Towards a real-time system for the analysis of the euro area cyclical situation

Gian Luigi Mazzi, Filippo Moauro, Gaetana Montana, and Rosa Ruggeri Cannata
Towards a real-time system for the analysis of the euro area cyclical situation

Gian Luigi Mazzi, Filippo Moauro, Gaetana Montana, and Rosa Ruggeri Cannata

2016 edition
Europe Direct is a service to help you find answers to your questions about the European Union.

Freephone number (*):
00 800 6 7 8 9 10 11

(*) The information given is free, as are most calls (though some operators, phone boxes or hotels may charge you).


ISSN 2315-0807
doi: 10.2785/787175
Cat. No: KS-TC-16-004-EN-N

© European Union, 2016

Reproduction is authorised provided the source is acknowledged.
For more information, please consult: http://ec.europa.eu/eurostat/about/our-partners/copyright

The information and views set out in this publication are those of the author(s) and do not necessarily reflect the official opinion of the European Union. Neither the European Union institutions and bodies nor any person acting on their behalf may be held responsible for the use which may be made of the information contained therein.
Abstract

This paper describes the construction of a real-time monitoring system of the euro area economy by complementing official statistics by advanced statistical and econometric techniques. It also shows how it is possible to build up an effective cyclical early warning system for the euro area by extracting relevant cyclical signals from official statistics by means of linear and non-linear filtering techniques. The construction of both the real-time monitoring and the cyclical early warning system is described in a step by step approach and at each step the results of empirical applications, usually based on real-time simulations, are presented. Despite the fact that improvements are still needed, the proposed applications show that they provide a reliable picture of the current economic situation. They also provide policy makers with stable and non-misleading information about the current cyclical situation.

Acknowledgements

Paper first presented at the 7th International Institute of Forecaster’s Workshop on “Flash Indicators” in Verbier, Switzerland (21-22 January 2011). The authors are grateful to the participants of the workshop for their useful comments, remarks and suggestions.

Keywords

Real-time monitoring; cyclical early warning systems; dynamic factor models; temporal disaggregation; nowcasting; density nowcasting; turning point detection; Markov-Switching models; output gap estimates; linear detrending filters.

Authors

Gian Luigi Mazzi(1), Filippo Moauro (2), Gaetana Montana(3) and Rosa Ruggeri Cannata(4)

(1) Gian Luigi Mazzi, Eurostat, Luxembourg; email: gianluigi.mazzi@ec.europa.eu
(2) Filippo Moauro, Istat, Rome, email: moauro@istat.it
(3) Gaetana Montana, European Parliament, Brussels, gaetana.montana@europarl.europa.eu
(4) Rosa Ruggeri Cannata, Eurostat, Luxembourg, email: rosa.ruggeri-cannata@ec.europa.eu
## Contents

Abstract ................................................................................................................................. 3
1. Introduction ......................................................................................................................... 3
2. Back-calculation ................................................................................................................ 8
3. Real-time database ............................................................................................................ 12
4. Increasing data timeliness: nowcasting techniques ............................................................ 14
   4.1. Density nowcasts ......................................................................................................... 17
5. Increasing data timeliness: coincident indicators ............................................................... 23
6. Euro-MIND: A euro area monthly indicator of the economic activity ................................ 26
   6.1. The dynamic factor model, its statistical treatment and temporal disaggregation ........... 27
   6.2. Extensions of Euro-Mind and ongoing activities .......................................................... 30
7. Euro area turning point detection ...................................................................................... 31
8. Growth cycle estimates ...................................................................................................... 36
9. Conclusions ....................................................................................................................... 37
References ............................................................................................................................. 38
1. Introduction

The regular monitoring of the economic situation requires, especially under uncertain conditions, a complete, timely and reliable set of statistics and statistical indicators giving a clear picture of economic movements.

The euro area system of short-term statistics has been built up in the 90’s with the adoption of a set of new legal acts. Since the beginning of the new century, with the definition of the Principal European Economic Indicators (PEEIs), Eurostat has launched an enhancement strategy mainly focusing on the reduction of publication delays and on the progressive harmonisation of data production processes at national and European level.

Furthermore, concrete actions have been undertaken to fill the availability gap especially in the areas of services and labour-market statistics. Nowadays, almost all PEEIs are available and tangible results have been achieved concerning the timeliness, especially thanks to the use of flash estimation techniques (i.e. GDP and HICP) or to early estimates based on an incomplete set of information (i.e. Retail Trade).

The identification of a comprehensive set of statistical indicators can constitute an important contribution to give a clear picture of the economic situation for a specific country or region. In this respect, Eurostat has decided to publish in 2007 a set of short term indicators to help users in correctly understanding the economic situation. This set of indicators is composed of 22 variables, most of them PEEIs, integrated and completed by few key monetary and financial indicators and business surveys ones. Table 1 presents a list of key indicators extracted from the PEEIs page as available in December 2010.

Table 1: Selected euro area(1) PEEIs (as available in December 2010)

<table>
<thead>
<tr>
<th>Indicator</th>
<th>Unit</th>
<th>2009q02</th>
<th>2009q03</th>
<th>2009q04</th>
<th>2010q01</th>
<th>2010q02</th>
<th>2010q03</th>
</tr>
</thead>
<tbody>
<tr>
<td>GDP in volume</td>
<td>% (Q/Q-1)</td>
<td>-0.1</td>
<td>0.4</td>
<td>0.2</td>
<td>0.4</td>
<td>1.0</td>
<td>0.4</td>
</tr>
<tr>
<td>Inflation (HICP all items)</td>
<td>M/M-12</td>
<td>1.4</td>
<td>1.7</td>
<td>1.6</td>
<td>1.8</td>
<td>1.9</td>
<td>1.9</td>
</tr>
<tr>
<td>Unemployment rate-total</td>
<td>%</td>
<td>10.0</td>
<td>10.0</td>
<td>10.0</td>
<td>10.0</td>
<td>10.0</td>
<td>10.1</td>
</tr>
<tr>
<td>Employment</td>
<td>% (Q/Q-1)</td>
<td>-0.6</td>
<td>-0.5</td>
<td>-0.2</td>
<td>0.0</td>
<td>0.0</td>
<td></td>
</tr>
<tr>
<td>Industrial producer prices</td>
<td>% (M/M-1)</td>
<td>0.3</td>
<td>0.2</td>
<td>0.1</td>
<td>0.3</td>
<td>0.4</td>
<td></td>
</tr>
<tr>
<td>Industrial production</td>
<td>% (M/M-1)</td>
<td>-0.1</td>
<td>0.1</td>
<td>1.1</td>
<td>-0.7</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Economic Sentiment Indicator</td>
<td>index</td>
<td>99.0</td>
<td>101.1</td>
<td>102.3</td>
<td>103.3</td>
<td>103.8</td>
<td>105.3</td>
</tr>
</tbody>
</table>

Source: Authors’ calculations

(1) Data refer to the EA-16 aggregate (euro area with 16 Member States) with the exception of Inflation (HICP all-items), which refers to the evolving euro area aggregate, i.e. euro area with 11 Member States till 31/12/2000, euro area with 12 Member States from 01/01/2001 till 31/12/2006, euro area with 13 Member States from 01/01/2007 till 31/12/2007, euro area with 15 Member States from 01/01/2008 till 31/12/2008, euro area with 16 Member States from 01/01/2009 till next euro area enlargement.

This has represented one of the main inputs for the Principal Global Indicators (PGIs) initiative, undertaken by the InterAgency Group after a G-20 request in 2009.
PGIs aim to ensure reliable follow-up of the economic situation for the larger and developed economies. The global financial and economic crisis of 2007-2009 has emphasised the importance of an accurate follow-up of the global economic situation and has stimulated further reflections towards a real-time monitoring and a construction of an early warning system. These topics have been extensively discussed in a series of seminars organised by UNSD and Eurostat together with other international organisations and with the participations of several national statistical offices from all around the world.

The main finding of these seminars is that, although PEEIs and PGIs constitute the core of an early warning system, they need to be integrated and completed by additional estimates and composite indicators to fulfil the requirement of real-time monitoring and to be considered as an effective early warning system.

To achieve such aim, the main improvements for PEEIs and PGIs can be synthesised as follows:

1) further improvements of timeliness are required without a significant decrease of accuracy (more flash estimates and nowcasting);

2) the time coverage of series is not always appropriate for economic modelling and business cycle analysis (definition of a strategy for the reconstruction of long time series);

3) an effective monitoring of the economic situation requires at least a monthly follow-up so that relevant indicators have to be available at least at monthly frequency; which is not always the case (construction of new monthly indicators);

4) compiling composite indicators to extract signals and to fill the specific gaps in data availability of official statistics (cyclical estimates, turning points dating and detection, coincident and leading indicators).

Point three requires some additional reflections. In the past, the Industrial Production Index was considered as the reference variable for the short-term analysis. With the continuous growth of the services sector and the consequently decrease of the industrial one, the role of the Industrial Production Index has also been substantially re-thought. Additionally, the empirical evidence of cyclical movements in the services sector has convinced business cycle analysts of the opportunity of finding a new and more comprehensive measure of the economic activity. This belief has been even reinforced by the fact that in the recent years, some industrial cycles have not generated fluctuations for the whole economy. Finally, the high degree of volatility of the Industrial Production Index, which could prevent a clear identification of signals, has also played a role in diminishing the relevance of this variable. For those reasons ongoing studies aim to produce a more significant monthly indicator for business cycle purposes likely based on GDP (monthly GDP or a proxy of it). Starting from this consideration, it is obvious that other relevant indicators need to be estimated on a monthly basis such as employment, balance of payments and wages indicators.

The issue raised in point 4 above is that statistics themselves cannot answer to all users' needs especially because data give a general measure of a phenomenon. Analysts, often, want to extract signals from data in order to emphasise some specific cyclical features. Statistics are also characterised by some gaps and specific indicators can be used to fill them and complete the information set to be put at disposal of policy makers and analysts.

Improvements of statistics by means of statistical and econometric techniques, as well as compiling new indicators and composite indicators using existing statistics, have not been considered traditionally part of official statistics. Nevertheless, in the recent years, more and more statistical agencies have been involved in such kind of activities. It appears also quite clear that official statisticians, due to their deep knowledge of data and of production systems, are in a privileged position. They can build up new estimates and indicators based on statistically sound methodologies, strong quality framework, and on transparent, replicable and well documented procedures.

This paper shortly presents some ongoing activities in Eurostat related to the improvements of PEEIs and to the construction of new indicators in order to help analysts and policy-makers to correctly interpret the economic situation.
In section 2 we discuss alternative ways to increase data coverage and we present some preliminary results concerning the reconstruction of long time-series for some PEEIs. In section 3 we describe the PEEIs real-time database which is an essential tool both for revision analysis and for real-time simulation of new indicators and estimates. In sections 4 and 5 we illustrate two alternative ways to increase the timeliness: the first one based on traditional nowcasting techniques and the second one on the construction of composite coincident indicators. In section 6 we describe our approach to the construction of a euro area monthly indicator of economic activity, which can be viewed as a very good proxy of the GDP at monthly frequency. Sections 7 and 8 present the statistical framework for business cycle analysis developed at Eurostat. This framework mainly consists of two parts: a turning points dating and detection system and a set of growth cycle estimates. Section 9 presents some conclusions and lines for further investigations and improvements. In each section, the proposed methodology is illustrated with some empirical applications to the euro area.
2. Back-calculation

The availability of a set of long time series covering, whenever possible, several cycles is essential for modelling purposes and for producing reliable cyclical estimates. Unfortunately, this sometimes contradicts the aim of official statistics to keep definitions, classifications and methods as much as possible in line with the current economic structure. Changes in methodology, definitions and classifications unavoidably shorten current time series due to the fact that it is not always possible to recalculate back in time the series according to the new methodology, definitions and classifications. In order to fill this gap in statistics, statistical offices, international organisations as well as final users, often perform back-calculation exercises based on statistical and/or econometric techniques. These techniques mainly rely on the hypothesis that different temporal segments of a given time series (e.g. the Industrial Production Index) are available, each one corresponding to calculation based on different methods, definitions and classifications according to the period of time covered, and that they are at least partially overlapping. These hypotheses allow then the reconstruction of homogeneous long time series based on several partially superposing segments. The reconstruction can be based on very simple statistical techniques such as those based on the ratio between two segments or on more sophisticated techniques based on regression models or time series methods. This approach to back-calculation is a very conservative one since it does not rewrite the past history but it keeps it unchanged. This is a big advantage because it preserves the coherence between the time series movements and the economic policy decisions taken in the past. When choosing the most appropriate approach in order to obtain acceptable results both from the statistical and economic point of view, it is essential to look if the series to be back-calculated are invertible or not according to the invertibility theorem, see Granger (1978). Obviously in case of non-invertibility, it is highly recommended not to use any dynamic specification in the back-calculation exercise. In the case of the euro area, the back-calculation exercise is particularly complex because euro area series are obtained by means of aggregation/estimation methods using national series as an input. In this context the back calculation exercise for euro area has to take into account the availability of the national long time series, the share of available long time series for each sub-period of time and the aggregation/estimation approach used to build up euro area figures for each macroeconomic variable. Eurostat has developed a sequential approach for the back-calculation of euro area aggregates, which is presented below. This approach can be applied to the EA 12 aggregate, which is the largest euro area aggregate consistent back in time even before the nineties, or to the moving EA aggregate.

The methodology for the Back-Calculation of Economic Time Series is presented in Caporin, Sartore "Methodological aspects of time series back-calculation" (Working Papers 2005). The methodology can be synthetically depicted by the following steps:

1) Planning

1.1 definition of the target series \( x_t \) and reference production standards. Chosen standards allow limiting the search for possible indicators among the series produced according to the selected standards.

1.2 definition of the target horizon, M, the back-calculation "optimal" horizon or a data-driven specification of M that allows the researcher to extract all the information from the available data.

2) Data retrieving.

This step is a fundamental one. Related series and indicators must be retrieved for a given reference series and target horizon, considering that:

2.1 sources: search must cover time series values from different data providers (e.g. National Statistical Institutes, National Banks, international organisations), as well as production methodologies (including information on the series definition, the possible adjustments for seasonality, working days, outliers).

2.2 series dimension: search must consider temporal, spatial and sectoral aggregated or disaggregated values of \( x_t \) as well as series measuring the same quantities but according to different definitions;

2.3 series related to \( x_t \): search must consider proxies of \( x_t \) which could be used when the analysis of \( x_t \) series dimension produces poor results.
Towards a real-time system for the analysis of the euro area cyclical situation

The collected data must be subjected to a first screening, at least graphical, in order to identify possible errors: typing, superimposition of series on different scales or reference years etc.

3) Selection of the strategy

Perform a comparative analysis to identify the best strategy. The model choice strictly depends on the available data, in particular on their coverage and on their quality. If we call Z the set of available information, then

3.1 Aggregation and disaggregation problems:

a) whenever Z contains a complete and exhaustive disaggregation of $x_t$ for $t \in \{-M, -M + 1, -M + 2, \ldots, -2, -1\}$ (violating condition 2.1) we are facing an aggregation problem (spatial or sectoral);

b) whenever Z contains a complete and exhaustive temporal aggregation of $x_t$ for $t \in \{-M, -M + 1, -M + 2, \ldots, -2, -1\}$ (violating definition 2.2) we are facing a temporal disaggregation problem;

c) whenever Z contains a complete and exhaustive temporal disaggregation of $x_t$ for $t \in \{-M, -M + 1, -M + 2, \ldots, -2, -1\}$ (violating definition 2.3) we are facing a temporal aggregation problem.

3.2 The preferred back-calculation model must be chosen among a set of possibilities considering that:

- the model should be defined on a data-driven basis
- the model can also depend on the ‘feasible’ target horizon.

When analysing a back-calculation exercise with indicators at the same frequency of the target series, we can initially consider two classes of regression models:

Static models, of the form

$$y_t = \alpha + \beta x_t + \epsilon_t$$

Dynamic models, of the form

$$y_t = \alpha + \delta y_{t-1} + \beta_1 x_t + \beta_2 x_{t-1} + \epsilon_t$$

In this second category we can further consider:

- ARIMA models
- ADL models
- ECM models

Unfortunately, none of such models can be fully satisfactory for a back-calculation exercise. The first one, expressed in levels, will be miss-specified if the target series and the indicators are characterised either by stochastic or deterministic trends which is quite often the case. Furthermore, the target variables and the indicators can also be co-integrated. The second model is inappropriate if the target variable is not invertible, which is usually the case. For this reason we have adopted a simple model which reduces the miss-specification effects of the first model and is compatible with the non-invertibility constraint. Such a model can be obtained as a restricted version of the dynamic one by assuming $\delta = \beta_1 = 1$ so that the model can be rewritten as:

$$\delta y_t = \alpha + \beta x_t + \epsilon_t$$

Where $\alpha$ can be 0, $\beta > 0$ and $\epsilon_t$ has MA(1) representation.

In the special case where $\beta = 1$, then the reconstructed pattern of the target variable will be identical to the one of the indicator. If $\alpha < 1$ then the reconstructed pattern of the target variable will show the same movements as the indicator but it will be less volatile while the opposite occurs when $\beta > 1$.

When dealing both with indicators at the same frequency of the target series and at a lower frequency, we face constrained back-calculation and temporal disaggregation or benchmarking techniques can be used.
4) Production and Maintenance

Once a particular model has been chosen, we can apply it to data. It is important to consider also an updating strategy; past values should generally be stable, however new past values for the used indicators could become available making necessary the run of a new back-calculation exercise.

Those principles have been applied in several back-calculation exercises. Caporin and Sartore (Working Papers 2005) present different strategies for back-calculating the Industrial Production Index of the European Union aggregate with 15 members (EU15). Several alternative methods were further explored to carry out high quality reconstructions of long time-series for PEEIs. In particular, a comprehensive validation strategy has been carried out for back-calculation of the Industrial Production Index, Producer prices, Turnover index, Unemployment, Employment, Building permits, Nights spent in Hotels and National Accounts and some components series back to the '70s, as detailed in the table below. For national accounts the consistency between current prices and volume series on the one hand, and annual and quarterly series on the other hand, have been carefully investigated.

Table 2: Experimental back-calculated series and relative horizons

<table>
<thead>
<tr>
<th>Indicator name</th>
<th>Breakdown</th>
<th>Starting date</th>
</tr>
</thead>
<tbody>
<tr>
<td>Industrial Production</td>
<td>NACE divisions and MiGs</td>
<td>1970</td>
</tr>
<tr>
<td>Producers Prices</td>
<td>NACE divisions and MiGs</td>
<td>1970</td>
</tr>
<tr>
<td>Turnover index</td>
<td>NACE divisions and MiGs</td>
<td>1974</td>
</tr>
<tr>
<td>Retail Trade</td>
<td>Food, Non food</td>
<td>1970</td>
</tr>
<tr>
<td>Unemployment</td>
<td>Male, Female, under and over 25</td>
<td>1970</td>
</tr>
<tr>
<td>Employment</td>
<td>Employees, self-employees NACE A6</td>
<td>1970</td>
</tr>
<tr>
<td>Building permits</td>
<td>Total</td>
<td>1970</td>
</tr>
<tr>
<td>Nights spent in Hotels</td>
<td>Resident/Not resident</td>
<td>1990</td>
</tr>
<tr>
<td>Wages and salaries</td>
<td>NACE A6</td>
<td>1971</td>
</tr>
<tr>
<td>National Accounts main</td>
<td>NACE A6</td>
<td>1970</td>
</tr>
<tr>
<td>aggregates</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Source: Author's calculations

Figure 1 presents the results of the back-calculated EA12 IPI series: in the first panel the monthly levels of the series are presented; the second panel is devoted to the analysis of the annualised indexes in terms of % growth rates; here the EA12 back-calculated series is compared to the old indexes of Germany, France and Italy; finally, the third panel compares the growth rates in % of the EA12 annualised back-calculated series with a proxy of the euro area aggregate computed with old data of all the Member States joining the euro area EA12 with the exception of Ireland. This proxy is denoted as EA11-old, it covers the sample period from January 1970 and it represents the most relevant related indicator to EA12 based on old data only.

Even if the described back-calculating approach has been originally developed for European aggregates, it is easily applicable also by Member States. In a medium term perspective, the use of a common back-calculated approach within the ESS could increase the comparability of national back-calculated series and the reliability of European aggregates.
Figure 1: Back-calculated monthly EA12 IPI and comparison of the annualised index with euro area and Member State related indicators of old data

Source: Author’s calculations
3. Real-time database

Beside the availability of long time series that has already been discussed in section two, it is crucial for the development and assessment of new indicators or nowcasts to have a real-time database that allows building up real-time simulation exercises on the base of data as available at the release date on a long enough period of time.

Central bankers, policy makers and economic forecasters have been paying increasing attention to the fact that their econometric models may be sensitive to revisions in several ways; revisions on independent variables can cause revisions to the value of the dependent variable; moreover, they could cause revisions to model parameters. Both events could therefore affect the usefulness of their models. This has generated a considerable demand for databases which provide access to first (and subsequent) releases of data for key economic variables.

When economists undertake research to construct an econometric model, they test their models’ performance over a given time span. Nevertheless, they are generally restricted to use latest available vintages of the variables they wish to include in the model. However, often out-of-sample data used to evaluate the model consist of time series where each data point will most likely have been revised from its first published value, whereas it is these first published values that would have been used by the model when applied in practice. Having access to first-published estimates for the out-of-sample analysis overcomes this problem and allows a more realistic assessment of the expected performance of the models. In other words, it is necessary to test models’ performance on historical originally released (and then subsequently revised) data, simulating the likely performance of models in real-time.

Furthermore, real-time data provide economists with the necessary information for testing the sensitivity of model parameters to revisions in the independent variables when constructing their models. When empirical results are sensitive to the vintage of the data, the validity of results or the method behind them should be better investigated. Sound empirical methods should be robust to data vintages and not oversensitive to small data changes.

Unfortunately, real-time databases currently available at euro area level are based upon snapshots of data at fixed intervals (usually at a different frequency of data updates). This has two side effects: suppose that an indicator is revised the day after the snapshot is added to the database, the new vintage will appear in the database just at the next snapshot, possibly with some delay. Moreover, when the indicator is revised several times in between two snapshots, the database will contain just one instead of several vintages. Daily snapshots will clearly overcome such limitations. Furthermore, there is a certain lack of real-time database at national level for euro area member states.

In order to meet users’ requirements in this context, Eurostat has decided to develop a real-time database for the euro area, the European Union and Member states starting from the Euroind database available on the Euro-indicators/PEEIs dedicated section of the Eurostat website (see “The PEEIs Real-Time Database”, Dominique Ladiray, Gian Luigi Mazzi and Rosa Ruggeri-Cannata, Cambridge, EABCN Workshop 27-28 March 2008). The Euroind database contains a wide range of macro-economic variables classified in the following eight main categories: Balance of payments, Business and consumer surveys, Consumer prices, External trade, Industry, commerce and services, Labour market, Monetary and Financial Indicators and National Accounts.

The Euroind database is mainly based on Principal European Economic Indicators (PEEIs), integrated and completed by some monetary and financial indicators (mainly produced by the European Central Bank) as well as business and consumer surveys (produced by the Directorate General for Economic and Financial Affairs in the context of the Joint Harmonised EU Program of the European Commission).
The real-time database is built up from the daily snapshots of the Euroind database, available from 16 November 2000 onwards. Such large availability of daily snapshots over the time allows Eurostat to reconstruct a real-time database covering a quite long time period (about ten years) which is particularly useful to study the revisions behaviour of macro-economic statistics, to perform simulation exercises concerning new estimates and/or indicators and to monitor the evolution of the database quality across time.

We are currently in the fine-tuning phase of the PEEIs real-time database; however, we have extensively used it to perform real-time simulations of all nowcasts and construction of new indicators presented in the following sections of this paper. The PEEIs real-time database turned out to be extremely useful in discriminating between alternative models and in supplying useful information on the behaviour over the time of nowcasts and new indicators. In particular, the feedback obtained in real-time simulations allowed refining, assessing and improving the chosen methodologies. The following sections will introduce several examples of those applications, covering real-time simulation, robustness analysis, comparison of alternative models and methodological improvements.
4. Increasing data timeliness: nowcasting techniques

Users and especially institutional ones consider the timeliness of official statistics as a key requirement. Theoretically they should welcome to receive real-time or quasi real-time information on main economic indicators. On the other hand, data producers are often reluctant to anticipate the release date because they worry about the loss of accuracy that such anticipation could determine.

The trade-off between timeliness and accuracy is a key aspect to be considered when deciding on the timing of macroeconomic statistics releases. For the euro area the timeliness is not only relevant in absolute terms but also in comparative terms with respect to the US release dates. The timeliness of US macroeconomic statistics is traditionally superior to the European one. The only exception is represented by the case of Consumer Prices where we anticipate the US release thanks to the production of our flash estimates.

The increase of timeliness can be achieved by different strategies:

1) the speeding up of the data production process, thanks to the use of advanced survey techniques, simplification of questionnaires, use of administrative information, etc.;
2) the use of statistical and econometric techniques to extrapolate indicators by using incomplete information set (flash estimates, nowcasting, etc);
3) the use of so called EU sampling techniques, which consist in the construction of European samples which do not necessarily generate national significant ones;
4) the construction of composite coincident indicators for main economic variables, which give an idea of their latest and current trends.

The first strategy is clearly a structural one which can be feasible in the medium-long term and which can be considered, from the producers' point of view, the most suitable way to approach this issue.

The third strategy implies a definition of a European sampling scheme applicable to the most relevant surveys on which key economic variables are based (e.g. the industrial production index, quarterly labour force survey, etc.). This solution is difficult to be implemented in statistical areas such as National Accounts which are mainly a synthesis of macroeconomic information coming from various statistical areas. Until now, Eurostat has obtained interesting results based, at least partially, on this approach in the field of Retail Trade Turnover. In this paper we mainly present some ongoing studies related to the second and the fourth strategy.

The use of statistical and econometric techniques can significantly contribute to the increase of timeliness in the short and medium term. In this context, the key tool is represented by a set of forecasting techniques adapted to estimate the recent past or the present instead of the near future. Since 2006, we have been working on a simulation study aiming to produce nowcasts for Producer Price Index at t+16, Employment at t+30 and, more recently GDP at t+0, t+15 and t+30 days. The nowcasts, which we are working on, are based on the following main principles agreed inside Eurostat:

- whenever partial information on the target variable, either at geographic or sectoral level is available, this has to be included in the flash estimation model;
- soft data (e.g. business and consumer surveys) can be integrated into the model under the condition that a minimum amount of hard data is available;
- in order to increase forecasting accuracy, statistically related indicators (e.g. conventional earnings in case of nowcasts of Labour Cost Index) can be used in the model either in case of unavailability of significant partial information on the target variable or to complement such partial information;
- any hypothesis based on economic theories has to be avoided in the model specifications;
Increasing data timeliness: nowcasting techniques

Towards a real-time system for the analysis of the euro area cyclical situation

- purely univariate models should not be taken into account but only used as a benchmark in a simulation exercise;
- the selected model should be as simple as possible, statistically sound, easy to use in the regular production process and characterised by a very simple dynamic specification where needed.

In our simulation exercises we pay particular attention to the identification of the most appropriate variable selection strategies and to the comparison of alternative model specifications to achieve the best performance of estimates. Combining forecasting techniques are also tested to increase the reliability of estimates.

In this section we will present two examples of nowcasts, based on the Eurostat nowcasting strategy for Producer Price Index and for the GDP. Basically, we use regression based methods to produce the nowcasts. This involves considering a reasonably large set of possible indicators and their lags, and then estimating many models. Importantly, national as well as Euro indicators variables are considered too. The BIC criterion is then used recursively to select the best single model. On the basis of these out-of-sample simulations we identify the best model specification and the variables to be retained.

Table 3: Euro area Flash Estimates of GDP Growth

<table>
<thead>
<tr>
<th>Year</th>
<th>Nowcast t+15 days</th>
<th>Eurostat First estimate</th>
<th>Eurostat Final estimate</th>
<th>Error First estimate</th>
<th>Error Final estimate</th>
</tr>
</thead>
<tbody>
<tr>
<td>2008q04</td>
<td>-0.94</td>
<td>-1.59</td>
<td>-1.86</td>
<td>-0.64</td>
<td>-0.91</td>
</tr>
<tr>
<td>2009q01</td>
<td>-1.55</td>
<td>-2.56</td>
<td>-2.01</td>
<td>-1.01</td>
<td>-0.47</td>
</tr>
<tr>
<td>2009q02</td>
<td>-1.13</td>
<td>-0.18</td>
<td>-0.15</td>
<td>0.96</td>
<td>0.99</td>
</tr>
<tr>
<td>2009q03</td>
<td>0.68</td>
<td>0.42</td>
<td>0.43</td>
<td>-0.26</td>
<td>-0.25</td>
</tr>
<tr>
<td>2009q04</td>
<td>0.74</td>
<td>0.04</td>
<td>0.2</td>
<td>-0.7</td>
<td>-0.54</td>
</tr>
<tr>
<td>2010q01</td>
<td>1.18</td>
<td>0.2</td>
<td>0.34</td>
<td>-0.97</td>
<td>-0.84</td>
</tr>
<tr>
<td>2010q02</td>
<td>0.83</td>
<td>0.95</td>
<td>0.95</td>
<td>0.12</td>
<td>0.12</td>
</tr>
<tr>
<td>2010q03</td>
<td>0.66</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Source: Author’s calculations

Table 3 shows for the period Q4 2008 – Q3 2010 the real-time simulation of our t+15 nowcast for the GDP against the Eurostat first and final estimates for the euro area.

For GDP the preferred model uses two months of within quarter IPI data and the latest quarter’s value of the Economic Sentiment Indicator (“soft” survey data, published by DG ECFIN).

There is no sign discordance between the nowcast from our model and the Eurostat estimates; however some significant errors are present. In particular, they indicate an underestimation during the recession phases (i.e. 2009q02) and overestimation in expansion ones (i.e. 2009q04, 2010q01). Probably the model, though correct, needs some improvements.

Concerning the under-estimation and over-estimation of the first and second quarter of 2009, it can be explained by the procedures of outliers correction within the Seasonal Adjustment methods used by the Member States, which are difficult to be incorporated into the nowcasting exercise. Finally, it has to be noted that the current practise to derive euro area seasonally adjusted data for GDP by summing up national seasonally adjusted data makes difficult the construction of a nowcasting model for the euro area as a whole using a direct approach.

Table 4 shows the real-time simulation of the nowcast of the Producer Price Index against the Eurostat first and final estimates for the euro area for the period April 2007 – August 2010. The selected model is based on German industrial output price and energy prices data according to the BIC selection criterion. Results appear quite encouraging: despite the presence of two cases of sign discordance, often the nowcast is quite close to Eurostat first estimates. The regression model reacts quite well to the change of regime even if a bit slowly. The variables have been selected among a large set of potential variables. Our
opinion is that this constitutes a very good starting point, which could be easily improved either by adding more national information or by including euro area Import Prices which were not yet available when we started the simulation.

Table 4: Euro Area nowcast of Producer Price Index

<table>
<thead>
<tr>
<th></th>
<th>Nowcast t+16 days</th>
<th>Eurostat First estimate</th>
<th>Eurostat Final estimate</th>
<th>Error First estimate</th>
<th>Error Final estimate</th>
</tr>
</thead>
<tbody>
<tr>
<td>2007m4</td>
<td>0.31</td>
<td>0.45</td>
<td>0.42</td>
<td>0.14</td>
<td>0.11</td>
</tr>
<tr>
<td>2007m5</td>
<td>0.28</td>
<td>0.29</td>
<td>0.43</td>
<td>0.01</td>
<td>0.15</td>
</tr>
<tr>
<td>2007m6</td>
<td>0.31</td>
<td>0.13</td>
<td>0.12</td>
<td>-0.18</td>
<td>-0.19</td>
</tr>
<tr>
<td>2007m7</td>
<td>0.25</td>
<td>0.26</td>
<td>0.29</td>
<td>0.01</td>
<td>0.04</td>
</tr>
<tr>
<td>2007m8</td>
<td>0.31</td>
<td>0.06</td>
<td>0.15</td>
<td>-0.25</td>
<td>-0.16</td>
</tr>
<tr>
<td>2007m9</td>
<td>0.24</td>
<td>0.36</td>
<td>0.41</td>
<td>0.12</td>
<td>0.17</td>
</tr>
<tr>
<td>2007m10</td>
<td>0.33</td>
<td>0.64</td>
<td>0.67</td>
<td>0.31</td>
<td>0.34</td>
</tr>
<tr>
<td>2007m11</td>
<td>0.78</td>
<td>0.90</td>
<td>0.92</td>
<td>0.12</td>
<td>0.14</td>
</tr>
<tr>
<td>2007m12</td>
<td>0.11</td>
<td>0.10</td>
<td>0.19</td>
<td>-0.01</td>
<td>0.08</td>
</tr>
<tr>
<td>2008m01</td>
<td>0.43</td>
<td>0.85</td>
<td>0.78</td>
<td>0.42</td>
<td>0.35</td>
</tr>
<tr>
<td>2008m02</td>
<td>0.74</td>
<td>0.66</td>
<td>0.61</td>
<td>-0.08</td>
<td>-0.13</td>
</tr>
<tr>
<td>2008m03</td>
<td>0.65</td>
<td>0.70</td>
<td>0.66</td>
<td>0.05</td>
<td>0.01</td>
</tr>
<tr>
<td>2008m04</td>
<td>0.97</td>
<td>0.79</td>
<td>0.77</td>
<td>-0.18</td>
<td>-0.20</td>
</tr>
<tr>
<td>2008m05</td>
<td>0.93</td>
<td>1.21</td>
<td>1.16</td>
<td>0.28</td>
<td>0.23</td>
</tr>
<tr>
<td>2008m06</td>
<td>0.90</td>
<td>0.96</td>
<td>1.06</td>
<td>0.06</td>
<td>0.16</td>
</tr>
<tr>
<td>2008m07</td>
<td>1.73</td>
<td>1.23</td>
<td>1.33</td>
<td>-0.50</td>
<td>-0.40</td>
</tr>
<tr>
<td>2008m08</td>
<td>-0.25</td>
<td>-0.50</td>
<td>-0.51</td>
<td>-0.25</td>
<td>-0.26</td>
</tr>
<tr>
<td>2008m09</td>
<td>0.23</td>
<td>-0.20</td>
<td>-0.18</td>
<td>-0.43</td>
<td>-0.41</td>
</tr>
<tr>
<td>2008m10</td>
<td>0.01</td>
<td>-0.80</td>
<td>-0.87</td>
<td>-0.81</td>
<td>-0.88</td>
</tr>
<tr>
<td>2008m11</td>
<td>-1.26</td>
<td>-1.96</td>
<td>-2.14</td>
<td>-0.70</td>
<td>-0.88</td>
</tr>
<tr>
<td>2008m12</td>
<td>-1.19</td>
<td>-1.58</td>
<td>-1.54</td>
<td>-0.39</td>
<td>-0.35</td>
</tr>
<tr>
<td>2009m01</td>
<td>-0.79</td>
<td>-0.88</td>
<td>-1.22</td>
<td>-0.09</td>
<td>-0.43</td>
</tr>
<tr>
<td>2009m02</td>
<td>-0.57</td>
<td>-0.43</td>
<td>-0.49</td>
<td>0.14</td>
<td>0.08</td>
</tr>
<tr>
<td>2009m03</td>
<td>-0.69</td>
<td>-0.71</td>
<td>-0.69</td>
<td>-0.02</td>
<td>0.00</td>
</tr>
<tr>
<td>2009m04</td>
<td>-1.25</td>
<td>-1.06</td>
<td>-0.87</td>
<td>0.18</td>
<td>0.38</td>
</tr>
<tr>
<td>2009m05</td>
<td>-0.18</td>
<td>-0.17</td>
<td>-0.05</td>
<td>0.01</td>
<td>0.13</td>
</tr>
<tr>
<td>2009m06</td>
<td>-0.09</td>
<td>0.31</td>
<td>0.42</td>
<td>0.40</td>
<td>0.50</td>
</tr>
<tr>
<td>2009m07</td>
<td>-1.16</td>
<td>-0.77</td>
<td>-0.69</td>
<td>0.38</td>
<td>0.46</td>
</tr>
<tr>
<td>2009m08</td>
<td>-0.32</td>
<td>0.48</td>
<td>0.51</td>
<td>0.80</td>
<td>0.83</td>
</tr>
<tr>
<td>2009m09</td>
<td>-0.27</td>
<td>-0.36</td>
<td>-0.35</td>
<td>-0.09</td>
<td>-0.08</td>
</tr>
<tr>
<td>2009m10</td>
<td>0.04</td>
<td>0.27</td>
<td>0.29</td>
<td>0.23</td>
<td>0.25</td>
</tr>
<tr>
<td>2009m11</td>
<td>0.11</td>
<td>0.18</td>
<td>0.17</td>
<td>0.07</td>
<td>0.06</td>
</tr>
<tr>
<td>2009m12</td>
<td>0.09</td>
<td>0.10</td>
<td>0.06</td>
<td>0.01</td>
<td>-0.03</td>
</tr>
<tr>
<td>2010m01</td>
<td>0.65</td>
<td>0.66</td>
<td>0.73</td>
<td>0.01</td>
<td>0.07</td>
</tr>
<tr>
<td>2010m02</td>
<td>0.10</td>
<td>0.13</td>
<td>0.13</td>
<td>0.03</td>
<td>0.03</td>
</tr>
<tr>
<td>2010m03</td>
<td>0.56</td>
<td>0.61</td>
<td>0.59</td>
<td>0.05</td>
<td>0.03</td>
</tr>
<tr>
<td>2010m04</td>
<td>0.80</td>
<td>0.93</td>
<td>0.98</td>
<td>0.13</td>
<td>0.17</td>
</tr>
<tr>
<td>2010m05</td>
<td>0.43</td>
<td>0.30</td>
<td>0.28</td>
<td>-0.13</td>
<td>-0.15</td>
</tr>
<tr>
<td>2010m06</td>
<td>0.59</td>
<td>0.30</td>
<td>0.34</td>
<td>-0.29</td>
<td>-0.25</td>
</tr>
<tr>
<td>2010m07</td>
<td>0.51</td>
<td>0.26</td>
<td>0.21</td>
<td>-0.25</td>
<td>-0.30</td>
</tr>
<tr>
<td>2010m08</td>
<td>0.08</td>
<td>0.11</td>
<td>0.11</td>
<td>0.03</td>
<td>0.03</td>
</tr>
</tbody>
</table>

Source: Author’s calculations
We have focused until now on so called “point” nowcasts, which are point estimates delivering no information about the uncertainty related to the estimates; in the next sub-section we will present a technique to take such uncertainty into account.

4.1. Density nowcasts

Timely and reliable information about macroeconomic variables is crucial for policy makers and business cycle analysts. Such need has been emphasised by the recent financial and economic crisis and has claimed for a speeding up of delivery of estimates on the current economic situation. Many studies have presented and compared alternative statistical methods for the computation of timely estimates or “nowcasts”. They can be produced using available soft and hard data, keeping in mind the trade-off between the timeliness and the accuracy of nowcasts. In fact, the quicker production of nowcasts by the use of less hard information and more soft data might deteriorate their quality.

Traditionally, those studies have focused on the "point" estimate, which does not provide any indication of the degree of uncertainty associated with the nowcasts.

A nowcast of, say, 2% of GDP growth means that people would be not surprised if actual growth turns out to be a little larger or smaller. And in times of economic uncertainty, they should not be surprised if it turns out to be much larger or smaller. Consequently to provide a complete description of the uncertainty associated with the point nowcasts, many analysts publish density nowcasts.

More formally, density nowcasts of the GDP growth, say, provide an estimate of the probability distribution of its possible future values. And, as nowcasts affect decisions, the “better” nowcasts allow delivering “better” decision. Publishing nowcasts implies also obvious advantages when communicating with the public: a) analysts themselves expect the point nowcast to be “wrong”; b) users assess the balance of risks associated with the nowcasts (e.g. they can see when Statistical Offices expect more downside than upside risks to GDP growth).

In our work (“Density nowcasts and model combination: Nowcasting Euro-area GDP growth over the 2008-9 recession”, Gian Luigi Mazzi, James Mitchell, Gaetana Montana (2010)), we nowcasted the euro area quarterly GDP growth by using density nowcasts and model combination. We constructed it ahead of Eurostat's Flash estimate which is at 45 days, over a period which includes the recent economic crisis. The exercise uses real-time data and allows for the staggered release of monthly information on indicator variables throughout the quarter.

We produced density nowcasts from a large set of components models, which differ in terms of the indicator variables they consider. We suggest the use of simple regression based methods. Density nowacasts from the component models are then derived analytically.

The nowcasts are produced by statistical models. These statistical models seek to explain and then nowcast GDP growth by exploiting information on indicator variables, which are meant to have a close relationship with GDP but available more promptly than the data they stand for as a proxy.

Following Giannone et al. (2008), we distinguished between quantitative (“hard”) and qualitative (“soft”) indicator variables, the latter typically published ahead of hard data. We focused on simple components models. We estimated a linear regression of quarterly GDP growth on a single indicator variable. Then we combined the component density nowcasts using the linear pool approach, as proposed by Timmerman (2006).

We produced nowcasts of quarterly GDP growth for the euro area to five time scales: t-30, t-15, t+0, t+15 and t+45 days; the last being produced for benchmarking with Eurostat flash estimate for the last quarter.

Since a monthly indicator, by construction, is released three times a quarter following Kitchen and Monaco (2003), we estimate three component models for each indicator following:

$$\Delta y_t = \beta_0 + \beta_1 x_{m_{k,t}} + e_t$$

(1)

where $\Delta y_t$ is the quarterly GDP growth, $e_t$ is assumed to be normally distributed, $x_{m_{k,t}}$ is the $k^\text{th}$ indicator variable from the information set $\Omega_j$; $m = 1,2,3$ denotes the month in the quarter $t$, $(t = 1, \ldots, T); j = 1, \ldots, 5$
Increasing data timeliness: nowcasting techniques

Towards a real-time system for the analysis of the euro area cyclical situation

denotes the first, second, third, fourth and fifth nowcast formed at t-30, t-15, t+0, t+15 and t+45 days. Each successive nowcast exploits an ever larger information set. This reflects the fact that with the passage of time more and more indicator data become available.

We did also experiment with a sixth nowcast at t+30 days, on receipt of retail trade data, but these data did not help and are not considered further.

The indicator variables considered were: as soft data the Economic Sentiment Indicator (ESI, published by the European Commission), and the spread between short term and 10 year Euro interest rates (available from the ECB). And the monthly indicator variables were transformed. Indeed, when relating the monthly variables \( x_{hard,t} \) and \( x_{soft,t} \) to quarterly GDP growth \( \Delta y \), there is an issue about how these monthly data should be transformed.

The quarterly transformation of the monthly survey data involves transforming \( x_{m,k,t} \) in a manner consistent with the quarterly variable \( \Delta y_t \) (which represents quarterly growth at a quarterly rate). This is achieved, for example, following Mariano & Murasawa (2003).

We then combined the density by the linear opinion pool approach, as in Timmerman (2006). Given \( i = 1, \ldots, N_j \) component models, the combination densities for GDP growth are given by the linear opinion pool:

\[
p(\Delta y_t) = \sum_{i} w_{i,t,j} g(\Delta y_t | \Omega^j_t)
\]

where \( g(\Delta y_t | \Omega^j_t) \) are the nowcast densities from component model \( i, i = 1, \ldots, N_j \) of \( \Delta y_t \), conditional on the information set \( \Omega^j_t \).

These densities are obtained having estimated the (1).

The non-negative weights, \( w_{i,t,j} \), which sum to unity, are constructed in two ways: equal weights (EW) strategy and the Recursive Weight (RW) strategy. The EW strategy attaches equal (prior) weight to each model with no updating of the weights through the recursive analysis:

\[
w_{i,t,j} = w_{i,j} = 1/N_j
\]

Secondly, we construct the weights \( w_{i,t,j} \) based on the fit of the individual model forecast densities: the Recursive Weight (RW) strategy. Following Jore et al. (2010) and Garratt, Mitchell & Vahey (2009), we use the logarithmic score to measure density fit for each model through the evaluation period. The logarithmic scoring rule is intuitively appealing as it gives a high score to a density forecast that assigns a high probability to the realised value. The recursive weights for the nowcast densities take the form:

\[
n_{h,t,j} = \frac{\exp \left[ \sum_{\tau=1}^{T-1} \ln g(\Delta y_{\tau} | \Omega^j_{t-\tau}) \right]}{\sum_{j=1}^{N} \exp \left[ \sum_{\tau=1}^{T-1} \ln g(\Delta y_{\tau} | \Omega^j_{t-\tau}) \right]}, \quad \tau = t, \ldots, T
\]

In constructing the combined densities using the linear opinion pool, we evaluate the density forecasts using the logarithmic score at each recursion. We emphasise that in deriving the weights based on this measure of density fit, the component models are repeatedly evaluated using real-time data. These weights provide an indication of whether the support for the component models is similar, or not, based on the score of the individual densities. A finding of similar weights across component models would be consistent with the equal-weight strategy.

The next step is the evaluation of nowcast density combination. A popular method, following Rosenblatt (1952), Dawid (1984) and Diebold et al. (1998), is to use the probability integral transforms (pits) \( z_t \) of the realisation of the variable with respect to forecast density:

\[
z_t = \int_{-\infty}^{\Delta y_t} p(u) \, du
\]

and the application of tests for goodness of fit as Likelihood Ratio (LR), as proposed by Berkowitz (2001), Anderson-Darling (AD) for uniformity of the pits (a modification of Kolmogorov-Smirnov) and
Increasing data timeliness: nowcasting techniques

The empirical application

We compare the accuracy of nowcasts of Euro-area GDP growth at the five horizons \((j = 1, \ldots, 5)\) in recursive out-of-sample experiments using real-time data. Specifically, we use the real-time data for real GDP and industrial production available from Eurostat. Nowcasts for GDP growth are computed recursively from 2003q2-2010q1 using the density forecast combination approach at \(j = 1, \ldots, 5\); i.e., \(t-30\), \(t-15\), \(t+0\), \(t+15\) and \(t+45\) days. We break our results into two parts: the RW weights on the soft indicators, the hard indicators and lagged GDP growth derived from the logarithmic score of the component forecast densities; and, the evaluations of the recursive weight, RW, and equal weight, EW, strategies for combination.

Figure 2 presents the Recursive Weights on the soft indicators (i.e., survey data and interest rate spread), hard indicators (i.e. IP) and lagged values of GDP growth for the five nowcasting horizons, \(j = 1, \ldots, 5\).

Figure 2: Recursive weights of soft and hard indicators for five nowcasting horizons

Source: Author’s calculations
Figure 2 shows interesting results:

- The weight on IPI increases as \( j \) increases: more hard data available, higher weight in the combined density (improvement out-of-sample density fit).
- At \( t+15 \) days IPI weight increases dramatically relatively to survey data, approaching one at the end of evaluation period.
- But during the recession, weights of soft data increases at \( t-30 \), \( t-15 \) and \( t+0 \), close to unity during the depth of the recession. While weights on IPI declined rapidly during the recession and rose as it ended. From \( t+15 \), weight remains high on IPI data even over the recessionary period.
- AR hard to beat when point nowcasting. With density nowcasting, weights of AR components non negligible, although declining as within quarter information accumulates: at \( t+45 \) less than half the weight than \( t-30 \).

There is always an issue about how to choose the length of the training period to calibrate the weights. In fact, there is a trade-off involved. The shorter the length of the training period the quicker the combined density can adjust to changes over time in the performance of the different models. But the longer the length of the training period the better the combination weights are estimated.

Figure 3 shows the recursively computed log score weights.

**Figure 3: Recursively computed log score weights**
In fact, Figure 3 shows more clearly than Figure 2 that:

- informational content of survey data increased suddenly during 2008;
- soft data picked up the recession more quickly than hard IPI data;
- when IPI data published, at t+45 days, survey data do not increase in importance during the recession;
- how the weight on the IPI data increases as $j$ increases.

Finally interesting results are also shown by the figures 4 and 5: Probability of recession.

Figures 4 and 5 extract from density nowcasts the implied probability of a (one period) recession. It is evident that:

- RW combined densities pick up earlier the recession than hard IPI data, as increasingly put higher weight on soft data during the recession.
- Soft data also correctly picked up the end of recession in euro area in a way not predicted by the combined densities. Their dependence on history to tune the weights means that they cannot adjust as quickly as soft data.
- But on average over the evaluation period, nowcast density produced by survey data alone is not well calibrated.
- Waiting for second month’s IPI data, RW combined densities clearly anticipate both the beginning and the end of recession earlier than EW combined density.

**Figure 4: Implied probability of one-step-ahead recession**

Source: Author’s calculations
The Relative importance of indicator variables switched suddenly in the recession. In fact, during the recession, soft data were more informative than hard data on IPI (when nowcasting with one month of within quarter information on IPI available). Equally important is the length of period used to tune the weights in combined densities: when abrupt switch in utility of different indicator variables, equal weighted combined densities deliver more accurate density nowcasts than recursive weighted ones. Equal weighted are more robust to uncertain instabilities, which are particularly acute nowcasting earlier within the quarter. At t+15 days, with second month of within quarter IPI data published, recursive weighted combined densities become more competitive because it shifts more weight, during the recession, from IPI data to the forward-looking survey data.
5. Increasing data timeliness: coincident indicators

Coincident indicators aim to forecast the evolution of economic variables during the reference period or just after it. For this purpose, they are based on the same principles as leading indicators. The main advantage of coincident indicators is that they are subject to fewer constraints than flash estimates or nowcasts, even if, once again, the use of economic relationship is not recommended. In the recent years, we have investigated alternative model specifications for GDP, IPI and more recently Employment. The results for the IPI are not very satisfactory also because of the high degree of volatility of this indicator. In this section we briefly present our approach to construct a coincident indicator of GDP.

We have used several models, belonging to the same category: regression models. They differ in the type of regressors they embed: individual series in bridge models against principal components in factor models. At the beginning bridge models were preferred to factor ones because they were based on an as large as possible data set. Afterwards, our interest was focussed on the type of data to be included in the models: soft data only or a combination of soft and hard data (i.e. IPI growth rate and the financial variable euro/dollar exchange rate growth).

Indeed, recent developments in the area of factor models pointed out that factor extraction could be better based on small data sets than large ones. This led us to focus on factor models, and we run them on real-time data.

Factor models are built in the spirit of Stock and Watson (2002), that is the data set is reorganised into principal components. Factors are then introduced in a regression to explain the target variable (for example, GDP growth) and the statistically significant ones are kept in the regression. The main difference between our factor models and the Stock and Watson ones is the size of data set: we use a small dataset whereas principal components are usually extracted from a large data set. The principal components extraction is carried out on standardised data, firstly stationarised if necessary. We then regress the target variable (GDP growth) on these principal components and an intercept. We finally select the significant principal component and add financial data to these factors.

The very first factor models were constructed on small data set with data series chosen by a subjective criterion based on past experience. Afterwards, we introduced a more objective data selection process, using the LARS algorithm introduced by Bai and Ng (2008). The LARS algorithm selects the targeted predictors (i.e. the most appropriate variables to estimate the quarterly GDP growth) not as aggressively as other stepwise algorithms do; in fact it allows keeping correlated series which is a desirable characteristic when the final objective is to extract principal components. Series in the initial data set are not eliminated; they are all ranked by decreasing predictive power according to the selection criterion of the LARS algorithm.

However, we used the LARS algorithm with some differences from Bai and Ng one. Firstly, some series are introduced several times in the data set (with different lags) as we expect the LARS algorithm will show coincident or leading features of the series. Secondly, soft data series can be introduced both in levels and variation. That can help to forecast; the LARS algorithm is able to rank rapidly a large number of series starting from the most predictive ones downward. An additional difference is that principal components are extracted from a dataset which does not contain financial series, series which are then introduced directly in the regression (we frequently observed that a financial series generates a principal component on its own).

We have built and compared three different models, see Charpin, Mathieu, Mazzi (2008): 1) a bridge model containing hard and soft data (named BHS), 2) a first factor model, with factors built with soft and hard data (named FHS), 3) a second factor model, with factors constructed exclusively with soft data (named FS). The latter is built in order to estimate coincident GDP growth when no hard data is available for the quarter. The bridge model includes the variables: Industrial Production Index, Construction Output Index, Consumer opinion over next 12 months, Employment expectations in construction, Construction confidence indicator and the euro/dollar real exchange rate. The factor models include survey data (Industry, Consumers, Construction, Retail Trade) and hard data (Industrial Production excluding construction, Construction production, exports, retail sales and unemployment rate).

For each quarter we produce three estimates of GDP: the first at the end of the second month (t-30), the
Increasing data timeliness: coincident indicators

Towards a real-time system for the analysis of the euro area cyclical situation

The following table shows the real time results obtained by using the three models for data selection over the period Q4 2005 – Q3 2010, complemented by the LARS algorithm from 2009. We compare the three estimates of the coincident indicator to the Eurostat flash. The results appear very encouraging.

Table 5: Coincident Indicator of the GPD Growth

<table>
<thead>
<tr>
<th>End of Quarter T</th>
<th>Models</th>
<th>Estimates T-30</th>
<th>Estimates T+0</th>
<th>Estimates T+30</th>
<th>Eurostat Flash T+45</th>
</tr>
</thead>
<tbody>
<tr>
<td>2005Q4</td>
<td>FS</td>
<td>0.65</td>
<td>0.72</td>
<td>0.72</td>
<td>0.31</td>
</tr>
<tr>
<td></td>
<td>BHS</td>
<td>0.50</td>
<td>0.58</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>FHS</td>
<td>0.41</td>
<td>0.47</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2006Q1</td>
<td>FS</td>
<td>0.77</td>
<td>0.82</td>
<td>0.82</td>
<td>0.59</td>
</tr>
<tr>
<td></td>
<td>BHS</td>
<td>0.83</td>
<td>0.80</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>FHS</td>
<td>0.81</td>
<td>0.78</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2006Q2</td>
<td>FS</td>
<td>0.93</td>
<td>0.97</td>
<td>0.97</td>
<td>0.88</td>
</tr>
<tr>
<td></td>
<td>BHS</td>
<td>0.81</td>
<td>0.90</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>FHS</td>
<td>0.89</td>
<td>0.87</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2006Q3</td>
<td>FS</td>
<td>0.65</td>
<td>0.69</td>
<td>0.72</td>
<td>0.52</td>
</tr>
<tr>
<td></td>
<td>BHS</td>
<td>0.73</td>
<td>0.82</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>FHS</td>
<td>0.74</td>
<td>0.89</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2006Q4</td>
<td>FS</td>
<td>0.68</td>
<td>0.67</td>
<td>0.66</td>
<td>0.90</td>
</tr>
<tr>
<td></td>
<td>BHS</td>
<td>0.49</td>
<td>0.41</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>FHS</td>
<td>0.69</td>
<td>0.69</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2007Q1</td>
<td>FS</td>
<td>0.52</td>
<td>0.55</td>
<td>0.55</td>
<td>0.57</td>
</tr>
<tr>
<td></td>
<td>BHS</td>
<td>0.50</td>
<td>0.49</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>FHS</td>
<td>0.59</td>
<td>0.53</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2007Q2</td>
<td>FS</td>
<td>0.56</td>
<td>0.57</td>
<td>0.57</td>
<td>0.34</td>
</tr>
<tr>
<td></td>
<td>BHS</td>
<td>0.32</td>
<td>0.28</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>FHS</td>
<td>0.53</td>
<td>0.47</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2007Q3</td>
<td>FS</td>
<td>0.75</td>
<td>0.66</td>
<td>0.65</td>
<td>0.71</td>
</tr>
<tr>
<td></td>
<td>BHS</td>
<td>0.76</td>
<td>0.96</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>FHS</td>
<td>0.72</td>
<td>0.93</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2007Q4</td>
<td>FS</td>
<td>0.35</td>
<td>0.34</td>
<td>0.34</td>
<td>0.41</td>
</tr>
<tr>
<td></td>
<td>BHS</td>
<td>0.32</td>
<td>0.33</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>FHS</td>
<td>0.38</td>
<td>0.43</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2008Q1</td>
<td>FS</td>
<td>0.12</td>
<td>0.13</td>
<td>0.13</td>
<td>0.79</td>
</tr>
<tr>
<td></td>
<td>BHS</td>
<td>0.28</td>
<td>0.41</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>FHS</td>
<td>0.43</td>
<td>0.55</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2008Q2</td>
<td>FS</td>
<td>0.09</td>
<td>0.03</td>
<td>0.02</td>
<td>-0.20</td>
</tr>
<tr>
<td></td>
<td>BHS</td>
<td>0.13</td>
<td>0.02</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>FHS</td>
<td>0.06</td>
<td>-0.03</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2008Q3</td>
<td>FS</td>
<td>-0.25</td>
<td>-0.35</td>
<td>-0.35</td>
<td>-0.19</td>
</tr>
<tr>
<td></td>
<td>BHS</td>
<td>-0.33</td>
<td>-0.09</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>FHS</td>
<td>-0.25</td>
<td>-0.09</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2008Q4</td>
<td>FS</td>
<td>-0.72</td>
<td>-0.99</td>
<td>-0.99</td>
<td>-1.46</td>
</tr>
<tr>
<td></td>
<td>BHS</td>
<td>-0.91</td>
<td>-1.06</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>FHS</td>
<td>-0.89</td>
<td>-1.10</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 5: Coincident Indicator of the GPD Growth (continued)
Increasing data timeliness: coincident indicators

Towards a real-time system for the analysis of the euro area cyclical situation

<table>
<thead>
<tr>
<th>End of Quarter T</th>
<th>Models</th>
<th>Estimates T-30</th>
<th>Estimates T+0</th>
<th>Estimates T+30</th>
<th>Eurostat Flash T+45</th>
</tr>
</thead>
<tbody>
<tr>
<td>2009Q1</td>
<td>FS</td>
<td>-0.95</td>
<td>-1</td>
<td>-1.06</td>
<td>-2.55</td>
</tr>
<tr>
<td></td>
<td>BHS</td>
<td>-1.4</td>
<td>-1.48</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>FHS</td>
<td>-1.56</td>
<td>-1.63</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2009Q2</td>
<td>FS</td>
<td>-0.61</td>
<td>-0.45</td>
<td>-0.45</td>
<td>-0.1</td>
</tr>
<tr>
<td></td>
<td>BHS</td>
<td>-0.76</td>
<td>-0.59</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>FHS</td>
<td>-0.56</td>
<td>-0.51</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2009Q3</td>
<td>FS</td>
<td>0.32</td>
<td>0.23</td>
<td>0.4</td>
<td>0.4</td>
</tr>
<tr>
<td></td>
<td>BHS</td>
<td>0.36</td>
<td>0.74</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>FHS</td>
<td>0.19</td>
<td>0.35</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2009Q4</td>
<td>FS</td>
<td>0.52</td>
<td>0.57</td>
<td>0.57</td>
<td>0.1</td>
</tr>
<tr>
<td></td>
<td>BHS</td>
<td>0.2</td>
<td>0.35</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>FHS</td>
<td>0.25</td>
<td>0.33</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2010Q1</td>
<td>FS</td>
<td>0.41</td>
<td>0.46</td>
<td>0.46</td>
<td>0.19</td>
</tr>
<tr>
<td></td>
<td>BHS</td>
<td>0.75</td>
<td>0.9</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>FHS</td>
<td>0.38</td>
<td>0.46</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2010Q2</td>
<td>FS</td>
<td>0.4</td>
<td>0.68</td>
<td>0.68</td>
<td>0.97</td>
</tr>
<tr>
<td></td>
<td>BHS</td>
<td>0.65</td>
<td>0.87</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>FHS</td>
<td>0.48</td>
<td>0.77</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2010Q3</td>
<td>FS</td>
<td>0.58</td>
<td>0.69</td>
<td>0.68</td>
<td></td>
</tr>
<tr>
<td></td>
<td>BHS</td>
<td>0.79</td>
<td>0.7</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>FHS</td>
<td>0.7</td>
<td>0.6</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Source: Author's calculations

At t-30 days, estimates are carried out by the factor model with soft data (FS). We can observe that only three quarters are poorly estimated: Q1 2008 and, during the crisis, Q4 2008 and Q1 2009. In those two quarters, there is an over estimation: survey data cannot track the strong fall in GDP. At t+0 and t+30 days, we can see that the magnitude of the recession (2008Q4, 2009Q1) has not been correctly anticipated whatever the model and the error between the estimates of the coincident indicator and the Eurostat flash one is of an unusual size. However, the two models including hard data predict better than the FS the 2009 Q1 fall. Globally, the factor model (FHS) gives better results than the bridge model (BHS). In general, the coincident indicators anticipate correctly the GDP growth in most cases. Further simulations are required, but the coincident indicator presented appears to be robust and reliable enough.
6. Euro-MIND: A euro area monthly indicator of the economic activity

GDP is obviously the ideal candidate as reference variable for short-term and business cycle analysis but, unfortunately, it is only available at quarterly basis and the production of a monthly GDP, completely based on National Accounts standards, still appears not feasible. For this reason, several studies have been recently undertaken to investigate alternative ways to construct monthly proxies of GDP. Examples of such indicators are available in Sweden, Finland, Estonia, U.K. as well as Canada. From 2006 onwards, we have investigated the possibility of constructing a euro area monthly indicator of economic activity as much as possible consistent with the GDP.

The methodology proposed in this section is presented in a detailed way in Frale C., Marcellino M., Mazzi G.L., Proietti T., "A Monthly Indicator of the Euro Area GDP" (2008). The methodology can be synthetically described by the following steps:

1. We base the construction of the monthly indicator of economic activity on a disaggregate approach represented by the output and expenditure breakdowns of the GDP at quarterly level;
2. For each disaggregate GDP component, a set of monthly indicators are carefully selected, including both macroeconomic variables and opinion survey data;
3. The indicator is based on information at both monthly and quarterly level, rather than monthly only, modelled with a dynamic factor specification cast in state-space form. In this step, for each component of GDP, a composite coincident indicator based on dynamic factor analysis is estimated;
4. The state space methodology has the flexibility of handling data with different frequency of observations. This is achieved by suitably defining the states of the system to convert temporal aggregation into a systematic sampling problem;
5. Since estimation of the multivariate dynamic factor model can be numerically complex, computational efficiency is achieved by implementing univariate filtering and smoothing procedures;
6. Special attention is paid to chain-linking and its implications for the construction of a monthly indicator of economic activity, via a multistep procedure that exploits the additivity of the volume measures expressed at the previous year prices;
7. The estimate of the euro area monthly indicator of economic activity is obtained by combining the estimates from the output and expenditure sides, with optimal weights reflecting their relative precision;
8. The resulting pooled estimator is more precise than each of its two components, paralleling the results on the usefulness of pooling in the forecasting literature. The resulting estimates are benchmarked to quarterly national accounts produced by Eurostat so that the full consistency between monthly and quarterly estimates is achieved;
9. We provide an explicit measure of uncertainty around the indicator, which is particularly relevant in a decision making context and for evaluation purposes.
6.1. The dynamic factor model, its statistical treatment and temporal disaggregation

The modelling strategy mentioned at point 3) refers to the Stock and Watson (1991) SW single index model. The fundamental idea behind this specification is to separate the dynamics which are common to a set of \( N \) coincident series, \( y_t \), that are \( I(1) \) but not cointegrated, from the idiosyncratic component, which is specific to each series. The level specification of the SW single index model here considered expresses \( y_t \) as the linear combination of a common factor, that will be denoted by \( \mu_t \), and an idiosyncratic component, \( \mu^*_t \). Letting \( \Phi_0 \) and \( \Phi_1 \) denote \( N \times 1 \) vectors of loadings, and assuming that both components are stationary in first difference and subject to autoregressive dynamics, we can write:

\[
y_t = \Phi_0 \mu_t + \Phi_1 \mu_{t-1} + \mu^*_t + Bx_t, \quad t = 1, ..., n,
\]

where \( \Phi(L) \) is an autoregressive polynomial of order \( p \) with stationary roots

\[
\Phi(L) = 1 - \varphi_1 L - ... - \varphi_p L^p
\]

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

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

with \( d_i(L) = 1 - d_{i1} L - ... - d_{ip} L^p \) and \( \Sigma_{\eta} = \text{diag}(\sigma_1^2, ..., \sigma_N^2) \). The vector \( x_t \) contains the value at time \( t \) of \( k \) deterministic regressors common to all the series, e.g. trading days and moving festivals regressors, and \( B \) is an \( N \times k \) matrix of regression coefficients. The disturbances \( \eta_t \) and \( \eta^*_t \) are mutually uncorrelated at all leads and lags.

The lag polynomial \( \Phi_0 + \Phi_1 L \) can also be rewritten as \( \theta_0 + \theta_1 L \), where \( \theta_0 = \Phi_0 + \Phi_1 \) and \( \theta_1 = -\Phi_1 \). The measurement equation can thus be re-parametrised as:

\[
y_t = \theta_0 \mu_t + \theta_1 \Delta \mu_t + \mu^*_t + Bx_t \quad (3)
\]

The model postulates that each series, in differences, \( \Delta y_{it} \), is composed of a mean term \( \delta_t \), an individual AR(\( p^* \)) process, \( d_i(L)^{-1} \eta^*_t \), and a common AR(\( p \)) process, \( \varphi(L)^{-1} \eta_t \). Both \( \mu_t \) and \( \mu^*_t \) are difference stationary processes and the common dynamics are the results of the accumulation of the same underlying shocks \( \eta_t \); 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.

Notice that (2) assumes a zero drift for the single index, moreover unit variance for its disturbances is an additional hypothesis. These identification restrictions can be removed at a later stage to enhance the interpretability of the estimated common index (we may alternatively restrict to unity one of the loadings in \( \theta_0 \) and include a nonzero drift in the common index equation, provided we impose one linear constraint on \( \beta \)).

The estimation of the model parameters, the factors and the disaggregate GDP monthly component series
is carried out by using the state space methodology (see Harvey, 1989).

The dynamic factor model (2) can be cast in a state space form, consisting of a measurement equation, relating the monthly series to a set of latent states (representing the latent factors and their lags), and a transition equation, describing the dynamic evolution of the states. The basic state space representation has to be modified taking into account the observational constraints imposed by temporal aggregation. The latter is such that we do not observe the monthly values of the GDP components, but only their sum across a quarter. Suppose that the set of coincident indicators, $y_t$, can be partitioned into two groups, $y_t = \{y_{1,t}, y_{2,t}\}$, where the second block gathers the flows that are subject to temporal aggregation, so that:

$$y^*_{2,\tau} = \sum_{i=0}^{\delta-1} y_{2,\tau+i},$$

$$\tau = 1, 2, [T/\delta]$$

where $\delta$ denotes the aggregation interval: for instance, if the model is specified at the monthly frequency and $y^*_{2,t}$ is quarterly, then $\delta = 3$.

The modified state space form is defined in terms of a partially cumulated monthly series, subject to missing values, that converts temporal aggregation into a systematic sampling problem. The cumulator variable, $y^c_{2,t}$, is defined as follows

$$y^c_{2,t} = \psi_r y^c_{2,r-1} + y_{2,t},$$

$$\psi_r = \begin{cases} 0 & t = \delta(r-1) + 1, \tau = 1, \ldots, [n/t] \\ 1 & \text{otherwise} \end{cases}$$

Notice that at times $t = \delta r$ the cumulator coincides with the (observed) aggregated series, otherwise it contains the partial cumulative value of the aggregate in the periods (e.g. months) making up the larger interval (e.g. quarter) up to and including the current one.

The Kalman filter for this state space representation enables the evaluation of the likelihood via the prediction error decomposition. This opens the way to maximum likelihood estimation of the unknown parameters by a quasi-Newton optimisation algorithm. Conditional on the maximum likelihood estimates and the available observations, the estimation of the unobserved components and the missing values (and thus the disaggregated series) is carried out by a suitable smoothing algorithm. For computational efficiency, the Kalman filtering and smoothing equations are implemented using sequential processing (see Anderson and Moore, 1979, and Koopman and Durbin, 2000).
Figure 6: Euro-MIND Growth rate on previous month

Figure 6 presents the growth rate of Euro-MIND from January 2005 to September 2010 as estimated in November 2010, together with their confidence interval at 95%. Looking at the graph, it is important to note that the evolution of the indicator is quite regular and it follows very well the cyclical movements. The estimates appear very stable and not volatile, that is also confirmed by analysing subsequent vintages for the same period. The main point on which the indicator still needs some improvements is represented by its behaviour in estimating the month of the current quarter especially in the recession phase. Our indicator delivers negative growth rates (e.g. January and February 2009), which appear too optimistic in comparison with the expected results. A most accurate specification of the model for the financial services sector and for the demand side component will probably improve the ability of the model to estimate the most recent months.
6.2. Extensions of Euro-Mind and ongoing activities

Several other projects are currently ongoing around Euro-Mind. The first one is the generalisation of this model with better forward looking properties. It is obtained using two factors in the construction of the composite indicators described in step three, where the second one contains business and consumer surveys data. This extended version is presented in Frale, Marcellino, Mazzi, Proietti (2010). This version of Euro-Mind increases its nowcasting and forecasting abilities at one to three steps ahead. This is achieved due to the fact that business and consumer surveys data are modelled as a separate factor from hard data. In fact, if both hard and soft data are modelled in a single factor, soft data are dominated by hard ones so that their contribution to the model is almost zero. A real-time simulation of the one factor based model has been carried out since 2006 with very encouraging results. In this simulation we are producing estimates at \( t+45 \) each month so that, in the month \( t \) we produce the estimate for the month \( t-2 \). At the same point in time, estimates for the month \( t-1, t, t+1 \) can be obtained by using the two factors version of the model.

A second extension of Euro-Mind is represented by an attempt to jointly estimate a monthly indicator of economic activity at euro area and member state level in order to assess the relevance of national information to increase the reliability of euro area estimates. This is a very huge project especially from the computational point of view and results are still very preliminary. Furthermore, we are investigating the possibility of reconstructing Euro-MIND back in time possibly to the eighties and even seventies. In this exercise, due to the lack of information, we are aiming to back-calculate only Euro-MIND and not its components. Nevertheless, the exercise is still quite complicated because many sources of information used in the Euro-Mind model were not available before the eighties.
7. Euro area turning point detection

The set of macroeconomic statistics regularly compiled by a statistical office represents a very useful instrument available to all users and analysts. Nevertheless, we have to recognise that not all the information needed by analysts is explicitly available from the analysis of statistics. Some signals need to be extracted in order to have a clearer picture of the cyclical evolution of the economy, complementing the information supplied by statistics. In this context, we have decided to launch several activities aiming to the definition of a coherent statistical framework for business cycle analysis. They include the construction of statistical turning point chronologies, the development of turning point coincident indicators and estimates of growth cycle (i.e. output gap in the case of GDP), which can support economic monitoring and decision making processes.

The methodology for the construction of a euro area turning point chronology and a system of coincident turning point indicators is presented in Anas, Billio, Ferrara, Mazzi (2008). The methodology can be synthetically described by the following steps:

1) simultaneous analysis of classical business cycle and growth cycle in the so-called ABCD framework, see Anas and Ferrara (2004);
2) statistical dating of classical business cycle and growth cycle euro area turning points by means of a simple non parametric dating rule;
3) comparison of euro area and member states dating chronologies for both classical business cycle and growth cycle to achieve a final statistical chronology ensuring the maximum degree of consistency between the two approaches and the fulfilment of the ABCD sequence. In this context we also investigate the synchronisation of turning points between euro area and member states and their diffusion. Some descriptive measures of cyclical movements such as length of cycle, deepness etc. are also produced. The chronology is updated on quarterly basis;
4) preliminary investigation of alternative models for the construction of turning point composite coincident indicators for classical business cycle and growth cycle, including the identification of appropriate number of regimes and thresholds;
5) concerning the growth cycle, the variable selection has been performed on the basis of the ability of a set of potential candidates series to correctly detect its turning points. For each series a set of transformations has been applied in order to choose the most appropriate one. After this exploratory investigation, five variables have been identified as components of the growth cycle's turning point composite coincident indicators: Employment expectation, Construction confidence indicator, Financial situation over the last 12 months, IPI, Imports of intermediate goods;
6) construction of the growth cycle coincident indicators (GCCI) as a weighted mean of the transition probability returned by the five univariate two regimes Markov Switching models fitted on each variable as it is shown in formula below:

\[
GCCI_t = \frac{1}{5} \sum_{k=1}^{5} Pr(\text{Recession})^k_t,
\]

where \( Pr(\text{Recession})^k_t \) is the probability that the \( k^{th} \) component of the GCCI is in a recession of the growth cycle at time \( t \), with \( k \in \{1,2,3,4,5\} \).

An equal averaging weighting scheme is used.

A K-regime Markov-Switching process, denoted by MS(K) – AR(p), can be defined by the following equation:

\[
Y_t - \mu(S_t) = \sum_{i=1}^{p} \phi_i(S_t)(Y_{t-i} - \mu(S_{t-i})) + \sigma(S_t)e_t,
\]

where the non-observed process \( S_t \) is an ergodic Markov chain of the first order and where \( e_t \) is a standardised white noise process; the parameters describe the dependence of the process \( Y_t \) on the current regime \( S_t \).
The associated transition probability of the process $S_t$ is defined by:

$$\Pr[S_t = j | S_{t-1} = i] = p_{ij}$$

For each release and each model the QPS and Concordance Index are computed as follows.

$$QPS = \frac{1}{T} \sum_{t=1}^{T} (P_t - RC_t)^2,$$

where, for $t \in \{1, ..., T\}, P_t$ is the filtered probability of being in recession in month $t$ and

$$CI = \frac{1}{T} \left[ \sum_{t=1}^{T} I_t \times RC_t + \sum_{t=1}^{T} (1-I_t) \times (1-RC_t) \right],$$

where $RC_t$ is the same variable already employed in the QPS, which represents the turning points of the reference chronology, while $I_t$ is a binary random variable that assumes value 1 if the coincident indicator is in the recessionary phase of the business cycle and 0 otherwise.

7) Concerning the classical business cycle, the variable selection has been performed on the basis of the ability of a set of potential candidates series to correctly detect its turning points. For each series a set of transformations has been applied in order to choose the most appropriate one. After this exploratory investigation, three variables have been identified as components of the business cycle turning point composite coincident indicators: IPI, New cars registrations and Unemployment rate;

8) construction of the business cycle coincident indicators (BCCI) as a weighted mean of the transition probability returned by the three univariate three regimes Markov Switching models fitted on each variable as it is shown below:

$$BCCI_t = \sum_{k=1}^{3} \omega^k \Pr(\text{Recession})^k,$$

where $\Pr(\text{recession})^k_t$ is the probability that the $k^{th}$ component of the BCCI is in a recession of the business cycle at time $t$ and $\omega^k$ is the weight given to it, with $k \in \{1,2,3\}$. 

The following weighting scheme is used IPI=0.34, Unemployment rate=0.46, New cars registration=0.20.

For each release of the BCCI, QPS and Concordance Index are computed in the same way as the GCCI in step 6.

It is important to note that filtered probabilities can be viewed as the probabilities of being in a recession/slowdown phase delivered by each component of the indicators. Indicators deliver the joint recession/slowdown probabilities. A high value of each indicator corresponds to a high probability of being in a recession/slowdown. The threshold (set at 0.5) corresponds to a decision rule: values exceeding the threshold indicate a recession/slowdown phase; values below the threshold correspond to an expansionary/recovery phase.

A real-time simulation of the two indicators against respectively the business cycle and growth cycle chronology has been carried out to check the reliability of the models as well as to discover possible false signals. The main results are that the two indicators do not show any significant evidence of false signals and that they are slightly lagging with respect to the corresponding chronologies. Each month we produce estimates of business cycle coincident indicators and growth cycle coincident indicators for the month $t-2$, based on filtered probabilities. Estimates for the month $t-1$ and $t$ are based on forecasted probabilities. Figures 7/8 and 9/10 show the behaviour of the two indicators GCCI and BCCI as estimated in November 2010. In both graphs the black bold line is the constant threshold equal to 0.5. When the indicators deliver
values higher than 0.5 we are in a slowdown phase of the growth cycle and in a recession phase of the business cycle respectively. On the contrary, when the indicators deliver values below 0.5, we are in a recovery phase of the growth cycle and in an expansion phase of the business cycle respectively. The blue lines show the values of the two indicators obtained by averaging the filtered probabilities of the components. The red part at the end of the line corresponds to the value obtained by averaging forecasted probabilities instead of filtered ones.

Looking at the indicators, the negative phase for the growth cycle started in April 2007 and ended in September 2009 (see Fig. 7/8). Concerning the business cycle, the recession started in October 2008 and ended in September 2009 (see Fig. 9/10). As already mentioned both indicators appear to be slightly lagging and this is particularly true for the BCCI. In fact, we are currently thinking that the business cycle recession has started in the first half of 2008. From this point of view, it is obvious that the BCCI still needs some improvements. Nevertheless, it has to be noted that it is preferable to have indicators detecting later turning points than indicators delivering false signals or too much anticipating turning points.

In order to enhance the cyclical monitoring of the euro area, several initiatives have been recently undertaken and some of them are still ongoing. The first one concerns the extension of the cyclical monitoring to include the acceleration cycle (also referred as growth rate cycle) following the approach \( \alpha AB\beta CD \). The chronology of these three cycles has already been assessed and a first attempt to construct turning point composite coincident indicators for the acceleration cycle is still ongoing. In order to improve the timeliness of BCCI and GCCI we have investigated also other non-linear specifications based on SETAR models. An extensive real-time comparison of indicators based on Markov Switching models and SETAR models has been carried out. The main outcome of this simulation can be synthesised as follows:

1. SETAR models are slightly timelier than MS ones.
2. the number of false signals returned by SETAR models is higher than in the case of Markov Switching ones, even when a censoring rule is applied.
3. composite indicators based on SETAR models are less stable over the time than those based on Markov Switching ones. For this reason it has been decided that SETAR models could only be used to complement the information supplied by Markov Switching ones but that they couldn't replace them. The main results of this study are presented in Billio, Ferrara, Guégan, Mazzi (2013).

Furthermore, the effect of alternative seasonal adjustments methods on detecting turning points as well as the behaviour of a composite coincident indicator for turning points based on non-seasonally adjusted data have also been analysed in real-time in Billio, Carati, Ladiray, Mazzi (forthcoming) and Billio, Carati, Mazzi and Montana (2010). The main outcome of this study is that the use of a common seasonal adjustment method for all components of the composite indicator is preferable with respect to the use of various seasonal adjustment methods.

The composite indicator based on non-seasonally adjusted data has given, in the case of growth cycle, interesting results which open a new interesting field of investigation. Finally, since some months, we are investigating the possibility of using a multivariate Markov Switching model to construct simultaneously a composite coincident indicator for the growth cycle and the business cycle. This approach, among others, has the advantage of explicitly imposing the constraints derived by the ABCD approach. Preliminary results of this work appear very encouraging and we are planning to finalise this experimental phase at the beginning of 2011. In case where the results will appear very positive the new multivariate approach will be used in the regular production and it will replace the two univariate composite indicators presented above.
Figure 7: Growth Cycle Coincident Indicators

Source: Author’s calculations

Figure 8: Growth Cycle Coincident Indicators

Source: Author’s calculations
Euro area turning point detection

Figure 9: Business Cycle Coincident Indicators

Source: Author’s calculations

Figure 10: Business Cycle Coincident Indicators

Source: Author’s calculations
8. Growth cycle estimates

The accurate and timely identification of turning points is a very important source of information for policy makers and analysts. Unfortunately, even the most accurate system of turning points detection does not supply details on the shape and on the main features of the cycle. This is especially true for the growth cycle which needs to be extracted from the seasonally adjusted version of indicators by means of detrending techniques. From the policy makers’ and analysts’ point of view, an accurate estimate of the growth cycle is of crucial importance especially for monitoring the inflationary pressures and for designing a monetary policy oriented to inflation targeting. The main problem we have to deal with when estimating the growth cycle is its instability at the end of sample, due to data revisions from one hand and to specific characteristics of most detrending filters from the other. During the last year, we have compared alternative univariate techniques for the estimation of growth cycle and currently we are regularly publishing three alternative growth cycle estimates based on Hodrick-Prescott filter, on Christiano-Fitzgerald filter and on Unobserved Components models filter in the Eurostatistics publication. We are accompanying such estimates with appropriate meta-information describing the characteristics of alternative procedures. Furthermore, we are investigating several multivariate detrending techniques based on structural VAR and multivariate unobserved component models which are presented in Mitchell, Moauro, Mazzi (paper presented at the 28th Symposium of Forecasters, Nice 2008) and in Lemoine, Mazzi, Monperrus-Veroni and Reynes (2010).

Figure 11 shows the latest EA GDP trend-cycle estimates from 1995Q1 to 2010Q2 using Hodrick-Prescott filter, obtained in November 2010.

As it appears clearly from figure 11, we are in a negative phase of the growth cycle, which confirms the results from the growth cycle coincident indicator shown in figure 8. Due to the endpoint estimation problems characterising detrending filters, some inconsistencies between the signals delivered by the GCCI and the growth cycle estimates cannot be excluded. Users should be informed about this and the producers of business cycle indicators should supply them their appreciation on the reliability of signals coming from different methods.

Figure 11: EA GDP trend-cycle decomposition using HP filter

Source: Author’s calculations
9. Conclusions

This paper has synthetically presented several ongoing Eurostat projects aiming to build up a system of rapid estimates giving a clear picture of the short-term economic situation at euro area level. The results presented in this paper can be considered as very preliminary ones so that further investigations are needed before taking a decision on the communication strategy of this kind of information. Nevertheless, some results are very encouraging and the approaches proposed can be considered methodologically sound, easy to be understood, as well as replicated and adapted to communicate results in a clear and transparent way. In order to improve the overall quality of the estimates presented in this paper, we are working on the following lines:

1) incorporating as much as possible national information into euro area models;
2) investigating, especially in the field of nowcasting, the possibility of constructing estimates using an indirect approach, working at Member States level instead of euro area one;
3) investigating the usefulness of introducing additional data sources into our models, especially in the case of Euro-MIND to increase the reliability of some components estimates;
4) analysing more sophisticated data and models selection techniques;
5) testing alternative specifications of our turning points coincident indicators to reduce their lagging characteristics especially for the BCCI;
6) constructing a chronology and a turning point indicator also for the acceleration cycle;
7) studying alternative solutions, recently proposed in the literature, to increase the reliability of endpoint estimates of detrending filters.

Finally, it is important to note that there are several synergies among the ongoing projects which still need to be better exploited. For example, the coincident indicators of the GDP growth could be used to improve the performance of the Euro-MIND for the current quarter and in perspective the Euro-MIND itself could replace the Industrial Production Index in the specification of both GCCI and BCCI.
References


18. Giannone, D., Reichlin, L., and Small, D., ‘Nowcasting: The Real-Time Informational Content of


HOW TO OBTAIN EU PUBLICATIONS

Free publications:
- one copy:
  via EU Bookshop (http://bookshop.europa.eu);
- more than one copy or posters/maps:
  from the European Union’s representations (http://ec.europa.eu/represent_en.htm);
  from the delegations in non-EU countries (http://eeas.europa.eu/delegations/index_en.htm);
  by contacting the Europe Direct service (http://europa.eu/europedirect/index_en.htm) or calling 00 800 6 7 8 9 10 11 (freephone number from anywhere in the EU) (*).

(*) The information given is free, as are most calls (though some operators, phone boxes or hotels may charge you).

Priced publications: