Available online at www.sciencedirect.com
SCIENCE DIRECT®
International Journal of Forecasting 20 (2004) 447–460
international journal of forecasting
www.elsevier.com/locate/ijforecast

Bridge models to forecast the euro area GDP☆

Alberto Baffigi[a], Roberto Golinelli[b,*], Giuseppe Parigi[a]

[a] Research Department, Bank of Italy, via Nazionale n. 91, 00184 Rome, Italy
[b] Department of Economics, University of Bologna, Strada Maggiore, 45, 40125 Bologna, Italy

Abstract

Quantitative information on the current state of the economy is crucial to economic policy-making and to early understanding of the economic situation, but the quarterly national account (NA) data for GDP in the euro area are released with a substantial delay. The aim of the paper is to examine the forecast ability of bridge models (BM) for GDP growth in the euro area. BM 'bridge the gap' between the information content of timely updated indicators and the delayed (but more complete) NA. In this paper, BM are estimated for aggregate GDP and components both area-wide and for the three main countries of the euro area. Their short-term (one- and two-quarter ahead) forecasting performance is assessed with respect to benchmark univariate/multivariate statistical models, and a small structural model. The paper shows that national BM fare better than benchmark models. In addition, euro area GDP and its components are more precisely predicted by aggregating national forecasts.

Keywords: Short-term GDP forecast for the euro area; Bridge model; Out-of-sample forecasting accuracy

1. Introduction

Information on the current state of economic activity is a crucial ingredient for policy making, as the choice of the appropriate policy stance relies on updated knowledge of the macroeconomic framework. Unfortunately, national account (NA) data become available with a non-negligible delay (currently, in the European countries, 70 days after the end of quarter), which weakens their role as a support to policy making and for an early understanding of the economic situation.

As a matter of fact, much information about short-term economic developments is available well before NA data. A number of short-term indicators, such as business or consumer surveys or the industrial production indexes, are released at a monthly frequency and can be used to get an early picture of the evolution of current economic activity. However, the extraction of a reliable signal from all the indicators available is not a straightforward task and requires complex methodologies (see Zarnowitz, 1992). In the NBER tradition, and in its 'high-tech' versions by Stock and Watson (1999) and Forni, Hallin, Lippi and Reichlin (2000) a major emphasis is attributed to the timing properties of the series-in terms of their leading, coincident or lagging properties-and to their correlation with respect to a reference variable. This

☆ Paper presented at the Cepr–Bocconi University Workshop "Understanding the Structure of the European Business Cycle", Milan, 11 June 2002 and at the Fall meeting of the Project LINK, Bologna, 7–11 October 2002.

* Corresponding author. Tel.: +39-051-209-2600/2638; fax: +39-051-209-2664; http://www.spbo.unibo.it/pais/golinelli/.
E-mail address: golinell@spbo.unibo.it (R. Golinelli).

0169-2070/$ - see front matter
doi:10.1016/S0169-2070(03)00067-0

approach aims at devising (coincident or leading) composite indicators which exploit the high frequency information of the underlying series in order to achieve (timely) detection of the business cycle turning points.

In this paper we take a different route. Our objective is to provide suitable tools to 'translate' the information content of short-term indicators into the more coherent and complete 'language' of NA. In particular, we refer to the 'bridge model' (BM) technique, which allows to compute early estimates of the more relevant NA variables by combining their dynamic properties with short-term indicators (see Parigi & Schlitzer, 1995)[1].

Since indicators cover a wide range of short-term macroeconomic phenomena, they can be used in different bridge equations for the main GDP components (namely, private consumption, government purchases of goods and services, fixed investment, inventory investment, exports, and imports), or directly at aggregate GDP level. In the first case, the model is labelled 'demand-side' BM (where GDP is predicted by the NA income–expenditure identity), in the second case, it is labelled 'supply-side' BM (where GDP is forecast by a single bridge equation[2]).

BM are not concerned with behavioural relations, as the structure underlying the BM is not a standard macroeconometric model: inclusion of specific explanatory indicators is not based on any causal relationship, but on the simple statistical fact that they embody timely updated information about the dependent NA variable. Since BM, in principle, require that the whole set of regressors (lagged endogenous and explanatory indicators) should be known over the projection period, they may be conceived as a tool providing an estimate of the current situation: a 'nowcast', rather than a pure forecast. In practice, however, this is seldom the case: usually only some realisations (months) of the indicators are known within the quarter of interest, and this fact leads to interpret the BM estimates as forecasts. The forecasting horizon of BM is of one or at most two quarters ahead, and in the paper we will present alternative forecasting experiments with BM in order to assess the model performance in situations as close as possible to the actual forecasting activity.

We present BM for France, Germany, Italy and the euro area. They are estimated over the period from 1980 1st quarter to 1999 4th quarter and the forecast exercises are computed over the period 2000 1st quarter to 2002 2nd quarter characterised by a number of relevant events such as the unfolding of the monetary union, the recession in the USA, and the September 11 terroristic attack.

Two main empirical issues are dealt with in this paper: (a) the assessment of the forecasting performance of our BM against some benchmark models (ARIMA, VAR and a structural model, described in Section 3); and (b) the choice of the more appropriate aggregation level to be adopted in the forecasting procedure. In this particular case, in Section 4 we evaluate two alternatives to forecast euro area GDP: either by means of a single area-wide bridge equation or by a suitable aggregation of country-level GDP predictions.

The main results of the paper show that BM fare better than benchmark models. In addition, euro area GDP and its components are more precisely estimated by aggregating national forecasts. Assessment and comparison of the forecasting performance of benchmark and BM are reported in Section 5. Section 6 concludes.

2. Empirical issues

2.1. Bridge versus benchmark models and our forecast exercises

In this paper we compare the one- and two-step-ahead forecasting ability of BM against three types of benchmark models: univariate ARIMA, multivariate VAR, and structural models. These benchmark models cover a wide range of forecasting tools (from simple to complex structures, and different amounts of information).

[1] The BM methodology is close to the 'flash estimate' procedure advocated by Eurostat to obtain an early picture of the economy based on a less complete set of information than that used for traditional quarterly NA (see Eurostat, 2000, p. 374). However, BM are more general: differing from flash estimates in that they are not constrained by the quarterly NA methodology.

[2] The bridge equation for GDP is like the monthly US indicator model proposed by Trehan (1989, 1992).

Simple univariate ARIMA models are purely statistical models that do not exploit additional information either from economic theory or from short-term indicators. VAR models exploit information from all the variables of interest, and their choice influences the results (see Lutkepohl, 1982). Finally, in the Vector Equilibrium Correction Model (VEqCM) approach it is possible to also exploit information from economic theory in order to identify structural relationships among the variables of interest, and to assess the usefulness of past level disequilibria in short-term forecasting (see Clements & Hendry, 1999). Our small model for the euro area belongs to the VeqCM approach, and embodies long-run structural cointegration relationships.

It is well known (see for instance Stock, 2001) that minimisation of the forecast error entails a trade-off between simple and complex models: simple ARIMA models reduce parameter uncertainty, but have a limited information set, so that the error of approximation in the forecast of the variable may be large. However, the advantages of VAR models with a lot of parameters to be estimated are obtained at a cost of greater parameter uncertainty. Finally, VEqCM models impose parameter restrictions that reduce VAR parameter uncertainty at a cost of specification error risks.

In the BM approach, the information set is based on both past observations (as in ARIMA–VAR–VEqCM cases), and some indicators. These should be at least partially available over the forecast horizon and allow a reduction in forecast error with respect to the benchmark models. When indicator data over the forecast horizon are not available, they can be predicted with univariate techniques, with a substantial increase in the number of parameters to be estimated.

BM performance is assessed in the following five exercises (where T is the last observation of the estimation sample).

(1) The nowcast: a one-quarter-ahead forecast where the indicator observations of quarter $T+1$ are completely known.
(2) The pure one-step-ahead forecast: a one-quarter-ahead forecast where no observations of the indicators for quarter $T+1$ are known.
(3) The two-step-ahead nowcast: a two-quarter-ahead forecast where the indicator observations of both $T+1$ and $T+2$ quarters are completely known.
(4) The mixed two-step-ahead forecast: a two quarter-ahead forecast where we know the observations of the indicators for quarter $T+1$, but not those for $T+2$.
(5) The pure two-step-ahead forecast: a two-quarter-ahead forecast where we do not know any realisation of the indicators for either $T+1$ or $T+2$.

With respect to benchmark models, the nowcasts (cases 1 and 3) are the most favourable situations for BM. On the other hand, the pure one- and two-step-ahead forecasts (cases 2 and 5) represent the most unfavourable situation to BMs: indicator data for $T+1$ (and $T+2$) are not available and, in our exercises, they are forecast with univariate AR(5) models. Case (4) is intermediate: the BM has some informational advantages (the indicators over the $T+1$ quarter) over the benchmark approaches.

In forecasting exercises, the problem of revisions of the provisional data requires the choice between real-time and latest-available data (see Stark & Croushore, 2002, and the comments therein). Since at present a real-time NA data-set is far from being fully available for Europe, we follow the traditional practice and use the latest-available data, leaving a full assessment of the impact of different data vintages for future research.

2.2. The choice of the aggregation level

Euro-area GDP may be forecast by ex post aggregation of the country forecasts, or by directly modelling the aggregate area-wide variables[3], as well as by ex post aggregation of the forecast of each

[3] Fagan and Henry (1998), Dedola, Gaiotti and Silipo (2001) and Golinelli and Pastorello (2002) present results for euro-area money demand, which show the superiority of the aggregate approach. In contrast, Espasa, Albacete and Senra (2002) find evidence against the use of area-wide models, and prefer to forecast euro-area inflation at country level. In Bodo, Golinelli and Parigi (2000), the performance of area-wide models is superior to national (disaggregated) models in forecasting the euro-area index of industrial production, while Zizza (2002) and Marcellino, Stock, and Watson (2001) obtain better results with disaggregated models.

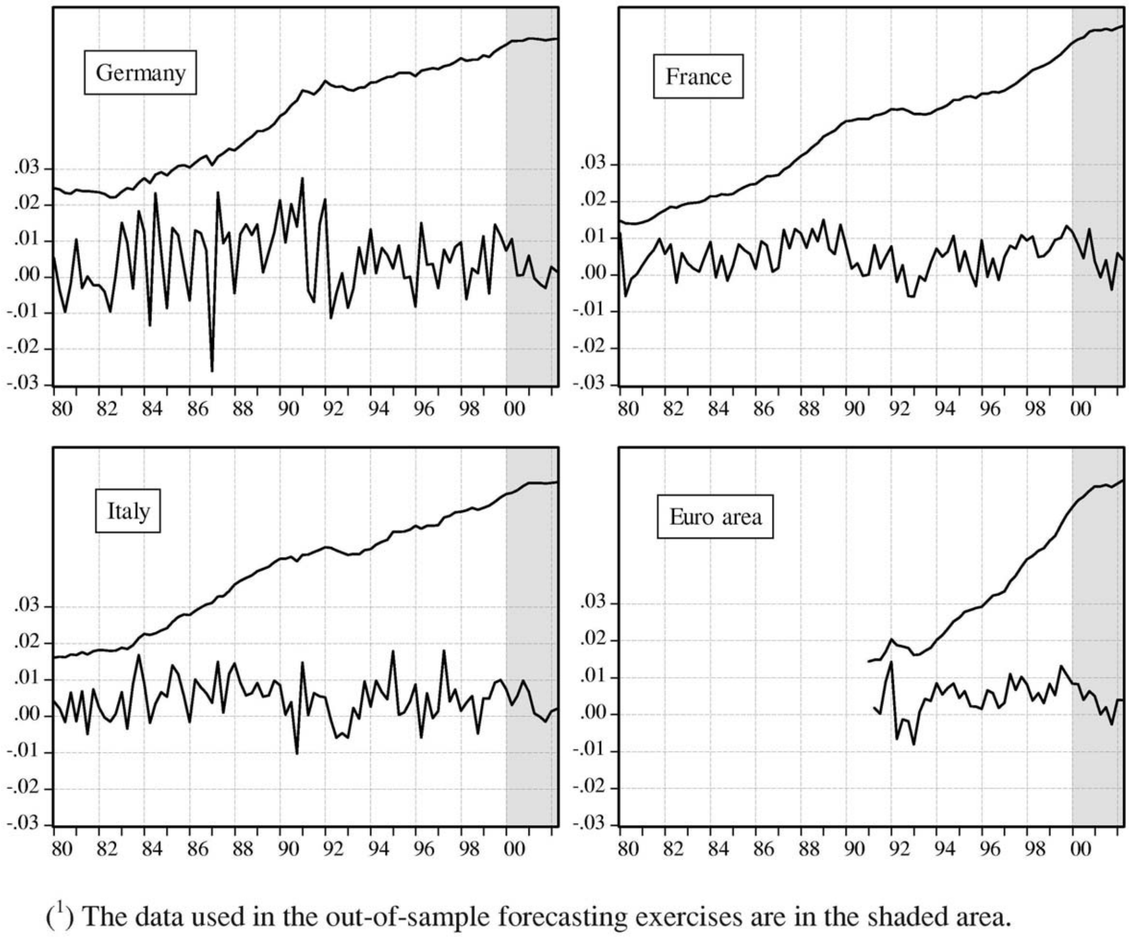

Fig. 1. GDP, log-levels and first differences (the data used in the out-of-sample forecasting exercises are in the shaded area).

component (demand-side models), or by modelling GDP directly (supply-side models)[4]. In the present paper, we compare the forecasting performance of supply and demand side benchmark and BM, estimated both area-wide and at country level.

Though the euro area includes 12 countries, we analyse single-country models only for France, Germany and Italy, since detailed and timely information is more easily available for these countries, which in any event account for almost three-quarters of total euro-area GDP. This implies that when using country-level equations, area-wide GDP cannot be forecast by a straightforward ex post aggregation. Indeed, the GDP forecasts by country are used as regressors in a second-step model aimed at predicting aggregate euro-area GDP (details on the 'aggregator' equation are in Section 4). It should be noticed that, in terms of forecasting performance, given the trade-off between 'more information' and 'less parameter uncertainty', the parsimony of this approach may offset the information loss arising from the exclusion of the remaining nine countries.

[4] For Italy, Parigi and Schlitzer (1995) show that both supply- and demand-side approaches perform satisfactorily, slightly better in the supply-side case.

3. The benchmark models for GDP

In this section the GDP benchmark models are derived. Section 3.1 reports unit root tests and Box–Jenkins univariate modelling of GDP, and Section 3.2 reports multivariate outcomes from a three-country GDP VAR model in differences; finally, Section 3.3 reviews a small model for the euro area, developed in the field of cointegrated VAR models.

3.1. Preliminary data analysis and the univariate models

Analysis of the levels and the first differences of GDP quarterly time series for France, Germany and Italy for 1980 1st quarter to 2002 2nd quarter (Fig. 1; see Appendix A for details on the statistical sources) suggests that non-stationarity is the main feature of the variables to be modelled. Fig. 1 also reports the euro-area GDP time series for 1991 1st quarter to 2002 2nd quarter (which is the largest officially available sample at area-wide level). Due to this data shortage, we have to be particularly careful in interpreting results of euro-area models. The benchmark models are estimated over the period from 1980 1st quarter to 1999 4th quarter, and the forecast comparison period 2000 1st quarter to 2002 2nd quarter is shaded in Fig. 1. The following analyses are based on the estimation period data only.

The average growth rate in Table 1 is similar among countries (about 2% on annual basis), while quarter-by-quarter variability is different (almost twice as great in Germany than in France or Italy). The path of first differences does not show relevant outliers, as suggested by plots in Fig. 1 and normality tests in Table 1. Fig. 1 also shows that, over the common 1991–1999 period, the variability of German GDP growth is still the highest (0.0072) though less than in the whole sample, while euro area, France and Italy volatility is about the same as for the whole period, ranging from 0.0046 to 0.0053.

As Newbold, Leybourne and Wohar (2001) argue, the distinction between trend-stationary and difference-stationary processes can be important for forecasting purposes (on this, see also Clements & Hendry, 2001). In addition, Christoffersen and Diebold (1998) and Diebold and Kilian (2000) show that unit root pretesting selects models with better forecasting accuracy. The results from both unit root and stationarity tests for single-country and area-wide GDP are clear-cut: over the sample period, log-GDP levels are $I(1)$, and must be put in differences to be stationary (test results are available on request from the authors). Then, the ARIMA models (1)–(3) are identified and estimated by following the Box–Jenkins specific-to-general approach (standard errors are in brackets below the parameter estimates, and e_t^{cc} are the residuals of the country cc model; cc=GE, FR, IT, EU).

$$\Delta\log(\mathrm{GDPGE})_t = \underset{(0.0012)}{0.0036} + \underset{(0.106)}{0.268}\,\Delta\log(\mathrm{GDPGE})_{t-4} + e_t^{\mathrm{GE}} \tag{1}$$

$$\Delta\log(\mathrm{GDPFR})_t = \underset{(0.0007)}{0.0033} + \underset{(0.108)}{0.362}\,\Delta\log(\mathrm{GDPFR})_{t-2} + e_t^{\mathrm{FR}} \tag{2}$$

$$\Delta\log(\mathrm{GDPIT})_t = \underset{(0.0006)}{0.0046} + e_t^{\mathrm{IT}} \tag{3}$$

The model for Italy is the simplest, since GDP levels are specified as a random walk with drift. The choice of

Table 1
GDP growth descriptive statistics[a] (1980 1st quarter–1999 4th quarter)

	Euro area[b]	Germany[c]	France	Italy
Mean	0.0045	0.0047	0.0050	0.0045
Standard deviation	0.0046	0.0094	0.0048	0.0056
Jarque–Bera test of normality, *P*-values	0.348	0.783	0.390	0.961

[a] First differences of log-levels.
[b] Euro area statistics refer to the period 1991 2nd quarter–1999 4th quarter.
[c] Pre-unification years refer to Western Germany only.

the best univariate model for the euro area is less clear; below, we report a simple ARIMA(1,1,0) model.

$$\Delta\log(\mathrm{GDPEU})_t = \underset{(0.0012)}{0.0031} + \underset{(0.173)}{0.365}\,\Delta\log(\mathrm{GDPEU})_{t-1} + e_t^{\mathrm{EU}} \tag{4}$$

Models (1)–(4) residual diagnostics do not show relevant specification problems (results are available on request).

3.2. The multicountry VAR model

The reduced-rank VAR models in levels are the multivariate extension of the integration-test-and-ARIMA-modelling approach of the previous section, though this may negatively affect the forecasting performance (see Clements & Hendry, 1999; Eitrheim, Husebo & Nymoen, 1999). In our case, given the absence of cointegration—as suggested by Johansen (1995) rank trace test—we start from a well behaved unrestricted VAR(4) (UVAR) in differences, simplify from general to specific, and obtain a model where: the German GDP equation is exactly the same as the univariate model (1); the French (2′) and Italian (3′) GDP equations present additional (and significant) explanatory variables. Full information maximum likelihood parameter estimates of the restricted VAR model are reported in Eqs. (1′), (2′) and (3′) (the main specification diagnostic are available on request):

$$\Delta\log(\mathrm{GDPGE})_t = \underset{(0.0011)}{0.0034} + \underset{(0.097)}{0.308}\,\Delta\log(\mathrm{GDPGE})_{t-4} + e_t^{\mathrm{GE}} \tag{1'}$$

$$\begin{aligned}\Delta\log(\mathrm{GDPFR})_t = {} & \underset{(0.0007)}{0.0031} + \underset{(0.088)}{0.409}\,\Delta\log(\mathrm{GDPFR})_{t-2} + \underset{(0.082)}{0.233}\,\Delta\log(\mathrm{GDPFR})_{t-3} \\ & - \underset{(0.091)}{0.214}\,\Delta\log(\mathrm{GDPFR})_{t-4} - \underset{(0.038)}{0.084}\,\Delta\log(\mathrm{GDPGE})_{t-2} \\ & + \underset{(0.058)}{0.233}\,\Delta\log(\mathrm{GDPIT})_{t-1} - \underset{(0.063)}{0.182}\,\Delta\log(\mathrm{GDPIT})_{t-3} + e_t^{\mathrm{FR}}\end{aligned} \tag{2'}$$

$$\Delta\log(\mathrm{GDPIT})_t = \underset{(0.0009)}{0.0054} + \underset{(0.057)}{0.138}\,\Delta\log(\mathrm{GDPGE})_{t-4} - \underset{(0.134)}{0.293}\,\Delta\log(\mathrm{GDPFR})_{t-4} + e_t^{\mathrm{IT}} \tag{3'}$$

The 27 zero-restrictions necessary to pass from UVAR to (1′)–(3′) are clearly accepted ($\chi^2_{27}=16.78$ with P-value$=0.938$). Even with the addition of significant parameters, the standard errors for (2′)–(3′) are only slightly smaller than for (2)–(3). In general, over the estimation period, the restricted VAR (1′)–(3′) has a statistical performance similar to that of univariate models (1)–(3).

3.3. The area-wide structural model

In this section we briefly sketch the main features of a small structural model for the euro area proposed by Bagliano, Golinelli and Morana (2002, 2003). The most appealing characteristics of the structural approach are: (i) it is explicitly based on economic theory; (ii) its empirical structure has been derived from the econometric evaluation of cointegration and parameter constancy. In particular, our model represents a way in which the P-star and the output-gap economic theories synthesise important issues such as dynamics, aggregation and equilibrium-correction (see Bagliano, Golinelli and Morana, 2003).

Though the model is made up of five equations, we report only the equation for GDP, the variable we want to benchmark. The list of explanatory variables includes the rate of capacity utilisation (qr), an indicator belonging to the information set of the BM described in Section 4.

$$\begin{aligned}\Delta\log(\mathrm{GDPEU})_t = {} & \underset{(1.15)}{4.11} - \underset{(0.048)}{0.199}(\Delta r_{t-4} - \Delta dp_{t-2}) \\ & + \underset{(0.081)}{0.223}\,\Delta qr_{t-1} + \underset{(0.047)}{0.137}\mathrm{qr}_{t-1}\end{aligned}$$

$$+\underset{(0.055)}{0.196}[m_{t-1} - \underset{(0.017)}{1.47}\log(\mathrm{GDPEU})_{t-1}] + e_t^{\mathrm{EU}} \tag{5}$$

where Δ is the first difference operator, r is the nominal long-term interest rate, dp is the HICP inflation rate, m is the logarithm of real M3 index, and e^{EU} represents the residuals. Parameters are estimated by Fiml over the 1980 1st quarter–1999 4th quarter period using partially reconstructed data (see Bagliano, Golinelli and Morana, 2002, 2003).

The model embodies a level relationship where the 'equilibrium' real money log-level has a long run elasticity to GDP equal to 1.47. Past imbalances between actual and target money (in squared brackets) exert a significant effect on short-term changes in GDP. This finding provides evidence of the usefulness of money in explaining GDP. In addition, short-term output changes are also explained by lagged changes in the real interest rate and in the rate of capacity utilisation. In the real rate definition, the expected inflation is proxied by a two-quarters lead of the realised inflation.

4. The bridge models

This section summarises the main estimation results for GDP single bridge equations (the supply-side model), and bridge equations for each GDP component (the demand-side model), both by country and area-wide. In the case of the GDP components, we explicitly model the following variables: total private consumption, public consumption, gross fixed capital formation, exports, imports, and variation in stocks. For Italy, a finer disaggregation is used in more detailed models, distinguishing between consumption of durable and non-durable goods and between investment in constructions and in other components.

All equation specifications are based on the general-to-specific methodology. The estimation period is 1980 1st quarter–1999 4th quarter, but sometimes data unavailability narrows the time span. Although the starting date can be different for each regression, end date is always the last quarter of 1999, thus leaving ten observations (up to the second quarter of 2002) for out-of-sample forecasting. The validity of model specifications has been checked through the usual battery of misspecification tests. In-sample stability of parameter estimates has been checked by estimating each regression over different sub-samples and by recursive least-squares. All variables have been transformed into logarithms (apart from the variation in stocks which has been considered as a percentage of total demand) in order to take account of different units of measurement.

Supply-side BM for Germany, France and Italy GDP are in Eqs. (6)–(8). The description and the sources of all the variables is in Appendix A.

$$\begin{aligned}\Delta\log(\mathrm{GDPGE})_t = {} & \underset{(0.196)}{0.658} \\ & - \underset{(0.051)}{0.167}\,\Delta_3\log(\mathrm{GDPGE})_{t-1} \\ & + \underset{(0.045)}{0.403}\,\Delta\log(\mathrm{IPGE})_t \\ & + \underset{(0.006)}{0.021}\,\Delta_3\log(\mathrm{RETGE})_{t-1} \\ & + \underset{(0.009)}{0.078}\,\Delta\log(\mathrm{ICTGE})_t \\ & + \underset{(0.010)}{0.045}\log(\mathrm{ICTGE})_{t-1} \\ & - \underset{(0.023)}{0.121}\log(\mathrm{GDPGE/IPGE})_{t-1} \\ & + \underset{(0.010)}{0.023}\,\Delta_2\,\Delta\log(\mathrm{CCIGE})_t \\ & + \underset{(0.010)}{0.023}\log(\mathrm{CARGED})_t \\ & - \underset{(0.010)}{0.011}\log(\mathrm{CARGED})_{t-1} + e_t^{\mathrm{GE}}\end{aligned} \tag{6}$$

$$\begin{aligned}\Delta\log(\mathrm{GDPFR})_t = {} & \underset{(0.127)}{1.083} - \underset{(0.101)}{0.458}\,\Delta\log(\mathrm{GDPFR})_{t-1} \\ & - \underset{(0.092)}{0.331}\,\Delta\log(\mathrm{GDPFR})_{t-2} \\ & + \underset{(0.018)}{0.153}\log(\mathrm{IPFR})_t \\ & - \underset{(0.016)}{0.142}\log(\mathrm{IPFR})_{t-1} \\ & + \underset{(0.010)}{0.031}\,\Delta_8(\mathrm{LEADFR})_{t-2} \\ & + \underset{(0.004)}{0.029}\,\Delta(\mathrm{EPROFR})_{t-1} \\ & + \underset{(0.019)}{0.069}\,\Delta\log(\mathrm{QCONFR})_t + e_t^{\mathrm{FR}}\end{aligned} \tag{7}$$

$$\begin{aligned}\Delta\, log(\mathrm{GDPIT})_t = {} & \underset{(0.877)}{2.968} - \underset{(0.103)}{0.459}\log(\mathrm{GDPIT})_{t-1} \\ & + \underset{(0.033)}{0.178}\log(\mathrm{IPIT})_{t-1} \\ & + \underset{(0.026)}{0.313}\,\Delta\log(\mathrm{IPIT})_t \\ & + \underset{(0.025)}{0.0844}\mathrm{TREND} \\ & + \underset{(0.054)}{0.109}\log(\mathrm{OCC})_{t-2} \\ & - \underset{(0.006)}{0.096}\mathrm{mav}(\mathrm{REALE},\ 6)_{t-1} \\ & + \underset{(0.011)}{0.035}\,\Delta\log(\mathrm{ESPIT/IMPIT})_{t-3} \\ & + \underset{(0.012)}{0.041}\,\Delta\,\Delta\log(\mathrm{CCIIT})_{t-4} \\ & + \underset{(0.029)}{0.056}\log(\mathrm{mav}(\mathrm{LEADIT},\ 6))_{t-1} \\ & + e_t^{\mathrm{IT}} \end{aligned} \tag{8}$$

where: mav(x, s) is the uncentered moving average of order s for the time series x, and Δ_k is the k-th difference operator. The supply-side BM by country are characterised by a strong link with the index of industrial production and a set of variables from the qualitative manufacturing surveys. The presence of other indicators is interpreted as a proxy for the lack of reliable information on the service sector.

The area-wide supply side equation is reported below[5].

$$\begin{aligned}\Delta\log(\mathrm{GDPEU})_t = {} & \underset{(1.047)}{5.498} \\ & - \underset{(0.119)}{0.290}\,\Delta\log(\mathrm{GDPEU})_{t-2} \\ & + \underset{(0.086)}{0.480}\log(\mathrm{GDPEU})_{t-1} \\ & + \underset{(0.039)}{0.256}\log(\mathrm{IPEU})_t \\ & + \underset{(0.003)}{0.097}\mathrm{TREND} \\ & - \underset{(0.020)}{0.051}\,\Delta\log(\mathrm{COMPE12})_{t-2} \\ & + \underset{(0.008)}{0.023}\,\Delta\log(\mathrm{BCIBE})_t + e_t^{\mathrm{EU}} \end{aligned} \tag{9}$$

A summary of misspecification tests of the regressions are available on request. Overall, the models appear to be fairly well specified. Though results are not perfectly comparable because the estimation periods are not the same, it is noticeable that the supply-side BM regression standard errors are always 40–50% lower than those of the benchmark models.

Given the large number of equations, the demand-side BM for France, Germany, Italy and the euro area are not reported, but simply summarised in words[6]. In the case of private consumption, the French (see Irac & Sedillot, 2002) and the German equations stress the importance of the retail sales index, while this is not the case for Italy because of reliability problems of the corresponding Italian indicator. A result common to all countries is the marginal role of the consumer confidence index, probably because other explanatory variables correlated with it are already included in the specification, such as the unemployment and the inflation rates and some proxy for the level of activity (see Carnazza & Parigi, 2001; Golinelli & Parigi, 2002). New car registrations confirm their validity for consumption of durable goods. Conditional on a proxy for the level of activity (such as the industrial production index for France and Germany or GDP for Italy), investment seems to be well tracked by survey variables, especially those related to the expected short-term evolution of demand (see Carnazza & Parigi, 2003). The results for Germany and Italy show that additional explanatory power may be achieved by considering the construction components expressly. The export and import equations are based on the corresponding trade variables, with some marginal role for the indexes of the real exchange rate, industrial production, and survey variables. In this context, however, the quality of the results is highly influenced by the well known statistical problems of customs data. Since variations in stocks are calculated in each country as NA identity residuals, there is no reliable indicator for this variable. Hence, a variety of specifications have been explored with differing results.

For almost all demand side equations of the euro area BM we have simple regressions of each area-

[5] For similar equations see also Grassman and Keereman (2001) and Runstler and Sedillot (2002).

[6] Detailed tables of the demand-side of the BM by country are available from the authors. The same is true for other statistical results not present in the paper but of interest to the reader.

wide series in differences on the intercept and the corresponding variables at country-level (in differences), the only exceptions are the investment and export equations where two indicators (respectively, the euro area business climate indicator and the real effective exchange rate) are included. These simple specifications in differences are the result of a search and are not imposed a priori: the corresponding variables of the three major countries in the euro area, along with the ordinary least-squares approximation, suffice to track the area-wide aggregate.

Country-specific GDP forecasts for Germany, France, and Italy are employed to predict euro-area GDP through the 'aggregator' (10).

$$\begin{aligned}\Delta \log(\mathrm{GDPEU})_t = {} & \underset{(0.00027)}{0.00087} \\ & + \underset{(0.045)}{0.392}\,\Delta \log(\mathrm{GDPFR})_t \\ & + \underset{(0.027)}{0.420}\,\Delta \log(\mathrm{GDPGE})_t \\ & + \underset{(0.037)}{0.179}\,\Delta \log(\mathrm{GDPIT})_t + e_t^{\mathrm{EU}}\end{aligned} \tag{10}$$

5. The GDP out-of-sample forecast comparison

In this section we compare the BM forecasting performance with those of the benchmark models presented in Section 3[7]. The comparison is based on one- and two-step-ahead forecasting exercises over the rolling fixed-length sample period January 2000–February 2002. The size of the window is equal to the number of observations used in the estimation phase. The advantages of the rolling approach are discussed by Tashman (2000), for example.

In Table 2 the root mean squared errors (RMSE) of the euro-area GDP growth rates of one-step-ahead forecasts are shown both for benchmark and BM. The results are robust to the use of alternative statistics such as the mean absolute error, the mean error, or Theil's inequality coefficient. The left-hand side of the table reports the RMSE of models using area-wide data: the ARIMA Eq. (4), the log-levels AR(5) model, the structural model (5) and the supply-side aggregate GDP Eq. (9). The right-hand-side results are based on single-country forecasts aggregated by Eq. (10). As explained in Section 2.1, the RMSE of the BM forecast are obtained in two different ways that are the lower and the upper bounds of the usual practice (best-case and worst-case scenario, respectively): complete information 'nowcast' (first column) and no information (second column) about the indicators over the $T+1$ quarter (missing $T+1$ indicator realisations are forecast by AR(5) models).

The main feature of Table 2 is the superior performance of the forecasts from the aggregation of national BM nowcasts, and specifically the supply-side BM nowcast. This result is consistent with the implications of Grunfeld and Griliches (1960) prescription to exploit as much disaggregate information as possible to model aggregate phenomena. All other performances of both benchmark and BM with unknown indicator data do not differ significantly, except the structural model that performs worse than the other models.

Since the one-quarter-ahead BM forecasting performance worsens when no indicator data is available, we can tentatively conclude that in one-step-ahead GDP forecasting there is no significant gain from BM when indicators are not updated promptly (see the results of two-step-ahead forecasting exercises in Table 4, however), but BM usefulness starts growing with the availability of some information on the quarter to be forecast (similar results are obtained by Trehan, 1992). The RMSE of the supply-side BM nowcast appears to be not only significantly lower than that of all benchmark models, but also to encompass all of them[8].

[7] Additional benchmark models were provided by AR(5) specifications in levels. In all forecasting exercises, these models play the role of automatic-univariate model, because they are not subject to specification search and are very easy to obtain. They are sufficiently flexible to represent the main dynamic features of the series, though the absence of unit root or MA parameter restrictions. The substance of reported forecast results about the AR(5) models does not change by using either AR(4) in differences or random walk models.

[8] The results of the RMSE equality tests and of the RMSE encompassing tests—both sets of tests are computed according to the Diebold and Mariano (1995) procedure modified by Harvey, Leybourne and Newbold (1997, 1998)—are available upon request.

Table 2
RMSE of the euro area one-step-ahead GDP forecasts (2000 1st quarter–2002 2nd quarter)

Area-wide models			Aggregation of national-models		
ARIMA Eq. (4)	0.34		ARIMA Eqs. (1)–(3)	0.39	
AR(5) model	0.40[a]		AR(5) models	0.45[a]	
Structural Eq. (5)	0.68		VAR Eqs. (1′), (2′) and (3′)	0.37	
	b	c		b	c
Supply-side Eq. (9)	0.28	0.37	Supply side BM (6)–(8)	0.15	0.36
			Demand side BM	0.24	0.45
			Supply–demand average	0.18	0.38

[a] Average of supply and demand forecasts.
[b] Nowcast exercise, where $T+1$ quarter indicator data are known.
[c] Pure forecast exercise, where $T+1$ quarter indicator data are not known. T is the last observation of the estimation sample.

The single-country contribution to the euro area forecasting performance described above is reported in Table 3, where we list the one-step-ahead RMSE of national GDP forecasts for benchmark and BM. Again, as in the case of the euro area, the superiority of the supply-side BM nowcasts emerges clearly, while all other performances are quite similar. The poor performance of demand-side specifications depends on both the quality of the indicators and the degree of misspecification for the single component equations. This feature is reinforced when the BM forecasts are computed without the help of additional information about the quarter T+1 indicators.

In order to further analyse the effects of timely availability of updated indicators on the forecasting performance, in Table 4 we report the RMSE of a two-step-ahead forecasting exercise both by country and by the corresponding area-wide aggregation. Along the rows there are the usual three approaches: supply side, demand side, and the average. To simplify the presentation, we only report the RMSE of the benchmark AR(5) model, which is not too different from that of the other benchmark models.

The results reported in Table 4 somewhat contrast with the one-step-ahead exercise. Even when no information on the indicators over the forecasting horizon is available, the RMSE of BM are in many cases lower than (though not significantly different from) those of the benchmark models. The performance of the BM improves as more pieces of information on the indicators become available: when indicators are known only for the T+1 quarter, the RMSE is well below the benchmark model (the German demand-side BM is the only exception), and it falls by about 50% in the two-step-ahead 'nowcast' exercise (fourth–fifth columns in Table 4)[9].

Often, the combination of supply and demand predictions delivers an RMSE lower than in all other cases. The better performance of the combination (the so called 'average' puzzle) may be related to the presence of some form of misspecification in both the demand and the supply-side models (see Hendry & Clements, 2002).

6. Conclusions

The policymaker needs timely information on the present state of the economy (essentially the short-term GDP path), but the release of reliable NA data requires a good deal of time. Can one devise econometric tools that provide timely and accurate data? That is the question addressed in this paper. More specifically, we assess the forecasting performance of a number of alternative models for the short-term prediction of euro-area GDP. The performance of traditional benchmark models (uni and multivariate) is compared with that of BM. The BM approach is an efficient way to exploit the timely but heterogeneous information contained in short-term indicators in the context of autoregressive distributed lag models. The results are clear-cut: BM performance is always better than benchmark models, provided that at least

[9] Indicators have been forecast by AR(5) models without any sort of specification search. Results may improve by exploiting more accurate forecasting models as advocated by Trehan (1992) and Runstler and Sedillot (2002).

Table 3
RMSE of single-country one-step-ahead GDP forecasts (2000 1st quarter–2002 2nd quarter)

	France		Germany		Italy	
ARIMA Eqs. (1)–(3)	0.46		0.57		0.38	
AR(5) equations[a]	0.47		0.66		0.47	
VAR Eqs. (1′), (2′) and (3′)	0.43		0.57		0.38	
	[b]	[c]	[b]	[c]	[b]	[c]
Supply side BM (6), (7), (8)	0.19	0.44	0.37	0.53	0.21	0.32
Demand side BM	0.57	0.63	0.61	0.68	0.36	0.41
Supply–demand average	0.35	0.46	0.43	0.56	0.25	0.34

[a] Average of supply and demand forecasts.
[b] Nowcast exercise, where $T+1$ quarter indicator data are known.
[c] Pure forecast exercise, where $T+1$ quarter indicator data are not known. T is the last observation of the estimation sample.

some indicators are available over the forecasting horizon.

As far as the choice of the level of model aggregation is concerned, our results are quite clear-cut: over a forecasting horizon one- to two-step-ahead, the aggregation of forecasts by country performs better in forecasting euro-area GDP and also offers information on the state of the single economies. On the other hand, the aggregation of forecasts by NA components (the demand-side approach) performs worse than modelling aggregate GDP data (the supply-side approach).

As Granger and Yoon (2001) point out, "In theory at least, more information in general leads to improved forecasts. However, due to difficulties associated with *model specification and estimation* among others, no general consensus is reached among various empirical results" (p. 18, emphasis added). This observation highlights the need to take the nature of the variables forecast explicitly into account. On the one hand, when the variable is 'easy to model' (as in the present case, where the usefulness of the economic indicators in explaining short-term GDP is evident), it is more efficient to exploit

Table 4
RMSE of two steps-ahead GDP forecasts (2000 1st quarter–2002 2nd quarter)

	Benchmark[a]	BM (pure forecast)[b]	BM (mixed forecast)[c]	BM (nowcast)[d]
Supply side models				
France	0.86	0.70	0.45	0.22
Germany	0.85	0.94	0.61	0.26
Italy	0.74	0.53	0.37	0.22
Euro area	0.63	0.63	0.43	0.19
Demand side models				
France	0.92	0.65	0.63	0.52
Germany	0.74	0.87	0.87	0.62
Italy	0.72	0.58	0.56	0.53
Euro area	0.62	0.60	0.50	0.27
Supply–demand average				
France	0.87	0.57	0.43	0.32
Germany	0.84	0.83	0.68	0.33
Italy	0.68	0.52	0.42	0.34
Euro area	0.61	0.57	0.46	0.21

[a] AR(5) models.
[b] Pure forecast, where neither $T+1$ nor $T+2$ quarter indicator data are available.
[c] Mixed forecast exercise, where only $T+1$ quarter indicator data are available.
[d] Nowcast exercise, where the indicator data are available over the whole forecasting horizon. T is the last observation of the estimation sample.

the advantages of disaggregate models (given data availability). In this context, a role is also played by the quarterly frequency of data, which may imply fairly low statistical noise. On the other hand, when the variables of interest are 'traditionally difficult to model' (like money and industrial production), aggregate modelling can often alleviate model specification problems (e.g. through statistical averaging effects).

Acknowledgements

We are grateful to Guja Bacchilega, Andrea Brandolini, Fabio Busetti, Michael Clements, Alfredo Coutiño, Catherine Doz, Clive Granger, Suleyman Ozmucur, Francesco Paternò, Federico Signorini, two anonymous referees, and an IJF associate editor for comments and suggestions. The usual caveats apply. Financial support from MIUR is gratefully acknowledged (R. Golinelli). The views contained here are those of the authors only and do not necessarily reflect those of the institutions for which they work.

Appendix A. Data

All NA data are in real terms, at 1995 prices; they are seasonally adjusted and, apart from Italy, working day adjusted. The frequency of all the series is quarterly, i.e. original monthly (or higher frequency) data are suitably transformed into quarterly data. The variable TREND is a linear trend. IMF and OECD denote the corresponding international organisations. BI stands for Bank of Italy.

Euro area	
BCIBE	Belgian business climate indicator. Source: European Commission
COMPE12	Real effective exchange rate, based upon production prices. Source: BI computations on national and IMF data
GDPEU	Gross domestic product. Source: Eurostat, NA
IPEU	Industrial production index (seasonally adjusted). Source: Eurostat
France	
EPROFR	Production expectations (balance). Source: INSEE
GDPFR	Gross domestic product. Source: Insee, NA
IPFR	Industrial production index (seasonally adjusted). Source: INSEE
LEADFR	Oecd leading indicator for industrial production. Source OECD, Main Economic Indicators
QCONFR	Household consumption of manufactured goods. Source: INSEE, NA
Germany	
CARGED	Car registrations. Source: OECD
CCIGE	Consumer confidence index. Source: European Commission
GDPGE	Gross domestic product. Source: Federal Statistical Institute, NA
ICTGE	Construction, production index. Source: Federal Statistical Institute
IPGE	Industrial production index (seasonally adjusted). Source: Federal Statistical Institute
RETGE	Retail sales index. Source: Federal Statistical Institute
Italy	
CCIIT	Consumer Confidence index. Source: ISAE
ESPIT	Exports of goods and services. Source: Istat, NA
GDPIT	Gross domestic product. Source: Istat, NA
IMPIT	Imports of goods and services. Source: Istat, NA
IPIT	Industrial production index (seasonally adjusted). Source: Istat
LEADIT	Leading indicator for the Italian economy. Source: Altissimo, Marchetti and Oneto (2000)
OCC	Total employment (standard unit of labour). Source: Istat, NA
REALE	TAIMPLQ less households' expected inflation rate. Source: BI computations on ISAE data, see also Parigi (1993)
TAIMPLQ	Nominal interest rate on loans (average). Source: BI

References

Altissimo, F., Marchetti, D. J., & Oneto, G. P. (2000). The Italian business cycle: coincident and leading indicators and some stylised facts. *Giornale degli Economisti e Annali di Economia*, *60*(2), 147–220.

Bagliano, F. C., Golinelli, R., & Morana, C. (2002). Core inflation in the euro area. *Applied Economics Letters*, *9*, 353–357.

Bagliano, F. C., Golinelli, R. & Morana, C. (2003). Inflation modelling in the euro area. In: R. Beetsma, C. Favero, A. Missale, V.A. Muscatelli, P. Natale & P. Tirelli, *Monetary Policies, fiscal policies and labour markets. Macroeconomic policies making in the EMU*. Cambridge University Press, forthcoming.

Bodo, G., Golinelli, R., & Parigi, G. (2000). Forecasting industrial production in the euro area. *Empirical Economics*, *25*, 541–561.

Carnazza, P, & Parigi, G. (2001). The evolution of confidence for European Consumers and businesses in France, Germany and Italy, Temi di Discussione. *Banca d'Italia*, 406.

Carnazza, P., & Parigi, G. (2003). Tentative business climate indicators for the Italian economy. *Journal of Forecasting*, forthcoming.

Clements, M. P., & Hendry, D. F. (1999). *Forecasting non-stationary economic time series*. Cambridge: MIT Press.

Clements, M. P., & Hendry, D. F. (2001). Forecasting with difference-stationary and trend-stationary models. *Econometrics Journal*, *4*, S1–S19.

Christoffersen, P. F., & Diebold, F. X. (1998). Cointegration and long-horizon forecasting. *Journal of Business and Economic Statistics*, *16*, 450–458.

Dedola, L., Gaiotti, E., & Silipo, L. (2001). Money demand in the euro area: do national differences matter? Temi di Discussione. *Banca d'Italia*, 405.

Diebold, F. X., & Mariano, R. S. (1995). Comparing predictive accuracy. *Journal of Business and Economic Statistics*, *13*, 253–263.

Diebold, F. X., & Kilian, L. (2000). Unit root tests are useful for selecting forecasting models. *Journal of Business and Economic Statistics*, *18*(3), 265–273.

Eitrheim, O., Husebo, T. A., & Nymoen, R. (1999). Equilibrium-correction vs. differencing in macroeconometric forecasting. *Economic Modelling*, *16*, 515–544.

Espasa, A., Albacete, R., & Senra, E. (2002). Forecasting EMU inflation: a disaggregated approach by countries and components. *European Journal of Finance*, forthcoming.

Eurostat (2000). *Handbook on quarterly national accounts*, Brussels.

Fagan, G., & Henry, J. (1998). Long run money demand in the EU: evidence for area-wide aggregates. *Empirical Economics*, *23*, 483–506.

Forni, M., Hallin, M., Lippi, M., & Reichlin, L. (2000). The generalized factor model: identification and estimation. *The Review of Economics and Statistics*, *82*(4), 540–554.

Golinelli, R., & Parigi, G. (2002). What is this thing called confidence? *A comparative analysis of consumer confidence indexes in Australia, Europe and the USA*, mimeo.

Golinelli, R., & Pastorello, S. (2002). Modelling the demand for M3 in the euro area. *European Journal of Finance*, forthcoming.

Granger, C. W. J., & Yoon, G. (2001). Self-generating variables in a cointegrated VAR framework, *UCSD Discussion Paper*, 4, February.

Grassman, P., & Keereman, F. (2001). An indicator based short-term forecast for quarterly GDP in the euro area, *European Commission Economic Paper*, 154.

Grunfeld, Y., & Griliches, Z. (1960). Is aggregation necessarily bad. *The Review of Economics and Statistics*, *42*(1), 1–13.

Harvey, D. I., Leybourne, S. J., & Newbold, P. (1997). Testing the equality of prediction mean squared errors. *International Journal of Forecasting*, *13*, 281–291.

Harvey, D. I., Leybourne, S. J., & Newbold, P. (1998). Tests for forecast encompassing. *Journal of Business and Economic Statistics*, *16*, 254–259.

Hendry, D. F., & Clements, M. P. (2002). Pooling of forecasts. *Econometrics Journal*, *5*, 1–26.

Irac, D., & Sedillot, F. (2002). Short run assessment of French economic activity using OPTIM. *Notes D'Etudes et de Recherche*, Banque de France, 88.

Johansen, S. (1995). *Likelihood-based inference in cointegrated vector autoregressive models*. Oxford University Press.

Lutkepohl, H. (1982). Non causality due to omitted variables. *Journal of Econometrics*, *19*, 367–378.

Marcellino, M., Stock, J. H., & Watson, M. W. (2001). Macroeconomic forecasting in the euro area: country specific versus area wide information. *European Economic Review*, forthcoming.

Newbold, P., Leybourne, S., & Wohar, M. E. (2001). Trend-stationarity, difference-stationarity or neither: further diagnostic tests with an application to US real GNP, 1875–1993. *Journal of Economics and Business*, *53*, 85–102.

Parigi, G. (1993). La determinazione del tasso d'interesse reale *ex-ante*, Banca d'Italia, mimeo, April.

Parigi, G., & Schlitzer, G. (1995). Quarterly forecasts of the Italian business cycle by means of monthly economic indicators. *Journal of Forecasting*, *14*, 117–141.

Runstler, G., & Sedillot, F. (2002). Short-term estimates of euro area real GDP by means of monthly data, mimeo, December.

Stark, T., & Croushore, D. (2002). Forecasting with a real-time data set for macroeconomists. *Journal of Macroeconomics*, *24*(4), 507–568.

Stock, J. H. (2001). Forecasting economic time series. In: Baltagi, B. H. (Ed.), *A companion to theoretical econometrics*. Blackwell, pp. 562–584.

Stock, J. H., & Watson, M. W. (1999). Diffusion indexes. *NBER Working Paper*, no. 6702.

Tashman, L. J. (2000). Out-of-sample tests of forecasting accuracy: an analysis and review. *International Journal of Forecasting*, *16*, 437–450.

Trehan, B. (1989). Forecasting growth in current quarter real GNP, Federal reserve Bank of S. Francisco. *Economic Review* (winter), 39–52.

Trehan, B. (1992). Predicting contemporaneous output, Federal reserve Bank of S. Francisco. *Economic Review* (2), 3–11.

Zarnowitz, V. (1992). *Business cycles*. University of Chicago Press.

Zizza, R. (2002). Forecasting the industrial production index of the euro area through forecasts of the three main countries. *Temi di Discussione*, Banca d'Italia, no. 441.

Biographies: Alberto BAFFIGI is an economist at the Research Department of the Bank of Italy, where he deals with short-term analysis for the euro area, and fellow of the Euro Area Business Cycle Network (EABCN). He did his graduate studies at Bologna University. His research interests include the study of business cycle, short-term forecasting, applied conjunctural analysis, and history of statistical methods.

Roberto GOLINELLI is Associate Professor in Applied Econometrics of the University of Bologna (Italy), and fellow of the Euro Area Business Cycle Network (EABCN). He did his main working experiences as head of research projects at the Associazione Prometeia (macroeconometric forecasts), and at the European Commission (HERMES model project, indirect taxes harmonisation, the cost of non-Europe). His research interests include time series forecasting, transition and accession countries analysis, poolability and heterogeneity.

Giuseppe PARIGI is deputy director in the Research Department of the Bank of Italy, and fellow of the Euro Area Business Cycle Network (EABCN). He did his post graduate studies at the Universities of Oxford and Warwick (UK) and has worked extensively at the Bank of Italy's macroeconometric model of the Italian economy. His research interests include macroeconomics, econometrics and time series methods and forecasting.