Extracting a Common Cycle from Series with Different Frequency

An Application to the Italian Economy

Edoardo Otranto*

Abstract

The extraction of a common signal from a set of time series is generally obtained using variables recorded with the same frequency or transformed in order to have the same frequency (monthly, quarterly, etc.). The econometric literature has not paid a great deal of attention to the study of alternative approaches. In this paper we extend an approach based on the use of dummy variables to the well known trend plus cycle model, in a multivariate context, using both quarterly and monthly data. This procedure is applied to the Italian case, using the variables considered by the Institute for Studies and Economic Analyses (ISAE) to provide a national dating. The results are compared with the ones obtained with the equivalent multivariate and univariate approaches using monthly data. The use of both quarterly and monthly data provides more consistent results with the ISAE ones than other approaches.

Key Words: Business cycle, State-space model, Time series, Trend, Turning points

JEL Classification: C32, C50

* Dipartimento di Economia, Impresa e Regolamentazione, University of Sassari, Italy; eotranto@uniss.it

I am grateful to two anonymous referees for their precious suggestions and I would like to thank Y. Murasawa and T. Proietti to have provided their codes to estimate the models described in Mariano and Murasawa (2003) and Proietti and Moauro (2006) respectively. I acknowledge the use of the RATS routine developed by G. Bruno for the Bruno and Otranto (2005) work. A particular thank goes to A. Gandin and R. Iannaccone for the technical support.

Résumé

L’extraction d’un signal commun d’un ensemble de séries chronologiques s’opère généralement à partir de variables présentant la même fréquence ou ramenées à la même fréquence (mensuelle, trimestrielle, etc.). Cette contrainte n’est guère remise en question dans la littérature économétrique. Dans cet article, l’auteur propose une approche du modèle bien connu « tendance + cycle », reposant sur l’utilisation de variables indicatrices dans un contexte multivarié, se fondant sur des données trimestrielles et mensuelles. Il applique cette procédure à l’économie italienne en utilisant les variables suggérées par un organisme de ce pays (l’ISAE) pour obtenir un datage et compare ses résultats à ceux des techniques multivariée et univariée équivalentes appliquées sur des données mensuelles uniquement. Il constate que l’utilisation simultanée de données trimestrielles et mensuelles aboutit à des résultats plus proches de ceux qui ressortent des chiffres officiels que les autres approches.

1 Introduction

In the statistical analysis the extraction of common signals, as a common trend or a common cycle, from a set of time series, is generally obtained using series with the same frequency (monthly, quarterly and so on). If the available data are recorded at different frequencies, for example a first group of monthly series and a second one of quarterly series, one of the two sets is transformed to obtain series with the same frequency. Simple aggregation methods are used to transform the monthly series into quarterly series; disaggregation is used to transform the quarterly series into monthly ones. In Italy, for example, the Institute for Studies and Economic Analyses (ISAE) extracts the common cycle from two quarterly and four monthly time series, transforming the quarterly into monthly series with a redistribution algorithm. The method developed by Altissimo *et al.* (2000) is then applied to create a coincident indicator.

The Kalman filter routines contain different procedures. For example, Azavedo *et al.* (2003) insert the GNP quarterly series with other monthly indicators in a state-space model, using the STAMP routines (Koopman *et al.*, 2000); at each step of the Kalman filter, the quarterly series are forecasted 2-periods ahead, creating artificial data.

The possibility to work with both types of data has not received an adequate attention in the statistical and econometric literature, maybe because the results derived from forecasting techniques are considered as a good approximation of the reality. Lately, Mariano and Murasawa (2003) dealt with this problem applying a model of the Stock and Watson (1991) type to estimate a coincident indicator. In their work, the missing monthly values of the quarterly series are not estimated, but are put in the Kalman filter without interpolations, using dummy variables.

The main purpose of this work is to apply the idea of Mariano and Murasawa (2003) to extract a common component from a set of time series with different frequency; in particular, we deal with the six series used by ISAE, extending this approach to the trend plus cycle model (see, for example, Harvey, 1985, 1990). This is one of the most used models in the literature, because of its flexibility and the possibility to have the well known Hodrick-Prescott filter (Hodrick and Prescott, 1997) as a particular case in the univariate framework. Extending this model to the multivariate case, we obtain a kind of multivariate Hodrick-Prescott filter.

This type of model has some advantages compared to a model of the Stock and Watson type. First of all, the main aim of the latter is the construction of a coincident indicator of the business cycle, whereas the other components of the series are included in the idiosyncratic parts. The multivariate trend plus cycle model provides a common cyclical indicator, and also separate trends and irregular components for each series. In other words, we propose a model that allows the extraction of the common signal and the decomposition of the single time series at the same time.

Second, the Stock and Watson model requires that the coincident series are $\mathrm{I}(1)$ but not cointegrated. Our model can be more flexible, because it is also possible to suppose a common trend, to some or all the variables, when cointegration is present. This is not the case of this paper (in our application the data are not cointegrated), but it is an important capability of this model.

We compare the results obtained in our model with similar approaches based on monthly data only. In particular we will consider the analogous multivariate model with monthly data and the synchronization of the cycles derived by six univariate trend plus cycle models (Harding and Pagan, 2006). On the other hand, we will estimate a Stock and Watson model, using the Mariano and Murasawa (2003) approach, to verify the existence of some differences in terms of detection of turning points. The data concern the period January 1972 to September 2002 (the same span used in the work of Bruno and Otranto, 2005, who compare alternative models with monthly data). Our empirical analysis shows that some variables seem to have more importance for the extraction of the common signal than others.

In the next section we describe the proposed model, emphasizing the technique based on dummy variables to avoid artificial data; in Section 3 we develop the application on the Italian economy, applying the four alternative models. Final remarks follow.

We will indicate with $\mathbf{I}_h$ the identity matrix of dimension $h \times h$ and with $\mathbf{0}_{h,s}$ a matrix of dimension $h \times s$ with all the elements equal to zero.

2 The Model

Let us consider n_1 time series recorded with frequency s_1 and n_2 with frequency s_2 (s_1, s_2 equal to 1 if the series are annual, 4 if the series are quarterly, 12 if they are monthly and so on); we suppose that $s_1 > s_2$. The aim is to extract a common cycle from these $n_1 + n_2 = n$ series. Let us indicate with $y_i(t)$ the i-th time series ($i = 1,...n$) observed at time t ($t = 1,...,T$). Supposing that the index t is referred to the s_1 frequency, we adopt a simple additive trend plus cycle model for each component (Harvey, 1985):

$$y_i(t) = \mu_i(t) + \psi(t) + \varepsilon_i(t), \tag{1}$$

in which $\mu_i(t)$ represents the proper stochastic trend of the variable i, $\psi(t)$ is the cycle common to all the variables, $\varepsilon_i(t)$ are Independent Identically Normally (IIN) distributed disturbances with zero mean and unknown variance σ_i^2 (constant for all t); moreover we suppose that the cycle is the only common element among the variables, so that the n trends are considered mutually independent, as well as the n series of disturbances.

The hypothesis of independent trends could appear too strong because it implies no cointegration among the variables. However many models aiming to extract a common cycle do not make use of cointegrated series (see, for example, Stock and Watson, 1991, Mariano and Murasawa, 2003, Proietti and Moauro, 2006). On the other side, if we can not adopt this hypothesis, it is possible to modify the model, considering also a common trend to all or some variables. This model will not be discussed in this paper because in the application part (Section 3) the variables are not cointegrated.

The trends and the common cycle are unobserved variables, which follow dynamics that are expressed by separate equations; each trend follows a linear model as:

$$\begin{aligned} \mu_i(t) &= \mu_i(t-1) + \beta_i(t-1) + \eta_i(t) \\ \beta_i(t) &= \beta_i(t-1) + \varsigma_i(t) \end{aligned} \tag{2}$$

where $\beta_i(t)$ is the slope of the trend and $\eta_i(t)$ and $\varsigma_i(t)$ are uncorrelated IIN disturbances with zero mean and variances respectively equal to δ_i^2 and ν_i^2. It is equivalent to an ARIMA(0,2,1) process. If $\delta_i^2 = \nu_i^2 = 0$ the trend is deterministic, whereas, if $\nu_i^2 = 0$ and $\delta_i^2 > 0$, the model is equivalent to a random walk with drift. The case with $\nu_i^2 > 0$ and $\delta_i^2 = 0$ represents a stationary trend in the second differences and it has the characteristic to be relatively smooth, which is a generally accepted idea of a trend component; the well known Hodrick-Prescott filter corresponds to a model as (1) without $\psi(t)$, with $\delta_i^2 = 0$ and the ratio ν_i^2/σ_i^2 fixed (Harvey and Jaeger, 1993). For example, Hodrick and Prescott (1997) show the value 1/1600 for quarterly series; some authors suggest to use other values (for example, Pedersen, 2001) or to estimate this ratio (for example, Otranto and Iannaccone, 2005). In this case, the cycle is included in the disturbance $\varepsilon_i(t)$. In our application we will estimate all the parameters.

Modelling the cyclical component explicitly, we take (Harvey, 1985):

$$\begin{aligned} \psi_t &= \rho\left[\cos(\lambda)\psi(t-1) + \sin(\lambda)\psi^*(t-1)\right] + \omega(t) \\ \psi^*(t) &= \rho\left[-\sin(\lambda)\psi(t-1) + \cos(\lambda)\psi^*(t-1)\right] + \omega^*(t) \end{aligned} \tag{3}$$

where $\psi^*(t)$ is an unobservable variable, $0 \le \lambda \le \pi$ is the frequency of the cycle, $0 \le \rho \le 1$ is a damping factor on the amplitude of the cycle; $\omega(t)$ and $\omega^*(t)$ are uncorrelated IIN disturbances with zero mean and the same variance κ^2 (this assumption is not forced, because generally it does not cause a real loss in goodness of fit; see Harvey, 1985).

The n equations expressed by (1) can be grouped in the vector $\mathbf{y}(t)$, whereas the trends and the slopes respectively in the vectors $\boldsymbol{\mu}(t)$ and $\boldsymbol{\beta}(t)$, the disturbances in the vector $\boldsymbol{\varepsilon}(t)$. A compact way to express these relationships is the following state-space model:

$$\begin{aligned} \mathbf{y}(t) &= \mathbf{A}\boldsymbol{\xi}(t) + \varepsilon(t) \\ \boldsymbol{\xi}(t) &= \mathbf{B}\boldsymbol{\xi}(t-1) + w(t) \end{aligned} \tag{4}$$

where the unobservable vector state is given by:

$$\boldsymbol{\xi}(t) = \left[\boldsymbol{\mu}(t)' \quad \beta(t)' \quad \psi(t) \quad \psi^*(t)\right]'.$$

The fixed matrices $\mathbf{A}$ and $\mathbf{B}$ are expressed by:

$$\mathbf{A} = \begin{bmatrix} \mathbf{I}_n & \mathbf{0}_{n,n} & \mathbf{c} & \mathbf{0}_{n,1} \end{bmatrix}; \qquad \mathbf{B} = \begin{bmatrix} \mathbf{I}_n & \mathbf{I}_n & \mathbf{0}_{n,1} & \mathbf{0}_{n,1} \\ \mathbf{0}_{n,n} & \mathbf{I}_n & \mathbf{0}_{n,1} & \mathbf{0}_{n,1} \\ \mathbf{0}_{1,n} & \mathbf{0}_{1,n} & \rho\cos(\lambda) & \rho\sin(\lambda) \\ \mathbf{0}_{1,n} & \mathbf{0}_{1,n} & -\rho\sin(\lambda) & \rho\cos(\lambda) \end{bmatrix},$$

whereas $\boldsymbol{\varepsilon}(t)$ and $\mathbf{w}(t)$ are $n \times 1$ and $2n \times 1$ vectors, containing respectively the disturbances in (1) and those in (2)-(3); they are mutually uncorrelated and IIN with zero means and covariance matrices $\boldsymbol{\Sigma}$ and $\mathbf{Q}$, expressed by diagonal matrices with elements respectively given by:

$$\boldsymbol{\sigma}^2 = \begin{bmatrix} \sigma_1^2, & \sigma_2^2, & \cdots & \sigma_n^2 \end{bmatrix}; \qquad \mathbf{q}^2 = \begin{bmatrix} \delta_1^2, & \cdots & \delta_n^2, & \nu_1^2, & \cdots & \nu_n^2, & \kappa^2, & \kappa^2 \end{bmatrix}.$$

We obtain the case of a smooth trend when the first n elements of $\boldsymbol{q}^2$ are equal to zero. The $n \times 1$ vector $\mathbf{c}$, contained in $\mathbf{A}$, is composed by n constants, representing the weights of the common cycle $\psi(t)$ to model the single equation (1). Another approach is the one of Harvey and Koopman (1997), named similar cycle model, where each equation in (1) has a proper cyclical component $\psi_i(t)$, but the n cycles have the same damping factor ρ and frequency λ. As one of the target of this paper is the extraction of a common component, we prefer to adopt our specification, allowing $\mathbf{c}$ to differentiate the presence of the common cycle on the single series. For identification, one of the elements of $\mathbf{c}$ has to be equal to 1.

Now, let us suppose that the first n_1 variables contained in $\mathbf{y}(t)$ are recorded with frequency s_1 and the remaining n_2 with frequency s_2. To simplify, let us also suppose that the last n_2 variables are stock variables, so that their values represent the total amount of the variable at that time (which is the case of the successive application of Section 3). It is possible to use the hypotheses adopted by Mariano and Murasawa (2003) for the case of flow variables.

Following Mariano and Murasawa (2003), we consider the n_2 variables with lowest frequency as variables recorded with frequency s_1 with missing values. For example, let $s_1 = 12$ (monthly frequency) and $s_2 = 4$ (quarterly frequency). In addition, let $x^*(t)$ be one of the n_2 quarterly series; then, it is observed at time $t,\ t+3,\ t+6,\ t+12$, ..., whereas values are missing for the other dates. To avoid the estimation of missing values, we can suppose that, for all t:

$$x(t) = \begin{cases} x^*(t) & \text{when } x(t) \text{ is observable} \\ z(t) & \text{otherwise} \end{cases}$$

were $z(t)$ are random variables IIN with distribution not depending on unknown coefficients. Using this hypothesis, the missing values will not affect the maximum likelihood estimators because $z(t)$ and $\mathbf{y}(1)$, $\mathbf{y}(2)$, ... $\mathbf{y}(T)$ are independent by construction. In this case the likelihood function L can be rewritten as:

$$L[\rho,\lambda,\kappa^2,\boldsymbol{\sigma}^2,\boldsymbol{q}^2 \mid \boldsymbol{y}^*(1),\boldsymbol{y}^*(2),...\boldsymbol{y}^*(T)] = L[\rho,\lambda,\kappa^2,\boldsymbol{\sigma}^2,\boldsymbol{q}^2 \mid \boldsymbol{y}(1),\boldsymbol{y}(2),...\boldsymbol{y}(T)]\prod_{i\in M} f[z(i)],$$

where M denotes the set of time instants in which the quarterly data are not observed and $\mathbf{y}^*(t)$ is the vector containing the n variables, with the last n_2 elements missing if $t \in M$, equal to $\mathbf{y}(t)$ otherwise. In other words, the likelihood function of the unknown parameters given the full data set $\mathbf{y}^*(1)$, $\mathbf{y}^*(2)$, ... $\mathbf{y}^*(T)$, is equivalent to the likelihood of the same parameters given the only data observed $\mathbf{y}(1)$, $\mathbf{y}(2)$, ... $\mathbf{y}(T)$ up to scale. As $z(t)$ does not affect the estimation procedure, we suppose, as in Mariano and Murasawa (2003), that f is the Normal distribution with mean $\mathbf{0}_{n_2,1}$ and covariance matrix $\mathbf{I}_{n_2}$ and that its realizations in our data set are always equal to zero.

The state-space representation is the same as (4), but the matrix $\mathbf{A}$, when $x(t)$ is not observable, will change in:

$$\mathbf{A}_1 = \begin{bmatrix} \mathbf{I}_{n_1} & \mathbf{0}_{n_1,n+n_2} & \mathbf{c}_1 & \mathbf{0}_{n_1,1} \\ & \mathbf{0}_{n_2,2(n+1)} & & \end{bmatrix},$$

where $\mathbf{c}_1$ is an $n_1 \times 1$ vector with the elements equal to the first n_1 elements of $\mathbf{c}$, whereas the covariance matrix of $\varepsilon(t)$ will change in a diagonal matrix $\mathbf{\Sigma}_1$ with elements given by:

$$\begin{bmatrix} \sigma_1^2, & \cdots & \sigma_{n_1}^2, & 1, & \cdots & 1 \end{bmatrix}.$$

Let:

$$y = \begin{cases} 1 & \text{when all the } n \text{ variables are observed} \\ 0 & \text{otherwise} \end{cases}$$

The final model can be written as:

$$\begin{aligned} &\mathbf{y}(t) = [\gamma\mathbf{A} + (1-\gamma)\mathbf{A}_1]\boldsymbol{\xi}(t) + \varepsilon(t) \\ &\boldsymbol{\xi}(t) = \mathbf{B}\boldsymbol{\xi}(t-1) + w(t) \\ \\ &\varepsilon(t) \sim \text{IIN}(\mathbf{0}_{n,1}, \gamma\mathbf{\Sigma} + (1-\gamma)\mathbf{\Sigma}_1) \\ &w(t) \sim \text{IIN}(\mathbf{0}_{2(n+1),1}, \mathbf{Q}) \end{aligned} \tag{5}$$

Note that the dummy variable γ is not present in the state equation, so that the trends and the common cyclical components are estimated for each time t.

3 Extracting the Italian Business Cycle

In this section, we use the method proposed in this paper – hereafter named the Quarterly-Monthly Multivariate Model (QMMM – to extract a common cycle and to get a dating. We will compare this dating with the ISAE results.[1] Then, we estimate the same multivariate model, but using monthly variables (disaggregating the quarterly series). In addition we estimate six univariate models. Finally, we will compare the trend plus cycle models with the Stock and Watson model, using both quarterly and monthly data.

These structural models can be considered as ARIMA models, but we do not use the Box and Jenkins (1976) approach to select and to validate them, as the structural models are subject to restrictions imposed by *a priori* evaluations, which could be not consistent with the information deriving by the correlogram (see Harvey, 1985 for details). So we prefer to compare the models in terms of detection of turning points.

To establish the dates of turning points, we adopt the automatic Bry and Boschan (1971) procedure for all the time series. Briefly, this non-parametric procedure can be applied to a single monthly time series, adjusted for seasonality, and it consists in the extraction of the points identified as local maxima/minima and satisfying certain censoring rules (see Bry and Boschan, 1971, for details). In the following sub-sections we describe shortly the data used, the other methods and finally we compare the results.

3.1 The data used

The six (seasonally adjusted) variables used are:

1. monthly index of industrial production (total industry excluding construction);
2. monthly quantity of goods transported on railways (in tons);
3. monthly percentage of overtime hours in large industrial firms;
4. monthly imports of investment goods (quantity);
5. quarterly investments in machinery and equipment at constant prices;
6. quarterly value added of the service sector, excluding mainly non-market branches (education, health services, public administration) at constant prices.

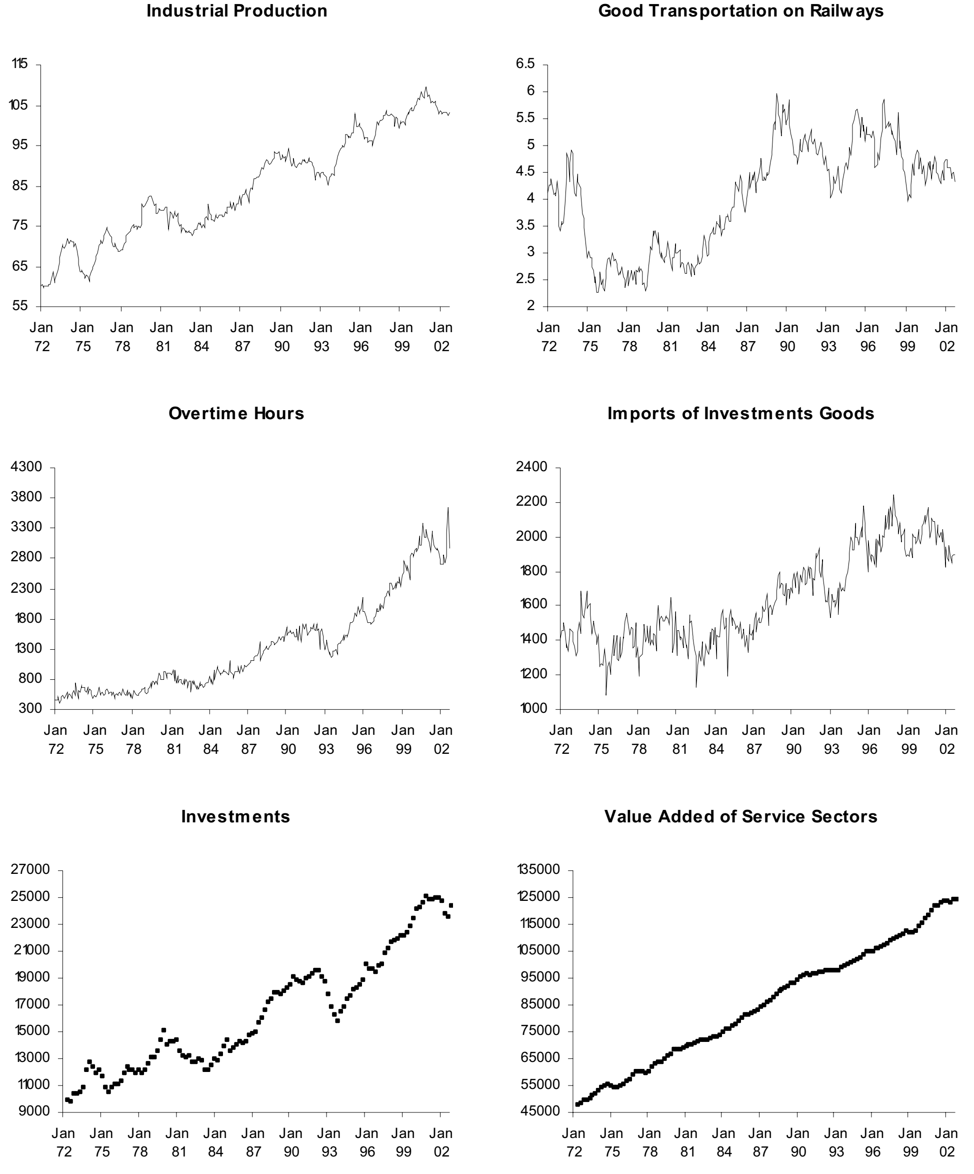
The source of the seasonally adjusted data, shown in Figure 1, is the Italian National Statistical Institute (Istat).[2] These data seem to have similar periods of expansion and contraction, but different trends. Furthermore their variability is very different; the overtime hours and the import of investment goods show a strong irregular component compared to

1 In Italy, the ISAE has established a business cycle dating, based on the NBER methodology.

2 The data were seasonally adjusted with the TRAMO-SEATS routine (Gómez and Maravall, 1997). In the TRAMO-SEATS decomposition an additive model was considered. For this reason we will not adopt the logarithmic transformation.

the other series (that means we expect large variances for the corresponding disturbances of equation (4)).

Figure 1 **Seasonally adjusted series**



Altissimo *et al.* (2000) have selected these variables from a set of 183 time series referring to the Italian economy, using a stepwise procedure with several restrictions. The choice is based on the coincidence behavior and the property to represent various aspects of economic activity (in fact the series selected represent the supply side including the service sector, the demand side and the labour market).

The last two variables are disaggregated in monthly frequencies with the procedure based on the Fernandez (1981) model.[3] This procedure assumes that the monthly data are generated by a random walk:

$$y^m(t) = y^m(t-1) + u(t)$$

where $u(t) \sim \mathrm{IIN}(0, \sigma^2)$. The quarterly data $y^q(t)$ are assumed to be observed without error. Moreover, the higher frequency data sum up to the lower frequency values across every quarter. Then the procedure estimates (with maximum likelihood) the $y^m(t)$'s which produce the right $y^q(t)$'s.

Note that Altissimo *et al.* (2000) did not insert a typical coincident variable, such as the Gross Domestic Product (GDP), among the variables selected. In this work we accept their choices, referring to them for other details; they use the GDP to express the judgemental aggregation of the turning points of the single series.

3.2 The ISAE procedure

The ISAE procedure is based on the NBER methodology. In practice, the turning points of the six series selected are detected with the Bry and Boschan procedure. The dating for the whole economy is obtained by aggregating the turning points of the single time series, based on a judgemental assessment. The results of this automatic procedure and the judgemental assessments supplied by business cycle experts, provide the Italian business cycle turning points.[4] This aspect is of paramount importance because it is possible to separate the actual from the apparent turning points; at the same time the limit of this procedure is that it needs the subjective judgement of a group of experienced business cycle analysts. In this way it is very difficult to replicate the ISAE results with a purely statistical procedure, but this dating can be assumed as a benchmark to evaluate the other methods proposed in this work. Bruno and Otranto (2005) have used the same six variables to evaluate several parametric and nonparametric procedures to extract the business cycle turning points in an automatic way for the period January 1972 to September 2002. Their results show that the methods provide similar results compared to the ISAE chronology in the period 1972-1983, characterized by the two oil shocks, and in the period 1993-2002, whereas

3 The corresponding algorithm is contained in the procedure DISTRIB.SRC of the software Winrats 32 (Doan, 2001).

4 As observed in Bruno and Otranto (2005), this is not an official dating in a strict sense, but it is considered by the users as a likely picture of the Italian business cycle dynamics.

they detect various extra-cycles in the 1984-1992 period, not indicated by ISAE. As we pointed out in the Introduction, in this work we will use the same period as in Bruno and Otranto (2005). A sub-product of this procedure is a coincident indicator of the Italian cycle. Hereafter this model will be denoted by ISAE.

3.3 The monthly multivariate model

To evaluate the performance of the model (1)-(2)-(3), estimated with both monthly and quarterly data, in comparison with a classical case in which all the series have the same frequency, we have estimated the analogous model (4), or the model (5) with $\gamma = 1$ for each t, using the monthly disaggregation explained in Section 3.1. The main interest in this case is to verify if the artificial data can produce extra-cycles or loose cycles detected by the contemporaneous use of monthly and quarterly series. Hereafter the abbreviation for this model will be MMM.

3.4 Univariate indirect approach

Another possibility is to estimate six separate univariate models for the monthly series:

$$\begin{aligned}
y(t) &= \mu(t) + \psi(t) + \varepsilon(t) \\
\mu(t) &= \mu(t-1) + \beta(t-1) + \eta(t) \\
\beta(t) &= \beta(t-1) + \varsigma(t) \\
\psi(t) &= \rho\left[\cos(\lambda)\psi(t-1) + \sin(\lambda)\psi^*(t-1)\right] + \omega(t) \\
\psi^*(t) &= \rho\left[-\sin(\lambda)\psi(t-1) + \cos(\lambda)\psi^*(t-1)\right] + \omega^*(t)
\end{aligned} \tag{6}$$

Practically, in an univariate framework, model (6) is equivalent to the model (1)-(2)-(3), providing separate cycles for each series. For each cycle, we extract the turning points following the Bry-Boschan procedure; then we aggregate, using an indirect way, the turning points with the procedure of Harding and Pagan (2006). In practice, this procedure consists in finding, for every t, a 6×1 vector of distances for the nearest peak (trough) for each time series considered. The median of this vector is interpreted as the mean distance from the nearest peak (trough) for the whole economy and the local minima of this series are candidates to be a peak (trough) for the whole economy. Then, the turning points are selected to alternate and the cycles and single phases last not less than 15 and 5 months respectively. This approach is useful because it is more similar to the ISAE one, being conducted in terms of single univariate analysis, but, at the same time, it uses the same trend model of the multivariate approaches considered in this paper. Hereafter this model will be called UIA.

3.5 The Stock and Watson model with quarterly and monthly series

Finally we estimate a particular version of the Mariano and Murasawa (2003) model, partially based on the extension proposed by Proietti and Moauro (2006). They suggest to specify the Stock and Watson model in terms of levels of the variables rather than in terms of changes; this provides the advantage to have the mean squared error of the estimated coincident index immediately available in real time. Moreover, dealing with the logarithms of the time series, they modify the state space representation to account for the nonlinear temporal aggregation of the flow variables involved by the logarithms.[5] We adopt the former extension whereas the latest does not concern with our application because the quarterly series are stock variables not transformed by logarithms (see footnote 2).

The model is given by:

$$\begin{aligned} \mathbf{y(t)} &= \mathbf{b}f(t) + \mathbf{u}(t) \\ \phi_f(L)f(t) &= v_1(t) \\ \mathbf{\Phi_u}(L) &= \mathbf{v}_2(t) \end{aligned}$$

where $f(t)$ represents the coincident indicator, $\mathbf{b}$ is a (6×1) loading vector, $u(t)$ is a (6×1) vector of idiosyncratic components, $\phi_f(L)$ is a p-th-order polynomial in the lag operator L, $\mathbf{\Phi_u}(L)$ is a q-th-order diagonal matrix polynomial in L, $v_1(t) \sim \mathrm{IIN}(0, \sigma^2_{v_1})$ and $v_2(t) \sim \mathrm{IIN}(0, \mathbf{\Sigma}_{v_2})$ are mutually uncorrelated ($\mathbf{\Sigma}_{v_2}$ is a diagonal matrix). For identification we suppose that the first element of $\mathbf{b}$ is equal to 1. We have used the BIC criterion (Schwarz, 1978) to identify the polynomials order, obtaining $p = q = 1$. For details on the state space representation and the estimation procedure see Mariano and Murasawa (2003) and Proietti and Moauro (2006). Hereafter this model will be denoted by SWQM.

3.6 Empirical results

The presented models require that the n coincident series are $\mathrm{I}(1)$ but not cointegrated; thus, we have to find empirical evidence about the integration and cointegration properties of the series. The standard univariate unit root tests of Dickey and Fuller (1979) fail to reject the null hypothesis that the six series are integrated (the variables are indexed with $i = 1, ..., 6$, which corresponds to the numbering of Section 3.1). Furthermore, the null hypothesis of none cointegrating vectors is accepted,[6] so that we can use separate trends for QMMM and MMM and we can estimate the SWQM model too. Details of the tests are shown in Table 1.

---

5 Moauro and Proietti (2004) have used this method to estimate a coincident indicator for the euro area.

6 The test statistic is the Likelihood Ratio statistic described in Johansen (1991) and (1995). As in King *et al.* (1999) and Kim and Piger (2002), we assume that each series has a linear trend, whereas the cointegrating equation has only intercepts.

Table 1 **Results of the Dickey-Fuller and Johansen cointegration tests**

Dickey-Fuller test						
5% Critical Value is -2.87						
1% Critical Value is -3.45						
Series	y_1	y_2	y_3	y_4	y_5	y_6
Test statistic	-1.26	-1.92	-0.66	-1.43	0.57	0.33

Johansen cointegration test	
5% Critical Value is 94.05	
1% Critical Value is 103.18	
Null Hypothesis	Test Statistic
No Cointegration Vectors	77.93

In a preliminary analysis, the variances δ_i^2 have resulted near to zero, so we have imposed the first equation in (2) as deterministic (as in the Hodrick-Prescott procedure), but without fixing the ratios ν_i^2/σ_i^2, which we will estimate for each variable. For identification we put $c_6 = 1$. The same holds for MMM and UIA. The final estimates are shown in Table 2.[7]

The first macroscopic difference of the QMMM approach with respect to the others, is related to the estimation of the variances of the quarterly series. Anyway, this is not unexpected because the monthly transformation described in Section 3.1 reduces the difference between $\beta_i(t)$ and $\beta_i(t-1)$ for each t in (2), so that the variance of $\varsigma_i(t)$ is artificially reduced. In the other estimates, the multivariate models show similar variances in the trend components (excluding the 5th variable). The univariate models provide different variances for the trend component. The trends of each variable obtained with the three different approaches are shown in Figure 2. Note that the dynamics of the trends deriving from the multivariate approaches are very similar; the only difference can be found in the investments series, in which the MMM approach provides a more irregular trend. The univariate models show the main differences compared to the multivariate models for the original monthly variables; they have a very smooth behavior. This is due to the fact that, not being the constraint of a common cycle, the univariate models provide smooth trends, assigning large part of the variance to the irregular or cyclical components (see Table 2). The last two variables (the quarterly transformed series) show a different behavior with components similar to the one obtained from the multivariate approaches. Vice versa, the multivariate approaches assign some movements to the trend components, that are assigned to the cyclical component in the univariate approaches; of course, this is due to the presence of a common component, representing the business cycle. The difference between the trend component of a multivariate approach and the trend of the univariate approach can be interpreted as an autonomous transitory part, peculiar to the dynamics of the series analyzed.

7 To save space we do not show the estimates of the SWQM model because they are not directly comparable with the trend plus cycle models. They are available on request.

Table 2 **Estimated parameters of the trend plus cycle models**

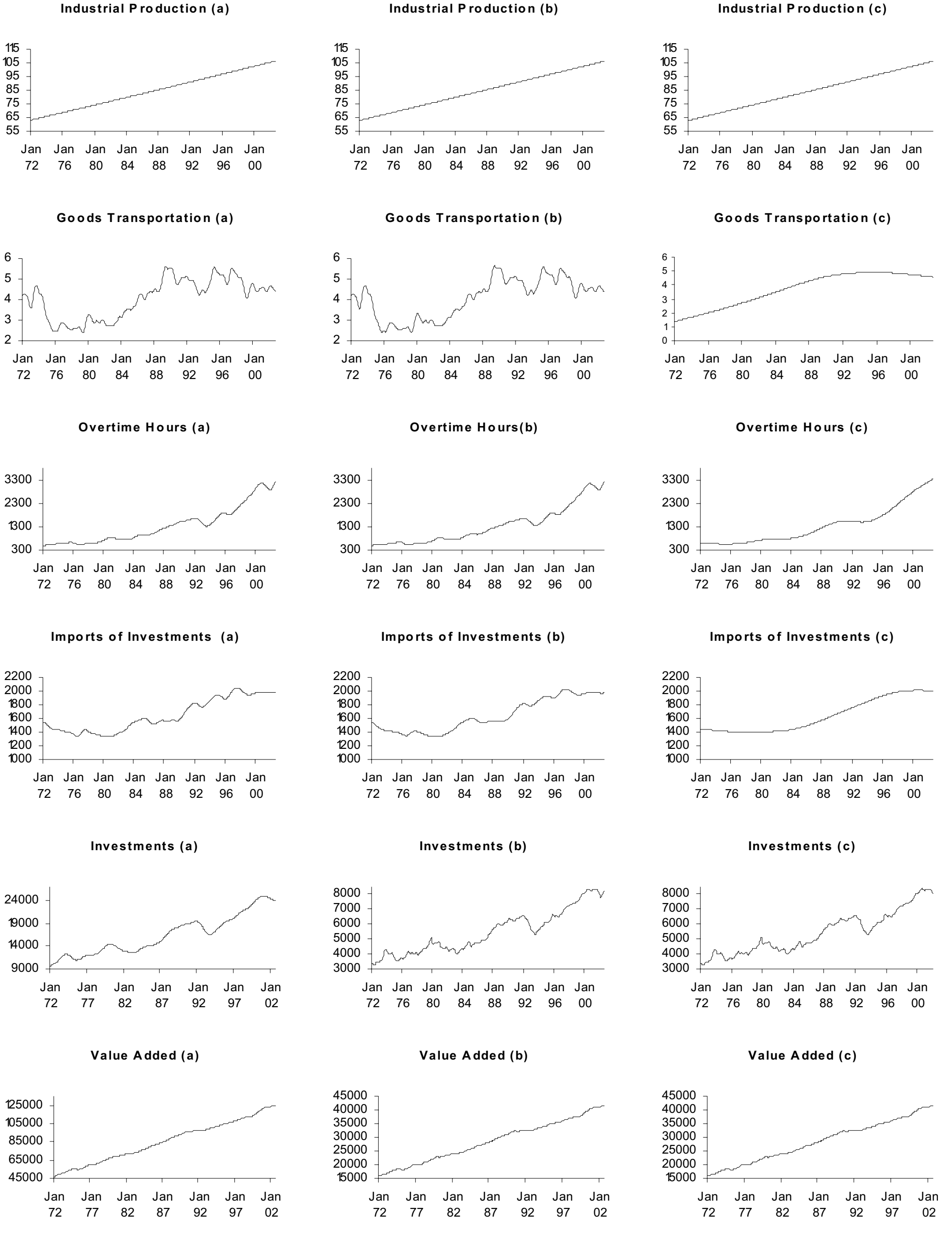
	QMMM	MMM	UIA					
			y_1	y_2	y_3	y_4	y_5	y_6
σ_1	0.432	0.336	0.422					
σ_2	0.134	0.133		0.055				
σ_3	74.361	74.368			70.751			
σ_4	59.636	58.682				54.099		
σ_5	271.87	0.301					0.082	
σ_6	292.57	0.142						0.068
c_1	0.149	0.799						
c_2	0.001	0.003						
c_3	1.565	11.054						
c_4	4.237	23.349						
c_5	1.226	1.220						
υ_1	0.001	0.000	0.000					
υ_2	0.063	0.064		0.001				
υ_3	8.153	8.872			1.734			
υ_4	4.632	2.538				0.341		
υ_5	25.742	36.019					34.715	
υ_6	51.095	49.539						47.992
ρ	0.953	0.963	0.970	0.952	0.971	0.944	0.963	0.963
λ	0.087	0.080	0.082	0.000	0.130	0.117	0.082	0.082
κ	5.579	1.128	0.830	0.178	29.576	30.28	0.330	0.334

QMMM = quarterly-monthly multivariate model

MMM = monthly multivariate model

UIA = univariate indirect approach

Figure 2 **Trends extracted with the QMMM (a), MMM (b), UIA (c) models**



It is interesting to observe that the trend of industrial production is a straight line, which implies that its fluctuations are totally due to the cyclical component; in addition it is a deterministic trend, as shown by the estimates of the standard deviation ν_1.

Compared to the cycle, the frequency λ is 0.087 in the QMMM and 0.080 in the MMM approach, which implies a period $2\pi/\lambda$ for the common cycle corresponding to 6 years for the former and 6.5 years for the latter. Note that in the univariate approach the only variables with a similar period of the cycle are the industrial production, the investment and the value added of the service sector; the others have different behavior, with the extreme case of goods transportation, with $\lambda = 0$. Furthermore, the anomalous behavior of this series, in cyclical terms, is confirmed by the null coefficient c_2 in the multivariate models, which in practice eliminates the $\psi(t)$ component. This is confirmed when we exclude this variable in the multivariate models; the results are the same as presented in Table 2 with the same inference on trends and cycles. In other words, using trend plus cycle models, the goods transported do not provide any relevant information about the common cycle and use of this variable seems inappropriate. However the following comments are valid for both the 6-variable and 5-variable models.

The variances of the cyclical component of the two multivariate approaches are quite different, but this does not imply different dynamics; in Figure 3 we can see that the cyclical components obtained with the two multivariate models[8] have similar dynamics with two evident differences; first, at the beginning of the series the cycle obtained by MMM shows the end of a decrease, whereas that derived by QMMM is increasing. Second, in the period 1994-1995 the MMM indicator shows a short cycle, that is not evident in the QMMM indicator. On the other hand there is a certain degree of similarity in the phases of growth and recession compared to those derived from the ISAE composite indicator.

The graph of the coincident indicator obtained by the SWQM model is also shown in Figure 3. We can observe a similar behavior as the ISAE composite indicator (except at the beginning and the end of the series), but the cyclical movements are less marked.

A more clear comparison can be made by using the turning points derived from each approach, obtained by the Bry and Boschan routine. Results are shown in Table 3.

First of all we comment the trend plus cycle models (QMMM, MMM and UIA). All the proposed procedures capture the two recessions in 1973-74 (first oil shock) and 1977; but in the observed period the MMM procedure identifies the trough in June 1972, whereas the other procedures do not identify this trough. This is one difference between the two multivariate methods, probably due to the use of quarterly data. In fact, in the univariate

8 The graphs plotted are obtained applying a band-pass filter to the component $\psi(t)$ obtained by the QMMM and MMM. This transformation is made to smooth the series and make the turning points visible, because in the original series they were obscured. The type of filter used is a Baxter and King (1999) type, modified for the end-of-sample values as contained in the routine Busy (Fiorentini and Planas, 2003). It is a centred and symmetric filter, so that it produces smoothing without moving the turning points. We thank an anonymous referee who made this point.

analysis with monthly variables, the estimated cycles of investments and value added of the service sector show a deep trough with a successive peak (top of Figure 4). We have also extracted the cycle directly from the original quarterly series and this behavior is not present (bottom of Figure 4). We expect that the disaggregated series have produced the trough in the multivariate analysis with monthly variables. The synchronization of turning points derived from the six univariate analyses does not provide this trough, as it is not present in the other series (except in the good transportation on railways). Table 3 does not show all the turning points obtained with the univariate analysis (these results are available on request).

Table 3 **Turning points derived from different approaches**

Turning Points	ISAE	QMMM	MMM	UIA	SWQM
Trough			jun-72		mar-72
Peak	mar-74	jan-74	jan-74	jan-74	dec-74
Trough	may-75	aug-75	aug-75	jun-75	jun-75
Peak	feb-77	dec-76	dec-76	nov-76	mar-77
Trough	dec-77	dec-77	dec-77	dec-77	dec-77
Peak	mar-80	mar-80	mar-80	jan-80	
Trough	mar-83	may-83	may-83	mar-83	
Peak			aug-84	nov-84	
Trough			oct-85	nov-86	
Peak		aug-89	aug-89	nov-88	
Trough				jul-90	
Peak	mar-92			jan-92	jul-92
Trough	jul-93	aug-93	aug-93	aug-93	jun-93
Peak	nov-95	aug-95	aug-95	sept-95	
Trough	nov-96	dec-96	dec-96	nov-96	
Peak		dec-97	dec-97	nov-97	
Trough		may-99	may-99	may-99	
Peak	dec-00	dec-00	dec-00	sep-00	sep-01
		Dissimilarity with respect to the ISAE dating			
		0.168	0.222	0.211	0.203

ISAE = procedure by the Institute for Studies and Economic Analyses
QMMM = quarterly-monthly multivariate model
MMM = monthly multivariate model
UIA = univariate indirect approach
SWQM = Stock and Watson model with quarterly and monthly series

Figure 3 **ISAE Composite Indicator and cycles extracted with QMMM and MMM**

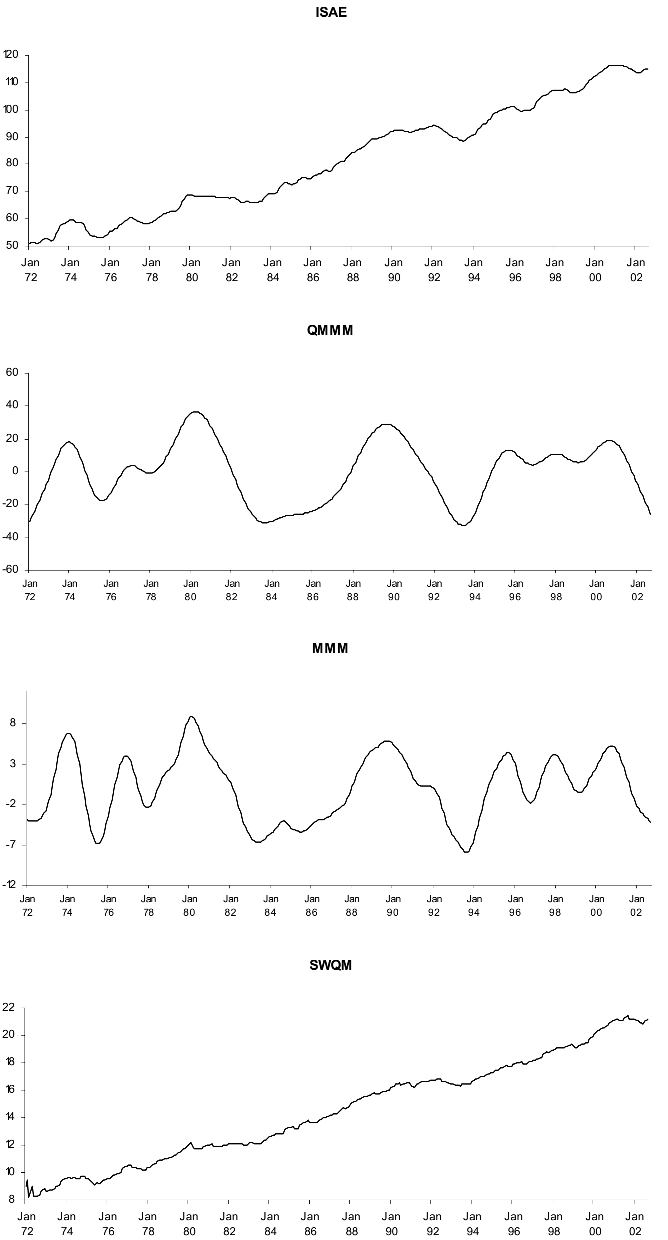
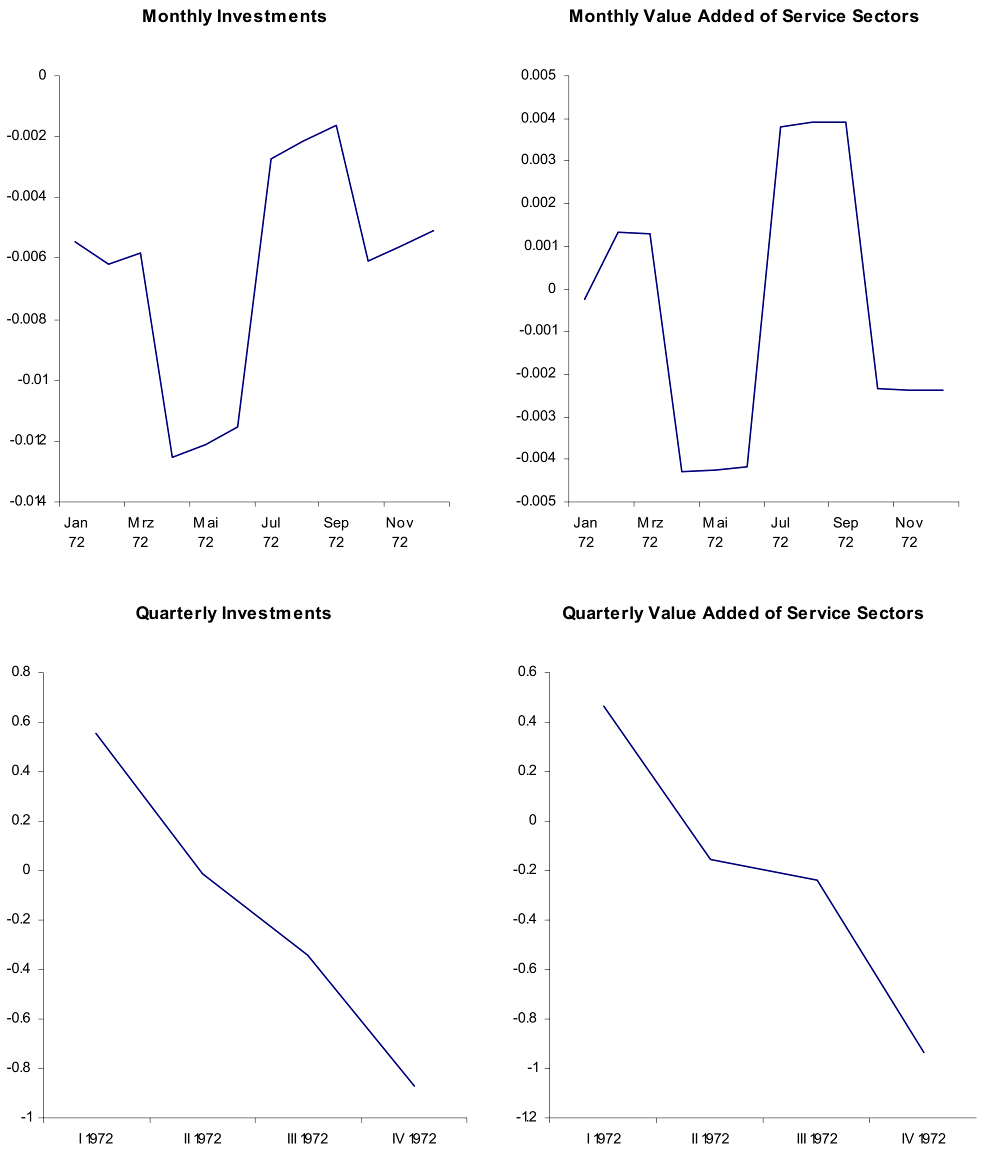


Figure 4 **Details of the cyclical components in 1972**



All methods agree in detecting a peak in the first half of 1980, starting a 3-years long recession. Another difference between QMMM and MMM arises in this period; in fact MMM shows an extra-cycle in the period 1983-1985, whereas QMMM establishes a long growth period without interruptions starting from 1983 until August 1989 (whereas ISAE until March 1992). In this case the difference is explained by the censoring rules of the Bry-Boschan routine; in fact, after the first screening, it identifies a peak in August 1984 for QMMM too, whereas the trough is placed in January 1985. This last one is dropped to ensure the constraint of the minimum phase duration of six periods and, as a consequence, the peak of August 1984 is deleted to ensure the alternation of turning points.

This period is a puzzling one due to the difficulty to establish an exact dating; in fact UIA follows a proper behavior with more extra-cycles.[9] During the nineties', the turning points derived from the three approaches are consistent with the ISAE dating, establishing a recession in 1995-96, as well as a peak at the end of 2000; but QMMM, MMM and UIA indicate an extra-cycle between the end of 1996 and the middle of 1999.

The SWQM model shows a different behavior compared to the other approaches. It is able to capture the two oil shocks, but it fails to individuate the movements at the beginning of the eighties' and the cycle between 1993 and 1996. As we said, this behavior is due to an excessive smooth indicator (Figure 3). Furthermore this procedure detects "anomalous" turning points at the beginning and the end of the series, putting a trough in March 1972 and a peak in September 2001.

Using the simple list of turning points it is not easy to evaluate the best performance in terms of detection of turning points among the four parametric methods. For this reason, we have calculated a loss function measuring the degree of similarity between the dating of a particular parametric method and the ISAE dating. This loss function is obtained as:

$$\frac{1}{T}\sum_{t=1}^{T}\left|P(t)^{M}-P(t)^{ISAE}\right| \tag{7}$$

where $P(t)^M$ is a dummy variable assuming value zero if at time t the parametric method M has identified a recession period (t is located between a previous peak and a subsequent trough), value one if it has identified a boom period (t is located between a previous trough and a subsequent peak).

The values assumed by (7) are shown in the bottom of Table 3; the two models using quarterly and monthly data have a better performance compared to MMM and UIA; QMMM shows the best index among the four models.

9 Bruno and Otranto (2005) registered the same difficulties using various parametric and non parametric methods.

4 Final Remarks

In this paper we have extended the idea of Mariano and Murasawa (2003) to extract a common cyclical component from a group of series composed by monthly and quarterly data, without transforming them to obtain homogeneous frequencies. Differently from Mariano and Murasawa (2003), who use the Stock and Watson (1991) procedure, we have extended the trend plus cycle model of Harvey (1985) to the multivariate case; this is one of the most used and flexible models created for this kind of analysis and provides directly a common cyclical component.

The model used in our analysis can be considered as a sort of multivariate Hodrick-Prescott filter in state-space form, alternative to the one proposed by Laxton and Tetlow (1992). The last one considers a local common trend model (without the common cyclical component) related to a main variable, whereas the other variables are used as regressors; the cycle is the residual series obtained as difference between the main series and the trend. The multivariate Hodrick-Prescott filter, in the version of Laxton and Tetlow (1992), has a state-space representation (see Boone, 2000), which is comparable to our extension of the model of Harvey (1985). Therefore, we can consider them as models belonging to the same family.

Another purpose was to verify the differences between our approach and the analogous one, obtained using monthly data (with a disaggregation of the quarterly series). This exercise was carried out by analysing cyclical components and by detecting turning points. Apart from the differences in terms of estimation, the cyclical components obtained with the two approaches are very similar and the only difference consists in two extra-cycles, detected by the MMM approach. In this case the QMMM approach is more consistent with the ISAE judgemental evaluation, and this is confirmed by the loss function (7).

The univariate analysis suggests some doubts about the coincident behavior of the six variables selected by Altissimo *et al.* (2000). In this case, only industrial production, investment and value added of the service sector have a similar cyclical frequency, which is consistent with the dynamics deduced by the multivariate models, whereas the railway transportation of goods does not seem to be useful to determine the common cycle in both of the multivariate approaches proposed here.

Finally, we want to emphasize the utility of the exercise developed in this work: the contemporaneous use of data with different frequency in multivariate models can be easily implemented and it can provide good results, without creating artificial data. Our model seems to be especially successful in terms of detection of turning points. It could be extended to all the multivariate models which can be represented in a state-space form.

References

Altissimo, F.; D. J. Marchetti and G. P. Oneto. 2000. "The Italian Business Cycle: Coincident and Leading Indicators and Some Stylized Facts," Temi di Discussione del Servizio Studi Banca d'Italia, p. 377.

Azevedo, J. V.; S. J. Koopman and A. Rua. 2003. "Tracking growth and the business cycle: a stochastic common cycle model for the Euro Area," Tinbergen Institute, Discussion Paper no. 03-069/4.

Baxter, M. and R. G. King. 1999. "Measuring Business Cycles: Approximate Bandpass Filters for Economic Time Series," *Review of Economics and Statistics* 81, pp. 575-93.

Boone, L. 2000. "Comparing semi-structural methods to estimate unobserved variables: the HPMV and Kalman filters approaches," OECD Economic Department, Working Paper no. 240.

Box, G. E. P. and G. M. Jenkins, G. M. 1976. *Time Series Analysis: Forecasting and Control.* San Francisco: Holden-Day.

Bry, G. and C. Boschan. 1971. "Cyclical Analysis of Time Series: Selected Procedures and Computer Programs," NBER Technical Papers no. 20.

Bruno, G. and E. Otranto, E. 2005. "Dating the Italian business cycle: a comparison of procedures," in *Business Cycles: Country Experiences,* V. V. Ramani, ed., (forthcoming) pp. 30-49. Hyderabad: The ICFAI University Press.

Dickey, D. A. and W. A. Fuller. 1979. "Distribution of the Estimators for Autoregressive Time Series with a Unit Root," *Journal of the American Statistical Association* 74, pp. 427-31.

Doan, T. A. 2000. *RATS for Windows Version 5.* Evanston: Estima.

Fernandez, R. B. 1981. "A Methodological Note on the Estimation of Time Series," *Review of Economics and Statistics* 63, pp. 471-78.

Fiorentini, G. and C. Planas. 2003. "Busy program User Manual," Joint Research Centre of European Commission Document.

Gómez, V. and A. Maravall, A. 1997. "Programs TRAMO and SEATS: Instructions for the User," Dirección General de Análisis y P. P., Ministry of Economy and Finance, Working Paper no. 97001.

Harding, D. and A. Pagan. 2006. "Synchronization of Cycles," *Journal of Econometrics,* forthcoming .

Harvey, A. C. 1985. "Trends and Cycles in Macroeconomic Time Series," *Journal of Business and Economic Statistics* 3, pp. 216-27.

Harvey, A. C. 1990. *Forecasting, structural time series models and the Kalman filter.* Cambridge: Cambridge University Press.

Harvey, A. C. and A. Jaeger. 1993. "Detrending, stylized facts and the business cycle," *Journal of applied econometrics* 8, pp. 231-47.

Harvey, A. C. and S. J. Koopman. 1997. "Multivariate Structural Time Series Models," in: *System Dynamics in Economic and Financial Models,* C. Heij et al., eds. Chichester: Wiley and Sons, pp. 269-98.

Hodrick, R. J. and E. C. Prescott. 1997. "Postwar U.S. business cycles: an empirical investigation," *Journal of money, credit, and banking* 29, pp. 1-16.

Johansen, S. 1991. "Estimation and Hypothesis Testing of Cointegration Vectors in Gaussian Vector Autoregressive Models," *Econometrica* 59, pp. 1551-80.

Johansen, S. 1995. *Likelihood-based inference in Cointegrated Vector Autoregressive Models.* Oxford: Oxford University Press.

Kim, C. J. and J. M. Piger, J. M. 2002. "Common Stochastic Trends, Common Cycles, and Asymmetry in Economic Fluctuations," *Journal of Monetary Economics* 49, pp. 1189-211.

King, R. G.; C. I. Plosser, J. H. Stock and M. W. Watson. 1991. "Stochastic Trends and Economic Fluctuations," *The American Economic Review* 81, pp. 819-40.

Koopman, S. J.; A. C. Harvey, J. A. Doornik and N. Shephard. 2000 *Stamp: Structural Time Series Analyser, Modeller and Predictor.* London: Timberlake Consultants Press.

Laxton, D. and R. Tetlow. 1992. "A simple multivariate filter for the measurement of potential output," Bank of Canada, Technical Report no. 59.

Mariano, R. S. and Y. Murasawa. 2003. "A new coincident index of business cycles based on monthly and quarterly series," *Journal of Applied Econometrics* 18, pp. 427-43.

Moauro, F. and T. Proietti. 2004. "An Index of Coincident Indicators for the Euro Area Based on Monthly and Quarterly Series," Eurostat, MOS, Modern Tools for Business Cycle Analysis.

Otranto, E. and R. Iannaccone. 2005. "Continuous Time Models to Extract a Signal in Presence of Irregular Surveys, *Statistica & Applicazioni,* forthcoming.

Pedersen, T. M. 2001. "The Hodrick-Prescott Filter, the Slutzky Effect, and the Distortionary Effect of Filters," *Journal of Economic Dynamics and Control* 25, pp. 1081-101.

Proietti, T. and F. Moauro. 2006. "Dynamic Factor Analysis with Non-linear Temporal Aggregation Constraints," *Journal of the Royal Statistical Society,* series C, forthcoming.

Schwarz, G. 1978. "Estimating the Dimension of a model," *The Annals of statistics* 6, pp. 461-64.

Stock, J. H. and M. W. Watson. 1991. "A probability model of the coincident economic indicators," in *Leading economic indicators,* K. Lahiri and G. H. Moore, eds. Cambridge: Cambridge University Press, pp. 63-89.

Contents

From:
Journal of Business Cycle Measurement and Analysis

Access the journal at:
https://doi.org/10.1787/17293626

Please cite this article as:

Otranto, Edoardo (2006), "Extracting a Common Cycle from Series with Different Frequency: An Application to the Italian Economy", *Journal of Business Cycle Measurement and Analysis*, Vol. 2005/3.

DOI: https://doi.org/10.1787/jbcma-v2005-art11-en

OECD