cross-sectional dimension providing a good temporal smoothing with a very short filter. Because of the filter shortness, the filtered series are only marginally affected by revisions at the end of the sample and, therefore, can be used to construct real-time indicators.

As a by-product, we propose a smooth composite leading indicator (SCLI) of the Italian industrial production defined as a weighted average of the cyclical common components of the industrial production indexes of selected leading sectors

$$SCLI_t = \frac{\sum_{l \in L} w_l \chi_{lt}^c}{\sum_{l \in L} w_l} \qquad (13)$$

where L is the set of selected leading sectors.

3.2 Estimation of SCLI

The estimation of the composite leading index consists of three distinct phases.

Phase 1: LEADING/LAGGING RELATIONSHIPS BETWEEN SECTORS AND THE REFERENCE CYCLE

Our reference cycle is the cyclical common component of the general industrial production index. The assessment of the leading/lagging relationship between the industrial sectors and the reference cycle is based on the estimated variance-covariance matrix of the common component at cyclical frequencies obtained as the inverse discrete Fourier transform of the common spectral density matrix on a proper frequency band (see previous Section). Industrial sectors have been first classified as pro- or counter-cyclical according to their phase angle with respect to the reference series evaluated at zero frequency. A variable is classified as pro-cyclical if the phase is zero (positive mean lag) and as counter-cyclical if the phase is π (negative mean lag). Besides, variables may be classified into lagging, coincident or leading. For pro-cyclical variables, attention is paid on the time displacement of the maximal positive correlation with the reference cycle: a series is classified as leading when such time displacement is smaller than -2, lagging when it is grater than 2 and coincident otherwise; the same criterion is used for counter-cyclical variables, the displacement being that of the minimal negative correlation.

Phase 2: INTERTEMPORAL SMOOTHING

The ingredients needed for the estimation of the cyclical components are provided by the two-steps estimation procedure described in the previous Section. In fact, the cyclical component of the i-th sector is obtained as the projection of χ_{it}^{C} on the present, m leads and m lags of the static factors, with projection coefficients derived by the covariance matrices of the cyclical components.

Precisely, setting

$$\mathbf{W} = \begin{pmatrix} \mathbf{V} & 0_{n\times r} & \cdots & 0_{n\times r} \\ 0_{n\times r} & \mathbf{V} & \cdots & 0_{n\times r} \\ \vdots & \vdots & \ddots & \vdots \\ 0_{n\times r} & 0_{n\times r} & \cdots & \mathbf{V} \end{pmatrix} \qquad (14)$$

and $\mathbf{x}_t^{*'} = (\mathbf{x}_{t+m}^{'} \ldots \mathbf{x}_t^{'} \ldots \mathbf{x}_{t-m}^{'})^{'}$, $\hat{\mathbf{F}}_t = \mathbf{W}^{'}\mathbf{x}_t^{*} = (\hat{F}_{t-m}^{'}, \cdots, \hat{F}_t^{'}, \cdots, \hat{F}_{t+m}^{'})$. The variance-covariance matrix of $\mathbf{x}_t^{*}$ is $\mathbf{M} = \begin{pmatrix} \hat{\Gamma}_0 & \hat{\Gamma}_1 & \cdots & \hat{\Gamma}_{2m} \\ \hat{\Gamma}_1^{'} & \hat{\Gamma}_0 & \cdots & \hat{\Gamma}_{2m-1} \\ \vdots & \vdots & \ddots & \vdots \\ \hat{\Gamma}_{2m}^{'} & \hat{\Gamma}_{2m-1}^{'} & \cdots & \hat{\Gamma}_0 \end{pmatrix}$ while $E(\chi_t^{C} \chi_{t-m}^{C'})$ can be estimated by $\mathbf{R} = (\hat{\Gamma}_{\chi^C}^{-m'} \ldots \hat{\Gamma}_{\chi^C}^{0'} \ldots \hat{\Gamma}_{\chi^C}^{+m'})$.

Finally, an estimate of the cyclical components is

$$\hat{\chi}_t^C = \mathbf{RW}(\mathbf{W}'\mathbf{MW})^{-1}\hat{\mathbf{F}}_\mathbf{t} \tag{15}$$

At the end of the sample, i.e. from $T-m$ onward, we have the problem that $\mathbf{x}_{T+m}$, $m>0$, is not available. Our estimate is then obtained by substituting the forecast of the common component, $\hat{\chi}_{T+m}$ (see (1)), to the unknown $\mathbf{x}_{T+m}$ and by applying the same formula.

Notice that, when $m=0$, (7) becomes $\hat{\chi}_t^C = \hat{\Gamma}_{\chi^C}^0\mathbf{V}(\mathbf{V}'\hat{\Gamma}_0\mathbf{V})^{-1}\hat{\mathbf{F}}_t$, which is the formula used in the application.

Phase 3: AGGREGATION

The composite leading indicator of the Italian industrial production is defined as the weighted average of the cyclical common component of the industrial production indexes of selected leading sectors where the weights are those assigned by ISTAT to each sector according to the ATECO 2002 classification scheme.

4 THE DATA-SET

The data-set considered consists of the time series of four-digit industrial production indexes plus the general industrial production index from 1990:1 to 2003:6. We only consider 178 four-digit indexes, which account for 97,5% of total industrial production according to the ISTAT weighting scheme[7]. Data were transformed in order to achieve stationarity applying the twelfth differences of the logarithm of the original series on the basis of preliminary unit root tests.

---

[7] The sectors excluded are 'sugar', because the series is too irregular, 'other products of wood', 'mortars, electronic valves and tubes and other electronic components', 'jewellery and related articles n.e.c.', 'metal secondary raw materials' and 'non-metal secondary raw materials', because the series start in 2001:01.

5 EMPIRICAL RESULTS

5.1 Multivariate vs univariate filtering

By applying bilateral filters, the estimate of the cyclical component at time T made at time T, $\hat{x}^{c}_{T/T}$ will be revised until the definitive estimate of x^{c*}_{T}, $\hat{x}^{c}_{T/T*}$ is obtained at time $T^{*} > T$. In univariate filters T^{*} is normally much greater than T.

As an example, Fig.1 shows succeeding estimates of the Baxter and King cyclical component of the general industrial production index corresponding to periods between 1,5 and 8 years. The first estimation is done using data till 2002:7, the second till 2002:8 and so on, the last using data till 2003:6. As we can see, even for $m = 8$, (m is the window length) revisions are important. Moreover, Fig.1 shows that, in order to get a sufficient degree of smoothness, m must be equal to 16, which implies revisions for a quite long period. This circumstance would be particularly awful if the aim of the analyst is the real time assessment of the business cycle because, using these filters, she will be able to have an idea of what the business cycle was at time *T* only with a large delay. As Fig.1 shows, the revisions in the estimate of the cycle component widened the amplitude of the last downturn and postponed the last turning point.

The use of cross-sectional information in the factors models framework reduces the filter length (we will show results for $m = 0$), thereby shortening the number of months needed to reliably asses the cyclical position at a certain point in time.

Fig.2 shows the normalized series (upper box) and the spectral densities (lower box) of the (12-month differences of the) industrial production index and of its cyclical component resulting from pure cross-sectional smoothing ($m = 0$), according to the procedure illustrated above. Indeed, even though the signal extracted is less smooth than the one obtained with the traditional filter and a large window, it is nonetheless cleaned from too high frequency dynamics.

As an example of the advantages obtained by using this procedure, two experiments are provided, namely:

a: pre-filtering the series using Baxter and King with $m = 16$ to obtain the cyclical components and applying the Generalized Dynamic Factors Model (GDFM) to clean them from the idiosyncratic noise, thus obtaining the reference cycle and the leading sectors as the common components of the input series;

Figure 1 The trade-off between smoothness and the length of the window.

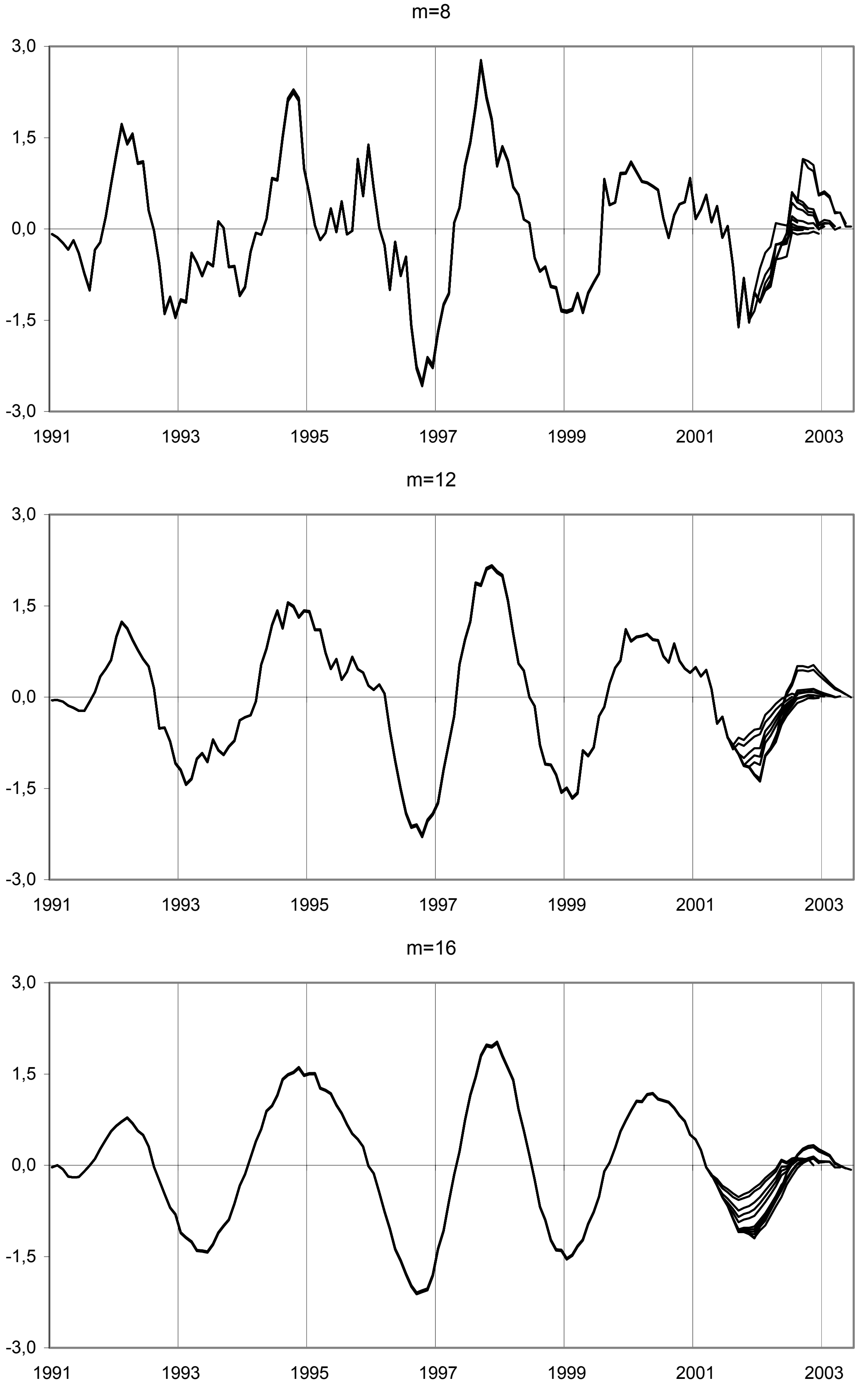
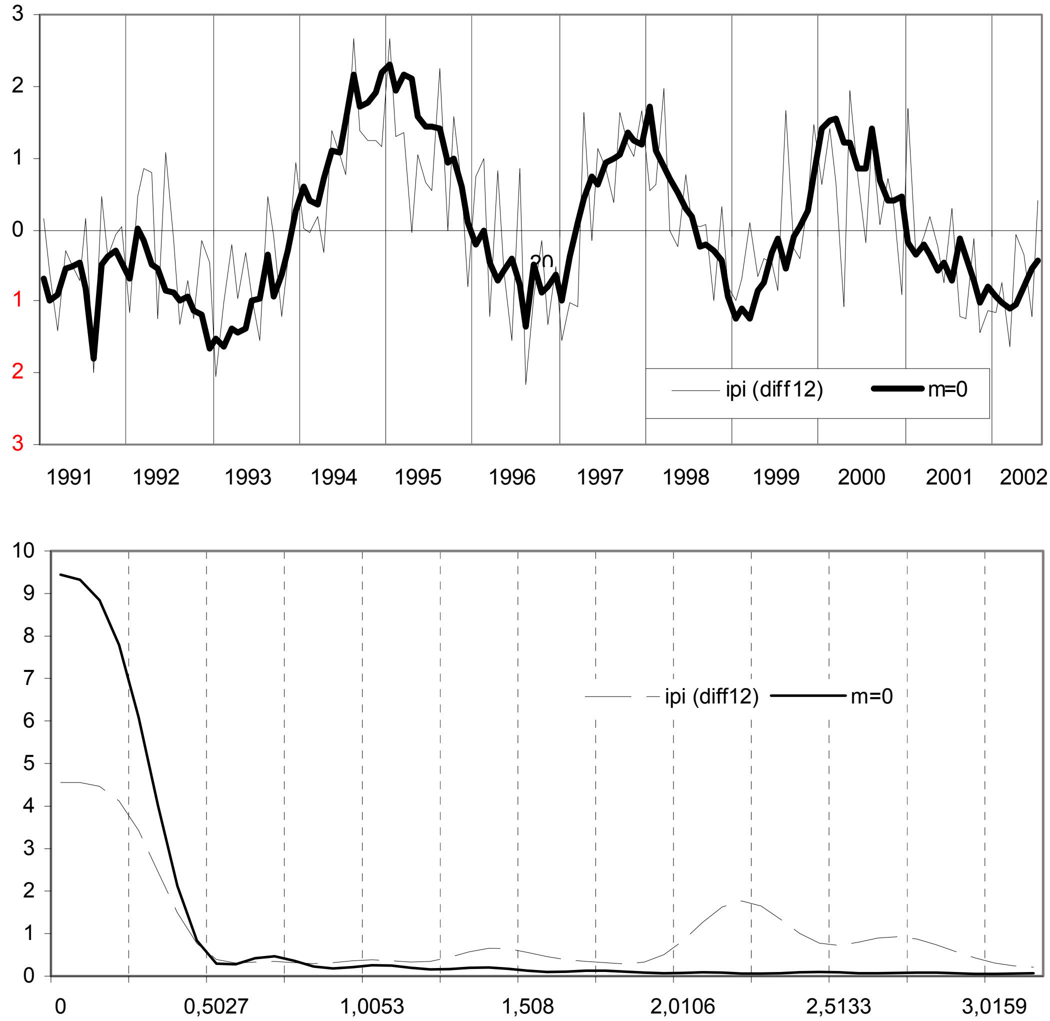


Figure 2 Series and spectral densities of raw and smoothed data



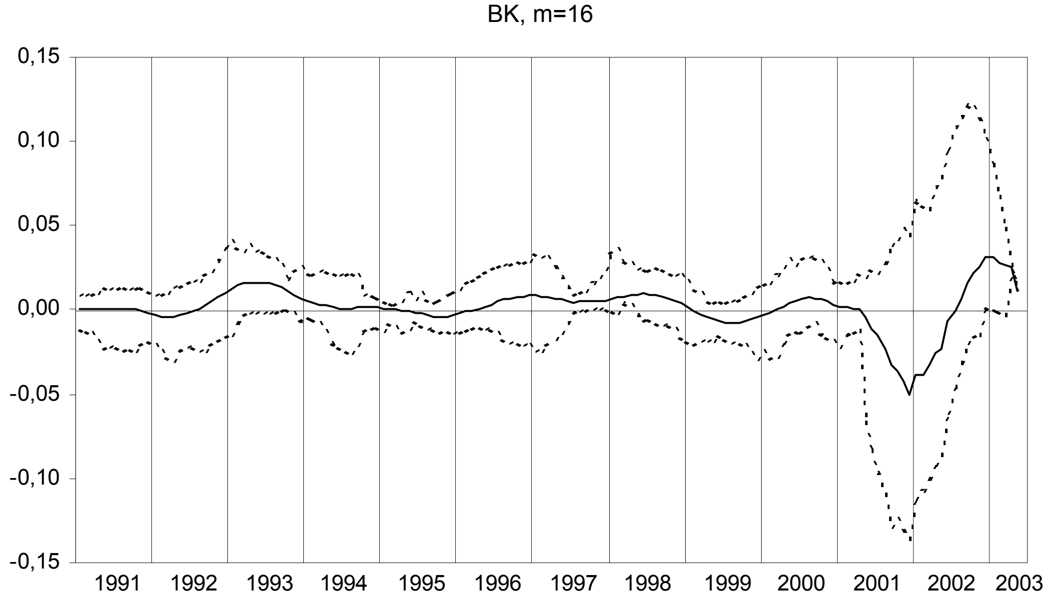
b: applying GDFM to the raw data to extract both their common and their cyclical components.

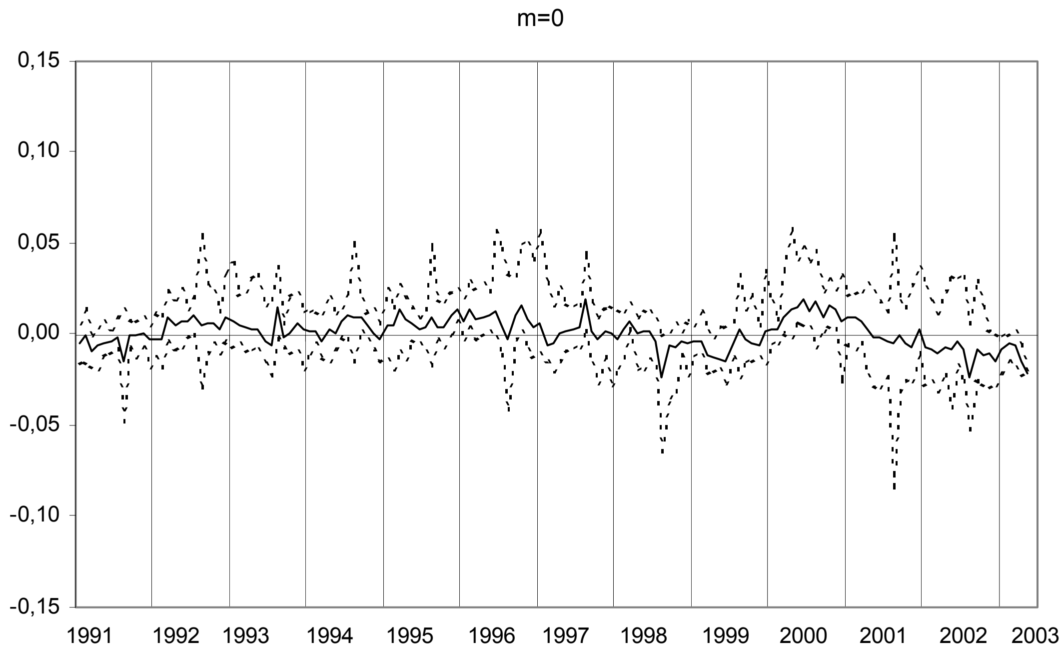
Both experiments are repeated 12 times. The first time using observations till 2002:7, the second time using observations till 2002:8 and so on, till the last one which is based on observations till 2003:6. For each date in the sample period (1991:1 - 2002:7), we have 12 estimates of the cyclical component[8]. The assessment of the extent of the end of sample revisions will be based on the period 2001:04 - 2002:07[9]).

[8] Obviously, for each of the periods after July 2002, we have, respectively, eleven, ten, ..., one estimates.

[9] Actually, a Baxter and King filter of length 16 provides definitive estimates of the cyclical component till February 2002 (the estimate obtained using all the observations available, i.e. till June 2003). Observations till November 2003 would provide definitive estimates of the cyclical component till July 2002.

Figure 3 The magnitude of the revisions over the sample period





The results obtained on the end-of-sample revisions concern the cyclical common component of the general industrial production index, that is our reference series. No comparison of the leading indicators is considered here, as it would entail the analysis of factors, other than the cyclical smoothing, which are outside the scope of this section.

Being a filter *a la* Baxter and King with $m=16$ the benchmark, the entire sample period appears to be composed by two subperiods: the period 1991:01-

2001:03, in which the revisions are not due, by construction, to the smoothing procedure[10]; the period 2001:4-2002:07 in which the revisions are also due to the filtering procedure. Fig. 3 shows the maximum, the minimum and the average revisions occurred running the two experiments.

Figure 3 shows that:

- both methods lead to similar revisions in the first sub-period where the errors are not caused by the smoothing procedure;
- pure cross-sectional smoothing does not entail particular end-of sample problems, the revisions being of the same magnitude throughout the whole sample period;
- pre-filtering with Baxter and King's filters induces very high revisions at the end of the sample compared with the other method.

Table 1 below confirms this evidence, showing that while for non pre-filtered data the extent of the revision is more or less the same in the two subperiods, for pre-filtered data the extent of the revision is particularly high in the second subperiod as emerges from the average difference between the last and the first estimation of the indicator for the same date and from the average of the absolute value of the revisions.

Table 1 **The magnitude of the revisions**

	in sample (jan-91 to mar-01)		end-of-sample (apr-01 to feb02)	
	m=0	**m=16**	**m=0**	**m=16**
Last-First [1]	0.07	0.05	0.04	0.29
Max-Min [2]	0.08	0.07	0.07	0.35
abs rev [3]	0.01	0.01	0.01	0.04

[1] Average of the absolute valued difference between the last (jun03) and the first (jul02) estimate of the cyclical component at time t.

[2] Average of the difference between the maximum and the minimum of the revisions of the estimate of the cyclical component at time t.

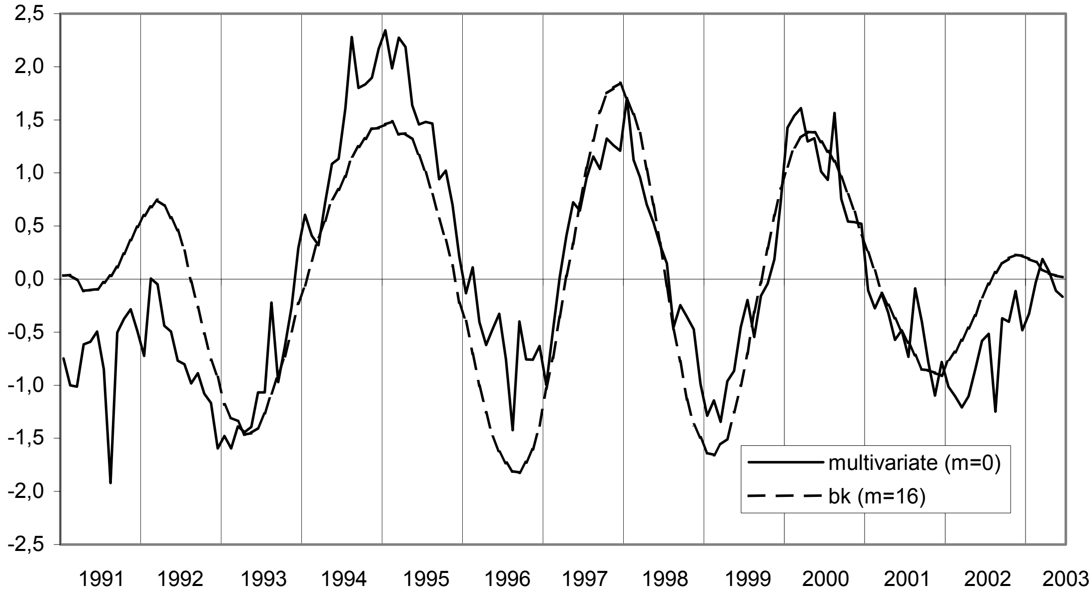
[3] Average of the mean absolute valued revisions.

[10] In-sample revisions may be due to the revision of the raw series (negligible for industrial production data) or to the centered nature of the output. The latter, however, decreases in magnitude with the length of the sample period and does not influence the cyclical property of the results since it affects all the observations in the same way.

Finally, Figure 4 shows that, in-sample, the turning points of the cyclical component extracted by exploiting the cross-section are perfectly 'coincident' with the ones obtained through the univariate filtering procedure, which is generally thought to provide the right signal. This can be considered an indirect proof of the fact that our cyclical component is not distorted.

Figure 4 The quality of the signal extracted with the multi-variate procedure

Definite estimates (june 2003)



5.2 The correlation structure

The estimation of the spectral density matrix of our panel of time series and of its dynamic principal components provides important information on its correlation structure, on the structural relationship among sectors and between sectors and the industry as a whole.

The series are strongly cross-correlated, as shown by the fact that the first two principal components capture more than 50% of their variance and even more at business cycle frequencies. Commonality is very heterogeneous across sectors. However, the correlation between the general industrial production index and the sectoral ones is always higher when considering common components rather than of the original series. Since the first two dynamic principal components account for about 90% of total variance for the general industrial production index, by considering common components of sectoral indexes means focusing on that part of the series which is more informative on the behaviour of the industrial production index.

5.3 The composite leading indicator

The leading-lagging relationships between sectors and the industry as a whole have been analyzed to assess whether it is possible to identify those sectors regularly anticipating the cyclical behaviour of the general industrial production index published by ISTAT, about 45 days after the end of the reference month. The sectors classification in leading, lagging and coincident was done according to the criteria described above. The results should, however, be taken with some caveats: a sector might anticipate the upturns and not the downturns or the lead may not be constant over time since the criteria used to classify the sectors (correlations at business cycle frequencies) do not enable an assessment of those characteristics. Indeed, to overcome this shortcoming, only those leading sectors that showed a regular lead across the whole sample period were chosen to enter the composite leading indicator. Moreover, sectors having a very large lead as well as those poorly correlated with the reference series or showing a low commonality were excluded. As a result, only 11 of the initial 46 leading sectors were used to construct the indicator.

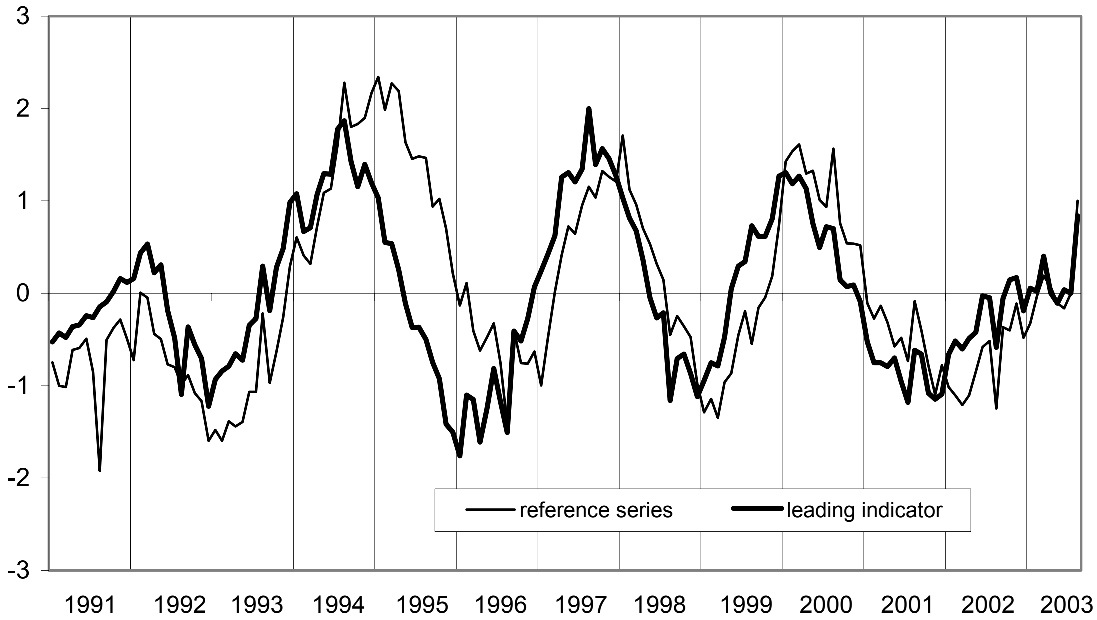
Table 2 **Selected leading sectors**

Code	Description	Peak corr	Lead
15.43	Margarine and similar edible fats	0,83	5
15.71	Prepared animal feeds for farm animals	-0,66	5
15.82	Rusk and biscuits; preserved pastry goods and cakes	0,60	3
15.93	Wines	0,69	7
21.11	Pulp	0,84	4
21.12	Paper and paperboard	0,74	5
24.17	Synthetic rubber in primary forms	0,88	4
25.24	Other plastic products	0,73	3
26.24	Technical ceramic wares	-0,61	5
29.21	Furnaces and furnace burners	0,59	4
29.53	Machinery for food, beverage and tobacco processing	0,55	6

Four leading indexes belongs to the food and beverages sector, which apparently reflects no economic rationale. However, it is worth noticing that, even among machinery, one of the leading sector is the production of food machines. The remaining indexes included in the composite indicator belong to the intermediate goods (five sectors) sector and to the capital goods (one sector). Their main characteristics may be found in Table 2.

In Figure 5, the indicator is plotted against the reference series. The indicator leads the reference series of four months on average over the whole sample period. It seems that the time lead is longer at the beginning of the sample (first six/seven years) and shorter at the end. However, the lagged correlations computed over the two subperiods show that the average time lead is constant over the sample period. Indeed, the procedure adopted to get the cyclical components of the indexes does not exclude the possibility of phase shifts between the original and the filtered series. However, an *a posteriori* analysis (of the lagged correlations) shows that this is not the case either for each of the 11 leading sectors, or for the general industrial production index.

Figure 5 **The composite leading indicator**



6 CONCLUSIONS

These pages explored the possibility of using the four-digit series of the industrial production indexes to obtain a smooth indicator of the industrial activity. In particular, exploiting the cross-sectional dimension proved to be very useful as it provided smooth indicators, whose estimates do not require revisions at the end of the sample. In addition, in-sample, the extracted signal is perfectly comparable with the one obtained trough two-sided univariate filtering and can therefore be considered as reliable. As a byproduct, a composite leading indicator of the industrial activity was obtained, based on 11 sectors, showing an average four-month lead as against the reference cycle.

Results on the degree of smoothness of the indicator and on the stability of the leading-lagging relationships between sectors can be improved. One of the weakness of the estimation procedure influencing the stability and reliability of the leading lagging relationships and the results in terms of smoothness is its sensitivity to outliers. Two alternative solutions are being explored: the former consists in pre-cleaning the series from outliers, as it is generally done in this type of application; the latter, which should induce smaller distortions in the data, consists in using a robust estimation method of or the covariance matrices. A further extension includes the possibility of using more aggregated data for industrial production (3-digit instead of 4-digit) and adding selected variables of different nature which are considered useful in assessing the development of the Italian industrial activity.

REFERENCES

Altissimo, F. et al. (2001), "Eurocoin: A Real Time Coincident Indicator of the Euro Area Business Cycle", CEPR, Discussion Paper n. 3108;

[Baxter, M. and R. G. King (1999), "Measuring Business Cycles: Approximate Band-Pass Filters For Economic Time Series", The Review of Economics and Statistics, Vol. 81 (4), pp. 575-593;

Christiano, L. and T. J. Fitzgerald (1999), "The Band Pass Filter", NBER, Working Paper n. 7257;

Forni, M. et al. (2000), "The Generalized Dynamic Factor Model: Identification and Estimation", The Review of Economics and Statistics, Vol. 82 (4), pp. 540-554;

Forni, M. et al. (2003), "The Generalized Dynamic Factor Model, One-Sided Estimation and Forecasting, downloadable from http://www.dynfactors.org.

Working Papers available:		
n. 25/02	M. CAGIANO DE AZEVEDO	Recent Developments in the Environmental Debate Before and after the Kyoto Protocol: A Survey
n. 26/02	M. BOVI	The Nature of the Underground Economy. Some Evidence from OECD Countries
n. 27/02	R. BASILE	Acquisition versus Greenfield Investment: the Location of Foreign Manufacturers in Italy
n. 28/02	G. BRUNO M. MALGARINI	An Indicator of Economic Sentiment for the Italian Economy
n. 29/02	G. ARBIA R. BASILE M. SALVATORE	Regional Convergence in Italy 1951-1999: a Spatial Econometric Perspective
n. 30/03	F. KOSTORIS PADOA SCHIOPPA	Il mercato e le politiche economiche in Italia
n. 31/03	S. DE NARDIS C. VICARELLI	The Impact of Euro on Trade: the (Early) Effect Is not So Large
n. 32/03	S. LEPROUX	L'inchiesta ISAE-UE presso le imprese del commercio al minuto tradizionale e della grande distribuzione: la revisione dell'impianto metodologico
n. 33/03	G. BRUNO C. LUPI	Forecasting Euro-area Industrial Production Using (Mostly)\ Business Surveys Data
n. 34/03	C. DE LUCIA	Wage Setters, Central Bank Conservatism and Economic Performance
n. 35/03	E. D'ELIA B. M. MARTELLI	Estimation of Households Income from Bracketed Income Survey Data
n. 36/03	G. PRINCIPE	Soglie dimensionali e regolazione del rapporto di lavoro in Italia
n. 37/03	M. BOVI	A Nonparametric Analysis of the International Business Cycles
n. 38/03	S. DE NARDIS M. MANCINI C. PAPPALARDO	Regolazione del mercato del lavoro e crescita dimensionale delle imprese: una verifica sull'effetto soglia dei 15 dipendenti

Working Papers available:		
n. 39/03	C. MILANA ALESSANDRO ZELI	Productivity Slowdown and the Role of the Ict in Italy: a Firm-level Analysis
n. 40/04	R. BASILE S. DE NARDIS	Non linearità e dinamica della dimensione d'impresa in Italia
n. 41/04	G. BRUNO E. OTRANTO	Dating the Italian Business Cycle: a Comparison of Procedures
n. 42/04	C. PAPPALARDO G. PIRAS	Vector-auto-regression Approach to Forecast Italian Imports
n. 43/04	R. DE SANTIS	Has Trade Structure Any Importance in the Transmission of Currency Shocks? An Empirical Application for Central and Eastern European Acceding Countries to EU
n. 44/04	L. DE BENEDICTIS C. VICARELLI	Trade Potentials in Gravity Panel Data Models
n. 45/04	S. DE NARDIS C. PENSA	How Intense Is Competition in International Markets of Traditional Goods? The Case of Italian Exporters
n. 46/04	M. BOVI	The Dark, and Independent, Side of Italy
n. 47/05	M. MALGARINI P. MARGANI B.M. MARTELLI	Re-engineering the ISAE manufacturing survey
n. 48/05	R. BASILE A. GIUNTA	Things change. Foreign market penetration and firms' behaviour in industrial districts: an empirical analysis