An index of coincident indicators for the Euro Area based on monthly and quarterly time series
Europe Direct is a service to help you find answers to your questions about the European Union.

New freephone number:
00 800 6 7 8 9 10 11

A great deal of additional information on the European Union is available on the Internet. It can be accessed through the Europa server (http://europa.eu.int).

Luxembourg: Office for Official Publications of the European Communities, 2003

ISBN 92-894-6856-4
ISSN 1725-4825

© European Communities, 2003
An index of coincident indicators for the Euro Area based on monthly and quarterly time series

Filippo Moauro, ISTAT, Direzione Centrale Contabilità Nazionale, Rome, Italy
Tommaso Proietti, Università di Udine, Facoltà di Economia, Italy
An Index of Coincident Indicators for the Euro Area based on Monthly and Quarterly Series

Filippo Moauro       Tommaso Proietti
ISTAT             Università di Udine

This version: 13th October 2003

Abstract

The paper proposes a new coincident index of business cycle for the Euro Area. The methodology we follow extends the proposal by Mariano and Murasawa (2002), which moves from the Stock and Watson probability approach and combines a set of monthly indicators with the quarterly real GDP. Nevertheless, our set-up is formulated in the levels instead of first differences of the representative series, modelling the logarithms like the original proposals. In this respect, the non-linear state space form for the treatment of temporal aggregation is developed. The application is carried out first to the well known US dataset, comparing our results with the estimates coming from the original proposals and then to the Euro Area. For the latter we use the OECD database for the quarterly series of GDP and total employment, along with the monthly industrial production and retail sales indexes. Finally, for both the applications we derive a monthly estimate of GDP.

Keywords: Temporal disaggregation; Kalman filter and smoother; Business Cycle; Nonlinear state space form.

JEL classification: C5, C22, E32.
1 Introduction

A prominent feature of the business cycle is the presence of similarities in the dynamics of several representative series or, following Lucas (1977), co-movements. According to the classical business cycle definition proposed by Burns and Mitchell (1946), the business cycle “is composed of expansions which take place almost at the same time in many economic activities, followed by equally general recessions, contractions and recoveries which merge with the expansion phase of the following cycle; this sequence of changes is recurrent but not periodical”.

Then the very nature of business cycles fluctuations implies that the “reference cycle” cannot be extracted from a single series, for instance the real measure of Gross Domestic Product (GDP), but an analysis of a range of relevant indicators of economic activity is required.

Stock and Watson (1991, SW henceforth) developed an explicit probability model for the composite index of coincident economic indicators. They proposed a dynamic factor model with a common difference stationary factor that defines the composite index. The reference cycle is assumed to be the value of a single unobservable variable, “the state of the economy”, that by assumption is the only source of the co-movements of four time series: industrial production, sales, employment, and real incomes.

On the other hand, GDP is perhaps the most important coincident indicator, although it is available only quarterly and it is subject to greater revisions than the four coincident series in the original SW model.

Moving from this approach, Mariano and Murosawa (2003, MM hereafter), propose to extend the SW model with the inclusion of quarterly real GDP, developing a method for the treatment of mixed-frequency series.

In this paper we improve upon MM in the following respects:

1. the model is formulated in terms of the levels instead of first differences of the series. It allows to derive the uncertainty of the real time and the final estimates of
the coincident indicator;

2. the nonlinear temporal aggregation constraint is approached consistently to the theory developed by Durbin and Koopman (2001), modelling the logarithms of selected indicators like the proposals of SW and MM. We develop a state space form of the linear Gaussian approximating model, which speeds up significantly computation;

3. the initialisation issue which arises for the level formulation of the coincident index model is solved adopting the exact Kalman filter developed by Koopman (1997).

The paper is organized as follows: in Section 2 the coincident index model is introduced. Section 3 presents the state space representation and Section 4 the temporal disaggregation issue. The solution to the nonlinear temporal aggregation constraint is proposed in Section 5, along with the formulation of the Maximum likelihood estimation. Estimation results for the US and the Euro Area exercises are considered in Section 6, with a comparison with the business cycle chronologies currently released. Section 7 concludes.

2 The coincident index model

The coincident index model proposed by SW has been developed with an approach which is a natural extension of the dynamic index model proposed by Geweke (1977) and Sargent and Sims (1977). The authors aim at rationalizing by a probabilistic model the judgmental procedure used by the Department of Commerce to build up a coincident indicator for the US economy. Their basic idea is to separate the dynamics which are common to a set of \( N \) coincident series \( y_t \) from the idiosyncratic component, which is specific to each series; this is formalized in the single index model, where \( y_t \) is assumed to be \( I(1) \) but not cointegrated.

The SW model features a common cyclical trend, that we shall denote by \( \mu_t \), which adds up to the vector of idiosyncratic component \( \mu_t^* \). The model can be represented as
follows:

\[
\begin{align*}
Y_t &= \theta \mu_t + \mu_t^*, \quad t = 1, \ldots, n \\
\phi(L) \Delta \mu_t &= \eta_t, \quad \eta_t \sim \text{NID}(0, \sigma^2_\eta) \\
D(L) \Delta \mu_t^* &= \beta + \eta_t^*, \quad \eta_t^* \sim \text{NID}(0, \Sigma_{\eta^*})
\end{align*}
\]

where $\phi(L)$ is an autoregressive polynomial of order $p$ with stationary roots:

\[
\phi(L) = 1 - \phi_1 L - \cdots - \phi_p L^p;
\]

and the matrix polynomial $D(L)$ is diagonal:

\[
D(L) = \text{diag} \left[ d_1(L), d_2(L), \ldots, d_N(L) \right],
\]

with $d_i(L) = 1 - d_{i1} L - \cdots - d_{ip} L^p$ and $\Sigma_{\eta^*} = \text{diag}(\sigma^2_1, \ldots, \sigma^2_N)$. The disturbances $\eta_t$ and $\eta_t^*$ are mutually uncorrelated at all leads and lags.

This slightly different representation is more useful since it is based on the index itself and eliminates the ambiguities in the interpretation of the real time (filtered) and smoothed estimates that arise when the model is formulated in terms of differences; for an account see also Mariano and Murasawa (2003, MM hereafter). Moreover, the representation (1) assumes a zero drift for the single index.

The complication is that the state vector has $N$ elements more with respect to the model on the differences, but the latter has a price to pay when temporal aggregation is present.

Note that both components are difference stationary processes and the common dynamics are the results of the accumulation of the same underlying shock $\eta_t$; moreover, the process generating the index of coincident indicators is usually more persistent than a random walk and in the accumulation of the shocks produces cyclical swings.

Alternative specifications, which separate the common trends from the (possibly common) cycles have recently been proposed in the literature. Further proposals discuss the possibility to consider similar cycle restrictions into the model. In this respect their inclusion in the representation (1) simply implies $d_i(L) = \phi(L)$.
3 State space representation

The treatment of model (1) through the Kalman filter (KF) requires the development of an appropriate state space form (SSF). We start from the model for the single index:

\[ \phi(L) \Delta \mu_t = \eta_t, \]

considering first the SSF of the stationary AR(\(p\)) model for the differences of \(\mu_t\), which is:

\[
\Delta \mu_t = e_{1p}^t a_t,
\]

\[
a_t = T_{\Delta \mu} a_{t-1} + e_{1p} \eta_t,
\]

where \(e_{1p} = [1, 0, \ldots, 0]^t\) and

\[
T_{\Delta \mu} = \begin{bmatrix}
\phi_1 & & \\
& \ddots & \\
& & \mathbf{I}_{p-1} \\
\phi_{p-1} & & \\
\phi_p & & 0^t
\end{bmatrix}.
\]

Hence, \(\mu_t = \mu_{t-1} + e_{1p}^t a_t = \mu_{t-1} + e_{1p}^t T_{\Delta \mu} a_{t-1} + \eta_t\), and defining

\[
\alpha_{\mu,t} = [\mu_t, a_t]^t,
\]

the Markovian representation of the model for \(\mu_t\) becomes

\[
\mu_t = e_{1p}^{t+1} \alpha_{\mu,t}, \quad \alpha_{\mu,t} = T_{\mu} \alpha_{\mu,t-1} + R_{\mu} \eta_t,
\]

where \(R_{\mu} = [1, e_{1p}^t]^t\).

Let \(p_i\) denote the order of the \(i\)-th lag polynomial \(d_i(L)\); a similar representation holds for each individual \(\mu_{it}^*\), with \(\phi_j\) replaced by \(d_{ij}\), then:

\[
\mu_{it}^* = e_{1p_i}^{t+1} \alpha_{\mu_{it}}, \quad \alpha_{\mu_{it}} = T_{i} \alpha_{\mu_{it-1}} + c_i + R_i \eta_{it}^*,
\]

where \(R_i = [1, e_{1p_i}^t]^t\), \(c_i = \beta_i R_i\) and \(\beta_i\) is the drift of the \(i\)-th series.
Combining all the blocks, the state vector $\alpha_t$ has dimension $\sum_i (p_i + 1) + p + 1$ becoming

$$\alpha_t = [\alpha_{\mu,t}, \alpha_{\mu_1,t}, \ldots, \alpha_{\mu_N,t}]$$

then, the SSF is as follows:

$$y_t = Z\alpha_t, \quad \alpha_t = T\alpha_{t-1} + c + Re_t$$

$$Z = \begin{bmatrix} \theta & \text{diag}(e_{p_1}', \ldots, e_{p_N}') \end{bmatrix}; \quad T = \text{diag}(T_{\mu}, T_1, \ldots, T_N).$$

### 4 Temporal disaggregation

In practical applications of the SW approach, it can happen that the coincident indicators are observed at different frequencies. For example, in the MM exercise to the US, quarterly GDP is considered along with 4 monthly indicators, the same originally selected by SW. Further, if one extends the same approach to other Countries the same circumstance might easily occur for important indicators like employment or income.

From a theoretical point of view, modelling mixed frequency time series requires a distinction between flows and stocks variables. When a time series refers to measures of flows, less frequent observations imply that the original data are temporal aggregated. Then, these data need to be distributed over time, assuming the frequency of the model as reference. Several solutions to the treatment of mixed frequency series have been proposed by the literature. In the applications related to the KF, it is considered as a missing observation problem. It requires the SSF to be appropriately adjusted with respect to the case in which temporal aggregation do not occur.

Conversely, when the variables under interest are stocks (e.g. money, debt or population), missing observations are treated through interpolation; in this case it can be...
sufficient the application of the KF to the unchanged SSF simply skipping the missing observations.

Then, let’s consider the case the set of coincident indicators $y_t$ includes mixed frequency series measuring flows. Then, $y_t$ can be partitioned into $[y_{1t}'; y_{2t}']'$, where the second block is the one subject to temporal aggregation. The approach we follow here consists in operating a suitable augmentation of the state vector (2) using an appropriately defined cumulator variable. In particular, the SSF (3)-(6) need to be augmented by the $N_2 \times 1$ vector $y_{2t}^\dagger$, generated as follows

$$y_{2t}^\dagger = \psi_t y_{2,t-1}^\dagger + y_{2t}$$

$$= \psi_t y_{2,t-1}^\dagger + Z_2 T \alpha_{t-1} + Z_2 c + Z_2 R e_t$$

where

$$\psi_t = \begin{cases} 0 & t = \delta(\tau - 1) + 1, \tau = 1, \ldots, T \\ 1 & \text{otherwise} \end{cases}$$

Further details on the treatment of the cumulator can be found in Harvey (1989, p. 314). A Markovian representation including the cumulator $y_{2t}^\dagger$ into the state vector can be given to the model; the part of the measurement equation referring to the second block is $y_{2t} = y_{2t}^\dagger$, $t = \delta \tau$, $\tau = 1, 2, \ldots$, else it is missing.

Define the augmented state vector $\alpha_t^* = \begin{bmatrix} \alpha_t^T \\ y_{2t}^\dagger \end{bmatrix}$; its dimension is $m^* = m + N_2$

Then the augmented SSF is:

$$y_t = Z^\ast \alpha_t^*, \quad \alpha_t^* = T^* \alpha_{T-1} + c^* + R^* e_t$$  \hspace{1cm} (5)$$

with

$$Z^\ast = \begin{bmatrix} Z_1 & 0 \\ 0 & I_{N_2} \end{bmatrix}, \quad T^* = \begin{bmatrix} T & 0 \\ Z_2 T & \psi I \end{bmatrix}, \quad c^* = \begin{bmatrix} I \\ Z_2 \end{bmatrix} c, \quad R^* = \begin{bmatrix} I \\ Z_2 \end{bmatrix} R.$$  \hspace{1cm} (6)
5 Nonlinear temporal aggregation constraint

Consider now the case when $y_t$ represents logarithms and that the second block of series is temporally aggregated, e.g., it is quarterly when the model is formulated at the monthly frequency ($\delta = 3$). The aggregation constraints is linear in $Y_{2t} = \exp(y_{2t})$, that is the aggregated series results from

$$Y_{2t} = \sum_{i=0}^{\delta-1} Y_{2,\tau\delta-i}$$

Then, the cumulator variable is defined recursively as follows:

$$Y_{2t} = \psi_t Y_{2,t-1} + \exp(y_{2t}) = \psi_t Y_{2,t-1} + \exp(Z_2\alpha_t)$$

where the last expression follows from the fact that the SSF does not feature a measurement error.

Let us consider now the linearization of the cumulator around the trial value $\tilde{\alpha}_t$, based on the first order Taylor series expansion:

$$Y_{2t} = \psi_t Y_{2,t-1} + \exp(Z_2\tilde{\alpha}_t) + \tilde{D}Z_2(\alpha_t - \tilde{\alpha}_t)$$

$$= \psi_t Y_{2,t-1} + \exp(Z_2\tilde{\alpha}_t) - \tilde{D}_t Z_2 \tilde{\alpha}_t + \tilde{D}_t Z_2 \alpha_t - \tilde{D}_t Z_2 c + \tilde{D}_t Z_2 R^2_t$$

where $\tilde{D}_t = \text{diag}(z_{2i}^\dagger \tilde{\alpha}_t)$ and $z_{2i}$ denotes the $i$-th row of $Z_2$.

The SSF is based upon the augmented vector $\alpha_t^\dagger = [\alpha_t^\dagger, Y_{2t}^\dagger]'$, with measurement equation

$$\begin{bmatrix} y_{1t}^\dagger \\ Y_{2t}^\dagger \end{bmatrix} = \begin{bmatrix} Z_1 & 0 \\ 0 & I \end{bmatrix} \alpha_t^\dagger$$

The left hand side lower block is observed only at $t = \tau\delta$. So the SSF is subject to missing values.

The transition equation is:

$$\alpha_t^\dagger = \begin{bmatrix} T & 0 \\ \tilde{D}_t Z_2 T & \psi_t I \end{bmatrix} \alpha_{t-1}^\dagger + \begin{bmatrix} c \\ \exp(Z_2\tilde{\alpha}_t) - \tilde{D}_t Z_2 \tilde{\alpha}_t + \tilde{D}_t Z_2 c \end{bmatrix} + \begin{bmatrix} I \\ \tilde{D}_t Z_2 \end{bmatrix} R\epsilon_t$$

(8)
Given $\tilde{\alpha}_t$, the linear Gaussian (LG) approximating model is given by (7)-(8). The output of the Kalman filter smoothing (KFS) is used to define a new $\tilde{\alpha}_t$ form $\alpha_t^\dagger$ and a new approximating model. Iterating ensures that the LG approximating model has the same conditional mode as the original nonlinear one.

Maximum likelihood estimation (MLE) and signal extraction are performed via linearizing the model, solving the model equation and evaluating the likelihood of the optimized linear Gaussian model.

An alternative representation that also uses the Gaussian likelihood of the linear approximating model considers the nonlinearity in the measurement equation, whereas the transition retains its linearity. Define the state vector $\alpha_t^* = [\alpha_t', c_t', c_{t-1}', \ldots, c_{t-\delta+1}']'$ where $c_t = Z_2 \alpha_t$. The measurement equation for the aggregated time series is:

$$Y_{2t} = \sum_{i=0}^{\delta-1} \rho_t \exp(c_{t-j})$$

for a suitable set of time-varying coefficients $\rho_t$.

This simplifies the inferences, at the expenses of a larger state vector, that features $N_2 \cdot (\delta - 1)$ elements in excess of the previous representation. Not very much if $N_2$ and $\delta$ are small.

5.1 Maximum likelihood estimation

As shown in Durbin and Koopman (2001),

$$\mathcal{L}(\psi) = \int p(y, \alpha^\dagger) d\alpha^\dagger$$

$$= \mathcal{L}_g(\psi) \int \frac{p(y, \alpha^\dagger)}{g(y, \alpha^\dagger)} g(\alpha^\dagger | y) d\alpha^\dagger$$

$$= \mathcal{L}_g(\psi) \int \frac{p(\alpha^\dagger)}{g(\alpha^\dagger)} g(\alpha^\dagger | y) d\alpha^\dagger$$

(9)

This is because $p(y | \alpha^\dagger) = g(y | \alpha^\dagger) = 1$ due to the absence of measurement noise, and $p(\alpha) = g(\alpha) = g(\alpha_0) \prod_{t=1}^{T} g(\alpha_t | \alpha_{t-1})$, which in turn implies

$$\frac{p(\alpha^\dagger)}{g(\alpha^\dagger)} = \frac{p(\alpha, Y_2)}{g(\alpha, Y_2)} = \prod_{t=1}^{T} \frac{p(Y_{2t} | \alpha_t, Y_{2t-1})}{g(Y_{2t} | \alpha_t, Y_{2t-1})}$$
Now, if $\alpha_t^\dagger$ denotes the solution of the mode equation, the ratio on the right hand side is equal to 1 and the likelihood of the nonlinear model is identical to that of the LG approximating model, that is $L(\psi) = L_g(\psi)$. The latter is evaluated by the KF.

The reason why it occurs is that when the iteration converges, $\alpha_t^\dagger - \tilde{\alpha}_t^\dagger$ implies

$$\tilde{Y}_{2t} - \psi_t \tilde{Y}_{2,t-1} - \exp(Z_2 \alpha_t) - \tilde{D}Z_2(\alpha_t - \tilde{\alpha}_t) = \tilde{Y}_{2t} - \psi_t \tilde{Y}_{2,t-1} - \exp(Z_2 \tilde{\alpha}_t)$$

6 Estimation results

The method discussed above has been applied first to the set of coincident indicators selected by SW for the US, plus the quarterly GDP (Figure 1) and then to the Euro area, considering the corresponding dataset (Figure 2). For the latter a complication consists in the fact that the information set is dramatically reduced, because the sample period is shorter, the series for Income is not significant and the indicator for employment is dispensed at quarterly frequency only. Our goal is twofold: first, we want to verify how consistent is the gain of the inclusion of GDP in the SW procedure and second we are interested in the monthly estimate of GDP.

All the results have been obtained using Ox 3.20 by Doornik (2001) and the package SsfPack 3.0 beta (see Koopman, Shephard and Doornik 2002).

6.1 Results obtained for the US

In this application we add the quarterly GDP to the four classical monthly indicators following, then, the strategy of MM. In the notation of previous Sections $y_t$ has therefore dimension $N = 5$. All the series are adjusted for seasonality and we denote them as follows:

1. IIP, the industrial production index 1997 = 100 from 1946.1 to 2003.2, source Board of Governors of the Federal Reserve System;
2. EMP, the employees on non-agricultural payrolls in thousands from 1946.1 to 2003.3, source Department of Labour, Bureau of Labor Statistics;

3. SLS, the manufactured and trade sales in millions of chained 1996 dollars from 1946.1 to 2003.1, source Department of Commerce, Bureau of Census;

4. INC, annual rate of the personal income less transfer payments in billions of chained 1996 dollars (?) from 1946.1 to 2003.2, source Department of Commerce, Bureau of Economic Analysis;

5. GDP, annual rate of the real Gross Domestic Product in billions of chained 1996 dollars from 1947 1st quarter to 2003 1st quarter, source Department of Commerce, Bureau of Economic Analysis.

In model parameter estimation we consider the logarithms of the original data for monthly indicators and the index 1947 1st quarter equal to 1 for GDP. We follow SW in the identification assumption that the variance of the common cyclical trend $\sigma_\eta^2$ is equal to 1 in Equation (1) and we adopt the $AR(2)$ specification for $\mu_t$ and all the elements of $\mu_t^t$. Therefore, the resulting SSF has dimension $(5 + 1) \times (2 + 1) + 1 = 19$. The sample period is taken from 1946.1 to 2003.1 ($n = 687$), starting 13 years before both the SW and MM applications.

The results of the estimates of model parameters are summarized in Table 1 where in parentheses are the asymptotic standard errors. The log likelihood for the approximating model is $\mathcal{L} = 10015.53$. The estimates of $d_1$ and $d_2$ for monthly indicators exhibits signs in line with the estimates of both SW and MM and a general coherence in terms of value. The differences are explained through the revisions occurred in the indicators over time and the longer sample period considered in our experiment. The autoregressive terms of $\mu_t$ show an higher value at lag two (0.1856) than both the SW and MM estimates (0.032...
Table 1: Estimated coincident index model for the US

<table>
<thead>
<tr>
<th>Parameters</th>
<th>IIP</th>
<th>EMP</th>
<th>SLS</th>
<th>INC</th>
<th>GDP</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\theta \times 100$</td>
<td>0.853</td>
<td>0.245</td>
<td>0.677</td>
<td>0.320</td>
<td>0.372</td>
</tr>
<tr>
<td></td>
<td>(0.035)</td>
<td>(0.011)</td>
<td>(0.041)</td>
<td>(0.021)</td>
<td>(0.021)</td>
</tr>
<tr>
<td>$\beta \times 100$</td>
<td>0.425</td>
<td>0.093</td>
<td>0.489</td>
<td>0.267</td>
<td>0.424</td>
</tr>
<tr>
<td></td>
<td>(0.106)</td>
<td>(0.017)</td>
<td>(0.100)</td>
<td>(0.035)</td>
<td>(0.154)</td>
</tr>
<tr>
<td>$d_1$</td>
<td>-0.174</td>
<td>0.166</td>
<td>-0.446</td>
<td>-0.014</td>
<td>0.175</td>
</tr>
<tr>
<td></td>
<td>(0.062)</td>
<td>(0.047)</td>
<td>(0.049)</td>
<td>(0.051)</td>
<td>(0.366)</td>
</tr>
<tr>
<td>$d_2$</td>
<td>-0.283</td>
<td>0.295</td>
<td>-0.221</td>
<td>0.056</td>
<td>-0.660</td>
</tr>
<tr>
<td></td>
<td>(0.061)</td>
<td>(0.059)</td>
<td>(0.047)</td>
<td>(0.052)</td>
<td>(0.172)</td>
</tr>
<tr>
<td>$\sigma_{\eta^*} \times 100$</td>
<td>0.570</td>
<td>0.181</td>
<td>0.776</td>
<td>0.296</td>
<td>0.396</td>
</tr>
<tr>
<td></td>
<td>(0.029)</td>
<td>(0.001)</td>
<td>(0.003)</td>
<td>(0.011)</td>
<td>(0.130)</td>
</tr>
</tbody>
</table>

$(1 - 0.3382L - 0.1856L^2) \Delta \mu_t = \eta_t, \quad \eta_t \sim N(0, 1)$

Note: In parenthesis are the standard errors of the estimates.

and 0.08 respectively). With respect to GDP the first autoregressive term is positive, whereas it is slightly negative for MM ($-0.04$).

Table 2 presents the diagnostics based on the KF innovations. For the 5 series we consider the measure of the Box-Ljung statistics $Q(15)$ and $Q(25)$ based, respectively, on 15 and 25 autocorrelations, the Bowman-Shenton normality test $\text{Norm}$ and the heteroskedasticity statistic $H(h)$, where $h = 229$ for the monthly indicators $\text{IIP, EMP, SLS}$ and $\text{INC}$ and $h = 76$ for $\text{GDP}$. The results suggests a satisfactory specifications for all the equations. The high values for $\text{IIP}$ and $\text{EMP}$ in the Normality test are explained with a limited number of outliers in the first two decades. However, neither we make adjustments for those, nor we consider alternative ARIMA specifications for the common and/or idiosyncratic components. The reason is that the model shows a general good fit and our interest goes much more in the reliability of the method in determining the business cycle and in distributing the aggregated values.
Table 2: Diagnostics on residuals for the US model

<table>
<thead>
<tr>
<th>Tests</th>
<th>IIP</th>
<th>EMP</th>
<th>SLS</th>
<th>INC</th>
<th>GDP</th>
</tr>
</thead>
<tbody>
<tr>
<td>$Q(15)$</td>
<td>22.131</td>
<td>31.953</td>
<td>22.767</td>
<td>27.916</td>
<td>10.520</td>
</tr>
<tr>
<td>$Q(25)$</td>
<td>42.736</td>
<td>60.365</td>
<td>42.010</td>
<td>33.964</td>
<td>29.153</td>
</tr>
<tr>
<td>Norm</td>
<td>163.491</td>
<td>2485.725</td>
<td>15.444</td>
<td>31.956</td>
<td>28.235</td>
</tr>
<tr>
<td>$H(h)$</td>
<td>0.321</td>
<td>0.204</td>
<td>3.794</td>
<td>2.931</td>
<td>0.319</td>
</tr>
</tbody>
</table>

Note: $Q(15)$ and $Q(25)$ are the Box-Ljung statistics based, respectively on 15 and 25 residual autocorrelations, Norm is the Bowman-Shenton Normality test and $H(h)$ is the test for heteroskedasticity ($h=76$ for GDP and $h=229$ for the other series).

In Table 3, we present the comparison of business cycle turning points of our index with the official NBER reference dates. Table 3 also considers the turning points identified by GDP, the experimental index of SW ($XCI$) and the proposal of MM related to the smoothing estimates ($CIMM$). The numbers in the table refers to lags, denoted with $+$, and leads, with $-$, with respect the official NBER business cycle chronology.

Our coincident index, denoted as $CI_t$, has been obtained through the fixed interval smoothing estimates of $\mu_t$, at which we add the monthly growth average rate of GDP computed as follows

$$CI_t = CI_{t-1} + \hat{\theta}_{GDP}\Delta\hat{\mu}_t + \left(1 - \hat{a}_1^{GDP} - \hat{a}_2^{GDP}\right)^{-1}\hat{\beta}_{GDP}, \quad CI_1 = 0, \quad (10)$$

where the values for $\hat{\theta}_{GDP}, \hat{\beta}_{GDP}, \hat{a}_1^{GDP}$ and $\hat{a}_2^{GDP}$ are the estimated model parameters in the last column of Table 1.

In the period from January 1959 to December 2000 the differences between $CI_t$ and $CIMM$ are generally not significant, considering that $CI_t$ gains 2 months in the peak of December 69 whereas the latter 1 month in the peak of July 81 and in the through of March 91. In the same period both the indexes performs better than $XCI$, which signals too early the peaks of April 60, December 69 and July 90, and late the troughs of March 1975 and November 1982. Conversely, GDP is perfectly in line with the official dates for the peaks of the sixties, whereas it leads the troughs of February 61 and November 82.
Table 3: The US turning points for alternative indexes of business cycle and GDP

<table>
<thead>
<tr>
<th></th>
<th>NBER</th>
<th>XCI</th>
<th>GDP</th>
<th>CI&lt;sup&gt;MM&lt;/sup&gt;</th>
<th>CIt</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Peaks</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>November 1948</td>
<td>×</td>
<td>−1</td>
<td>×</td>
<td>−1</td>
<td></td>
</tr>
<tr>
<td>July 1953</td>
<td>×</td>
<td>−2</td>
<td>×</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>August 1957</td>
<td>×</td>
<td>0</td>
<td>×</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>April 1960</td>
<td>−3</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>December 1969</td>
<td>−2</td>
<td>0</td>
<td>−2</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>November 1973</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>January 1980</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>July 1981</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>+1</td>
<td></td>
</tr>
<tr>
<td>July 1990</td>
<td>−1</td>
<td>−1</td>
<td>0</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>March 2001</td>
<td>−6</td>
<td>+2</td>
<td>×</td>
<td>−3</td>
<td></td>
</tr>
<tr>
<td><strong>Troughs</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>October 1949</td>
<td>×</td>
<td>0</td>
<td>×</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>May 1954</td>
<td>×</td>
<td>−1</td>
<td>×</td>
<td>−2</td>
<td></td>
</tr>
<tr>
<td>April 1958</td>
<td>×</td>
<td>−2</td>
<td>×</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>February 1961</td>
<td>0</td>
<td>−2</td>
<td>−2</td>
<td>−2</td>
<td></td>
</tr>
<tr>
<td>November 1970</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>March 1975</td>
<td>+1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>July 1980</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>November 1982</td>
<td>+1</td>
<td>−1</td>
<td>−1</td>
<td>−1</td>
<td></td>
</tr>
<tr>
<td>March 1991</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>−1</td>
<td></td>
</tr>
<tr>
<td>November 2001</td>
<td>0</td>
<td>−3</td>
<td>×</td>
<td>−1</td>
<td></td>
</tr>
</tbody>
</table>

Note: XCI denotes the Stock and Watson coincident index.
Note that since both $CI_t$ and $CI^{MM}$ takes GDP as a further common component, the resulting turning points might be considered an average between the single behavior of $XCI$ and GDP. This evidence is even confirmed in the period subsequent to 2000 where, for example, $XCI$ signals 6 months before the peak of March 2001, GDP is late of 2 months, whereas $CI_t$ results in the middle ($-3$).

The plot of the exponential of $CI_t$ along with $XCI$ and the monthly estimate of GDP is shown in Figure 3. The evidence is that $XCI$ emphasizes the amplitude of cycles because it is not linked with the level of economy as $CI_t$. Finally, Figure 4 presents the comparison between $CI_t$ and $XCI$, splitting the sample period in the four sub-periods $1946-58$, $1959-71$, $1972-1984$ and $1985-2003$. Note that in general $CI_t$ is smoother than $XCI$.

6.2 Results obtained for the Euro Area

In this application we make use of the following seasonal adjusted time series database:

1. $IIP$, the monthly industrial production index from 1980.1 to 2003.6, $1995 = 100$, source OECD. This series is corrected for the well known outlier in June 1984;

2. $SLS$, the monthly retail sales series in volume from 1995.1 to 2003.6, $1995 = 100$, source OECD;

3. $EMP$, the quarterly employment measure from 1980.1 to 2003.1, source European Central Bank (ECB);

4. $GDP$, quarterly rate of the real Gross Domestic Product at 1995 prices from 1980.1 to 2003.1, source OECD.

As the US exercise we take the logarithms of the original data, following the same identification assumptions and adopting an $AR(1)$ specification for both $\mu_t$ and $\mu^*_t$. The
Table 4: Estimated coincident index model for Euro Area

<table>
<thead>
<tr>
<th>Parameters × 100</th>
<th>IIP</th>
<th>SLS</th>
<th>EMP</th>
<th>GDP</th>
</tr>
</thead>
<tbody>
<tr>
<td>θ</td>
<td>0.795</td>
<td>0.244</td>
<td>0.058</td>
<td>0.407</td>
</tr>
<tr>
<td>(0.091)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>β × 100</td>
<td>0.154</td>
<td>0.175</td>
<td>0.006</td>
<td>0.306</td>
</tr>
<tr>
<td>(0.059)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>d_1</td>
<td>-0.319</td>
<td>-0.471</td>
<td>0.884</td>
<td>-0.837</td>
</tr>
<tr>
<td>(0.222)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>σ_{η^∗} × 100</td>
<td>0.538</td>
<td>0.764</td>
<td>0.043</td>
<td>0.221</td>
</tr>
<tr>
<td>(0.126)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

\[(1 + 0.491L) \Delta \mu_t = \eta_t, \quad \eta_t \sim N(0, 1)\]

Note: Standard errors in parenthesis.

resulting model is then characterized by \( N = 4 \) and \( n = 102 \), with a SSF of dimension 12; the estimation is carried from 1980.1 to 2003.6. Results of estimation are in Table 4:

As in Table 1 we consider in parentheses the asymptotic standard errors. The log likelihood for the approximating model is \( \mathcal{L} = 2083.46 \). The model shows a general good fit and the parameters are all significant with the exception of the autoregressive coefficient of \( IIP \). With respect to the model for the US, the simpler specification used in this case derives from the shorter estimation period and a higher irregularity in the series representative of the Euro Area.

The autoregressive coefficients are all negative but the sign related to employment (\( EMP \)), which shows a strong persistence (0.884). The negative sign is particularly high for \( GDP \) (−0.837) and for the coincident index \( \mu_t \) (−0.491).

Table 5 presents the diagnostics based on the KF innovations. For the 4 series we consider the measure of the Box-Ljung statistics \( Q(8) \) and \( Q(12) \), together with \( Norm \) and \( H(h) \) where both take the same meaning of Table 2 and \( h = 93 \) for \( IIP \) and \( SLS \) and \( h = 31 \) for \( EMP \) and \( GDP \). The results suggests a satisfactory specifications for all the equations. The Normality test is generally good but the series of Employment, probably
Table 5: Diagnostics on residuals for the Euro Area model

<table>
<thead>
<tr>
<th>Tests</th>
<th>IIP</th>
<th>SLS</th>
<th>EMP</th>
<th>GDP</th>
</tr>
</thead>
<tbody>
<tr>
<td>Q(8)</td>
<td>18.411</td>
<td>48.407</td>
<td>8.079</td>
<td>7.062</td>
</tr>
<tr>
<td>Q(12)</td>
<td>29.960</td>
<td>50.700</td>
<td>10.336</td>
<td>9.659</td>
</tr>
<tr>
<td>Norm</td>
<td>2.444</td>
<td>0.995</td>
<td>28.073</td>
<td>1.271</td>
</tr>
<tr>
<td>H(h)</td>
<td>0.799</td>
<td>1.704</td>
<td>0.377</td>
<td>0.556</td>
</tr>
</tbody>
</table>

Note: Q(8) and Q(12) are the Box-Ljung statistics based, respectively on 8 and 12 residual autocorrelations, Norm is the Bowman-Shenton Normality test and H(h) is the test for heteroskedasticity (h=93 for IIP and SLS and h=31 for the other series).

for a further outlier at the beginning of the eighties. On the opposite, the residual autocorrelation in the series of IIP and SLS can be explained for the simplified model adopted for the idiosyncratic component. Finally, the heteroskedasticity test is never significant.

Despite these diagnostics, we do not adjust the AR(1) component specification because, as the US exercise, we look at the good fit of the model, at its simplicity, at the meaning of turning points identified by the coincident index and, finally, at the results obtained in terms monthly distributed GDP.

The distributed values of GDP are shown in Figure 5 together with the coincident index. The first is obtained through the decumulated smoothing estimates related to GDP, and the latter adding to the smoothing estimates of the driftless coincident index, the drift estimated for GDP. In this regard the expression (10) is applied.

A comparison of our results with reference dates is quite complicated because official dates are not yet available for the Euro Area. Consider that the Eurozone only come into being on the 1st January 1999 and the study of business cycles needs a larger sample than three-and-a-half years.

However, we compare our results with the quarterly estimates of GDP coming from the following International Institutions: the first is given by GDP_{ECB}, i.e. the ECB estimates by Fagan et al. (2001) available for the period 1970-1998; the second is GDP_{Eurostat}, i.e.
Table 6: Turning points for alternative Euro Area indicators

<table>
<thead>
<tr>
<th>CEPR</th>
<th>GDP_{ECB}</th>
<th>GDP_{OECD}</th>
<th>GDP_{Eurostat}</th>
<th>GDP_{Monthly}</th>
<th>C_{t}</th>
</tr>
</thead>
<tbody>
<tr>
<td>Peaks</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1974q3</td>
<td>1974q3</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>1980q1</td>
<td>1980q1</td>
<td>1980q1</td>
<td>-</td>
<td>1980m1</td>
<td>1980m1</td>
</tr>
<tr>
<td></td>
<td>-</td>
<td>1982q2</td>
<td>-</td>
<td>1982m4</td>
<td>1982m4</td>
</tr>
<tr>
<td>1992q1</td>
<td>1992q1</td>
<td>1992q1</td>
<td>1992q1</td>
<td>1992m3</td>
<td>1992m2</td>
</tr>
<tr>
<td></td>
<td>-</td>
<td>-</td>
<td>2001q3</td>
<td>2001m3</td>
<td>2001m3</td>
</tr>
<tr>
<td>Troughs</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1975q1</td>
<td>1975q1</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>-</td>
<td>1981q1</td>
<td>-</td>
<td>1981m1</td>
<td>1981m1</td>
</tr>
<tr>
<td>1982q3</td>
<td>1982q4</td>
<td>1982q3</td>
<td>-</td>
<td>1982m12</td>
<td>1982m12</td>
</tr>
<tr>
<td>1993q3</td>
<td>1993q1</td>
<td>1993q2</td>
<td>1993q1</td>
<td>1993m2</td>
<td>1993m2</td>
</tr>
<tr>
<td></td>
<td>-</td>
<td>-</td>
<td>2001q4</td>
<td>2001m11</td>
<td>2001m11</td>
</tr>
</tbody>
</table>

the official Eurostat series available from 1991 only; the third is the quarterly GDP used in our exercise, denoted as GDP_{OECD}. Furthermore, we consider the chronology recently established by the Centre for Economy and Policy Research, (CEPR hereafter, see Artis et al., 2003) for the Euro Area in the period 1970-98.

Table 6 presents this simple comparison between the CEPR chronology and the turning points of GDP_{ECB}, GDP_{OECD}, GDP_{Eurostat}, along with our monthly estimates of GDP (GDP_{Monthly}) and the Coincident Index (C_{t}). The chronology of the first group is in terms of quarters, whereas the second is in terms of months.

A first evidence is that there is a substantial coherence among Institutions in turning points of GDP, which are perfectly coincident apart the troughs of 1982 and 1993. For these, there is a difference of one quarter between the ECB and OECD estimates.

The CEPR chronology is in line with movements registered by GDP. However it does not consider the weak cyclical fluctuations occurred to the Euro Area GDP in the period 1981-82. Apart this, CEPR establishes quite late the end of the recession of 1992-93 (3th quarter of 1993) with respect to the trough registered by the alternative measures of GDP.
Even the monthly estimates obtained applying our approach confirm this evidence; in fact both GDP and $CI_t$ show a trough at February 1993. The end of recession of 1982 seems to go to December, in line with the GDP measure of ECB. Finally, going to the most recent years, the monthly measures indicate a weak peak at March 2001, followed by a trough at November 2001, perfectly in line with the US chronology.

For a comparison between the US and the Euro Area business cycle look at Figure 6; it shows the two coincident indexes obtained using our proposal. The overall appearance is that the index for the US is generally smoother and it shows more clearly all the turning points. For recent years the Euro Area does not experienced the same decline and recovery as the US. This phenomenon can be explained by the pattern of employment, which in Europe has continued a slow growth in 2001, whereas in the US has not.

Looking at the first decades, in the eighties the two short recessions occurred in the US are quite undefined for the Euro Area, whereas in the nineties the recession has been experienced first in the US. In general, the analysis of the coincident indexes here computed confirm that there is not a clear lead-lag relation between the US ant the Euro Area.

7 Conclusion

The paper develops a methodology of a new coincident index of business cycle following the SW probability approach. It has been applied to the US and for the first time to the Euro Area. The method combines a set of monthly and quarterly series, considering the GDP along with the coincident series originally selected by SW.

The inclusion of GDP in the set of coincident series has been originally proposed by MM (2002). Nevertheless, we have considered an alternative and more attractive solution based upon a nonlinear state space model. The SW coincident index has been expressed in the levels of the the series, modelling the logarithms like the original proposals. The approach has revealed several advantages: the most important are the nonlinear treatment of temporal aggregation for models in logarithms and the monthly estimate of GDP.
The application to the US popular dataset reaches results in line with the original proposals giving, however, a monthly estimate of GDP from 1946; the application to European data derives for the first time a coincident business cycle indicator for the Euro Area.

References


[2] Boschan and Banerjee (1988);


[6] Engle and Watson (1981);


[8] Geweke (1977);


[14] Moore (1958);


[16] Shiskin (1961);

Figure 1: Coincident indexes and GDP for the US

- Industrial Production
- Employment
- Retail Sales
- Income
- GDP
Figure 2: Coincident indexes and GDP for the Euro Area
Figure 3: Comparison between different coincident indexes and monthly GDP for the US
Figure 4: Comparison of alternative Coincident indexes for the US in four subsamples
Figure 5: The Coincident Index and monthly GDP for the Euro Area
Figure 6: Coincident indexes for the US the Euro Area in the period 1980.1-2003.6