BANCA D'ITALIA

Temi di discussione

del Servizio Studi

Forecasting the industrial production index for the euro area through forecasts for the main countries

by Roberta Zizza

Number 441 - March 2002

FORECASTING THE INDUSTRIAL PRODUCTION INDEX FOR THE EURO AREA THROUGH FORECASTS FOR THE MAIN COUNTRIES

by Roberta Zizza*

Abstract

The aim of the present work is to obtain short-term predictions of the monthly volume of the industrial production of the euro area. Preliminary information on the behaviour of this variable is needed, since the index is released with a lag of about two months. A model based on the US industrial production index and on the single-country forecasts of the production indices for the main euro-area countries is proposed.

JEL classification: C22, C53.

Keywords: prediction, industrial production, forecast combination, encompassing.

Contents

* Bank of Italy , Economic Research Department.

1. Introduction[1]

The aim of the present work is to obtain short-term predictions of the monthly volume of the industrial production of the euro area, which is commonly adopted as a business cycle indicator. Preliminary information on the evolution of this variable is needed since the index[2] is released by Eurostat with a lag of about two months with respect to the reference period.

Several approaches have been followed in the literature: among these, autoregressive models and models based on "external" variables whose relationship with the level of activity has proven to be strong. In this paper I propose two models based on the forecasting of the production indices of the main euro-area countries (Germany, France and Italy): in the first, each country's forecast will be separately included in the model; in the second, the forecasts for the different countries will be combined with appropriate weights. I then extend the models by introducing the forecast for the other member countries considered as a whole. Further, I seek to exploit all the information available by constructing a model in which the euro-area index is regressed against the US industrial production index, the Business Confidence Index (BCI) from the European Commission harmonized survey on manufacturing firms, and the single-country indices.

To evaluate the forecasting performance of these models, my benchmark is the model proposed by Bodo, Golinelli and Parigi (2000), based on the US index and on the BCI.

Finally, I effect a suitable combination of the forecasts obtained from some of the models; this operation – as documented by several studies (see, for example, Clemen (1989) and the references listed therein) – is expected to enhance the forecasting power with respect to the predictions taken individually.

[1] I wish to thank Giuseppe Parigi for his constant guidance and advice, and Alberto Baffigi, Fabio Busetti, Massimo Caruso, Riccardo Cristadoro, Roberto Golinelli and L. Federico Signorini for useful comments. E-mail: zizza.roberta@insedia.interbusiness.it.

[2] According to Council Regulation No. 1165/98 of 19 May 1998 concerning short-term statistics, Member States are required to transmit to Eurostat a working-day adjusted index; the index for the area is then calculated by aggregating the national indices using appropriate weights. The weight for country j is given by the ratio between the value added at factor cost of industry excluding construction of j and the sum of the value added of industry excluding construction of all countries in the area.

2. Single-country models

The ultimate objective of this study is to obtain predictions of the industrial production index for the euro area. One way of doing this is to estimate an area-wide model directly for the aggregate euro-area index, computed as the weighted average of the national indices adjusted for the number of trading days.

Since the member countries still present some national specificities, there is a rationale for a multi-country (disaggregated) approach.[3] Consistently with the method used to compute the euro-area index, one should construct a model for each member country, use it to forecast the production index one-step ahead and obtain the euro-area forecast as a linear combination of the single-country forecasts by applying the weights provided by Eurostat. However, official production indices for some countries are released with a considerable lag and, in several cases, even after the index for the euro area.[4]

Accordingly, at this stage I attempt to estimate separate models only for the main countries, i.e. France, Germany and Italy, which together account for 75 per cent of total euro-area output. An important by-product will be the forecasts of the industrial production index for each of these countries.

Predicting the level of activity in a given country – or aggregation of countries such as the euro area – in the short run with structural models is often difficult, since most of the explanatory series are not released quickly, sometimes even after the industrial production index itself.

Only a few variables are actually available before the industrial production index and can thus be used as regressors; among these, the data derived from qualitative surveys on firms and households, such as the harmonized survey carried out by the European Commission. Apart from their timeliness, the ability of these indicators to explain, and in some cases, anticipate the behaviour of manufacturing has been widely documented (see, for

[3] Several studies (for example, Grunfeld and Griliches (1960)) report the poor quality of disaggregated statistics as a justification for using models for aggregated variables. But since the index for the euro area is nothing but the weighted average of the national indices, this line of reasoning seems not to apply here.

[4] Some member countries provide preliminary estimates by the release date for the euro-area index; missing values for EMU and EU aggregates are estimated by Eurostat.

example, Altissimo *et al.* (2001), Baffigi and Bassanetti (2001), Carnazza and Parigi (2001) and references cited therein).

Another promptly available variable whose role as an indicator of the industrial production has been emphasized (see, for example, Terasvirta (1984) and Bodo and Signorini (1987)) is electricity consumption. This variable has been successfully applied to the Italian case, as Marchetti and Parigi (2000) point out. Other short-term indicators, such as production indices from other countries, competitiveness indicators, and the like, will be considered.

This study deals with regression models for single countries (or aggregations of them); regarding the specification of the models, I adhere to the general-to-specific modelling approach originally associated with the London School of Economics. In the words of Hoover and Perez (1999, p. 168), "A sufficiently complicated model can, in principle, describe the salient features of the economic world. Any more parsimonious model is an improvement [...] if it conveys all of the same information in a simpler, more compact form […]. The art of model specification in the LSE framework is to seek out models that are valid parsimonious restrictions of the completely general model, and that are not redundant in the sense of having even more parsimonious models nested within them that are also valid restrictions of the completely general model". In this context, the current and lagged values of all the variables, exogenous and endogenous, are considered in the initial specification, with seasonal dummies and a time trend.

The production of the industrial sector excluding construction is expressed with an index (base 1995=100); the series have been transformed into logarithms for France and Germany.[5]

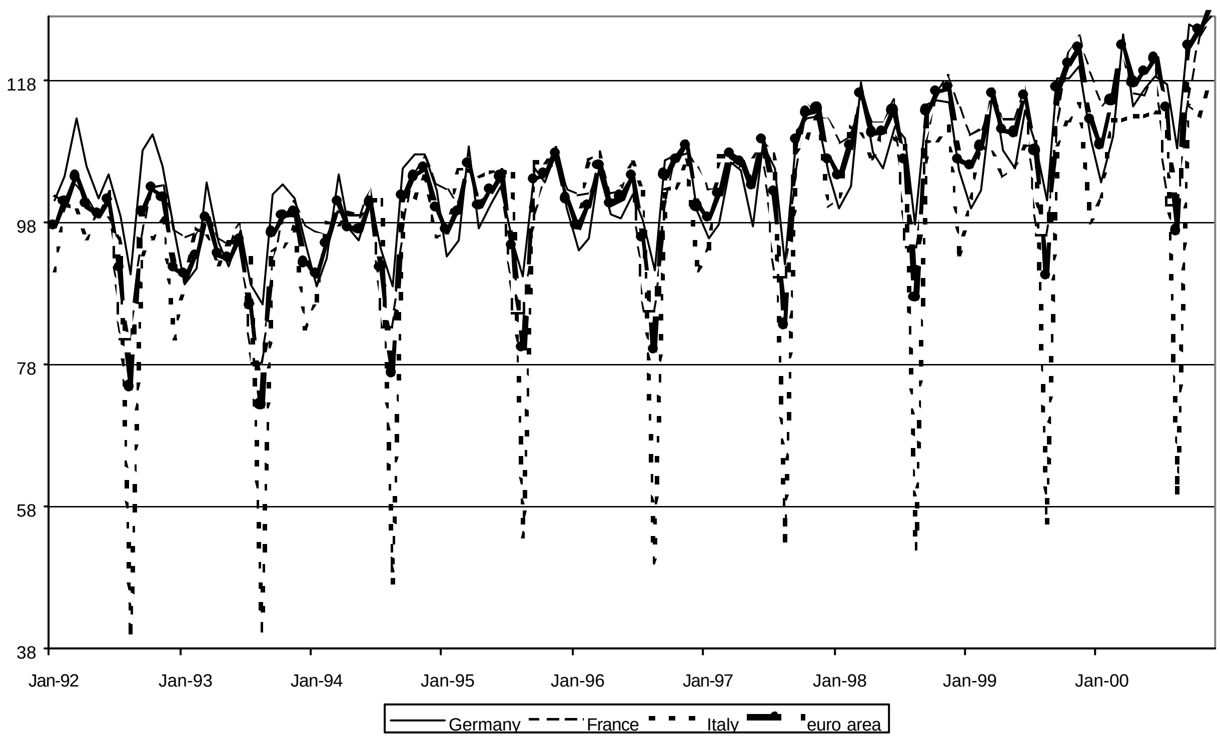
Dummies (called *seasonals*) are included to adjust the series for seasonal effects (the monthly series are adjusted for the number of working days but not seasonally adjusted). From Figure 1 it is easy to observe the especially strong seasonality of the Italian data in

[5] The choice between linear and log-linear models was made using the test implemented in the TRAMO-SEATS package. The evidence was in favour of a logarithmic transformation for all variables except the Italian index.

correspondence with August holidays. Similar periodical behaviour is actually evident for all series. The coefficients attached to seasonal dummies are in general highly significant for all months for the main economies. Seasonality for the euro area is less pronounced, possibly reflecting the mixture of countries with different habits.

Figure 1

INDUSTRIAL PRODUCTION INDICES



Sources: ISTAT (Italy); INSEE (France); Bundesbank (Germany); Eurostat (euro area). The data are working-day adjusted. The base year is always 1995 = 100. Range: 1992.01 – 2000.11.

2.1 Italy

For Italy I borrow the model estimated by Marchetti and Parigi (2000) and adopted by the Bank of Italy in order to forecast the industrial production index. The authors introduced the total electricity consumption[6] as a regressor (*elecons*), using data on weather temperature (in degrees centigrade, called *weather*) to account for domestic uses of electricity. They identified the following model, where y_t represents the level of the index:

$$y_t = a + b\sum_{i=1}^{2}\frac{y_{t-i}}{2} + g\ \log(elecons_t) + l\cdot trend + m\cdot trend^3 + n\cdot weather + x\cdot weather^2 + seasonals + u_t$$

2.2 France

As regards France, the appeal of the model proposed lies in the joint use of two pieces of information: electricity consumption (source: *Eurostat*) and replies to qualitative questions concerning the current and expected state of different aspects of the economy. Several variables were tried as regressors; among them, the order-book level, stocks of finished products, the industrial and consumer confidence index and the competitiveness indicator proved to have no role in explaining the evolution of the level of activity. In the end, expectations for the months ahead for production (*prodex*) and selling price (*pricex*) were chosen; these components were combined to form a sort of confidence index (*euindex* hereafter), which is different from the BCI[7] released by the European Commission. Figure 2a shows that the cyclical pattern of *euindex* tends to lead that of BCI.[8]

The log-level of the industrial production index (y_t) is linearly related to its own lagged values, to current electricity consumption transformed into logarithms[9] and to the lagged values of the European Commission indicators (indices, 1995 = 100).

[6] Electricity consumption data, expressed in Megawatt hours (MWh), are provided by ENEL, Italy's largest power company, immediately after the end of the reference month.

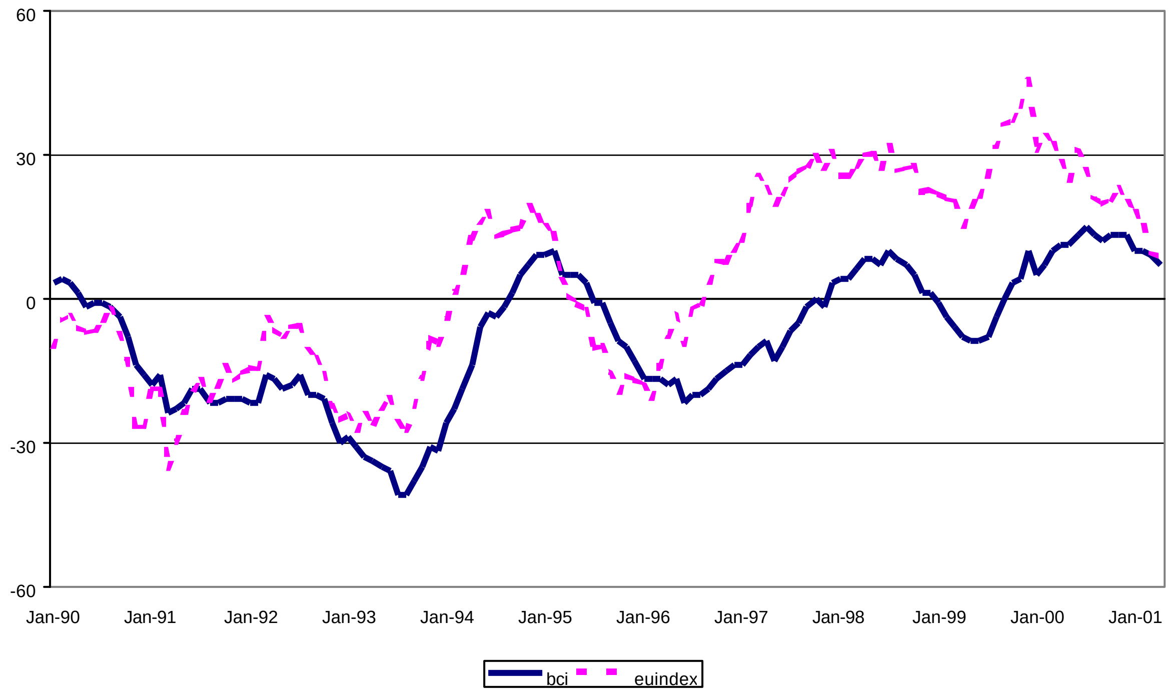
[7] The Business Confidence Index is calculated by the European Commission as an arithmetic average of the replies to the questions regarding expectations of production, level of demand and stocks of finished goods (with inverted sign). The leading behaviour of *euindex* with respect to BCI could follow from the fact that two expectation components – prices and production – are included.

[8] The correlation between *euindex* and BCI is equal to 0.75; the correlation between *euindex* and BCI one lag ahead rises to 0.80.

[9] As in the Italian case, the information about industrial consumption for France is not currently available. Marchetti and Parigi, as mentioned above, considered data on climate to adjust the series for non-manufacturing uses of electricity (e.g. households). This kind of information is difficult to collect for France, but the performance of the model does not seem to suffer from this drawback.

Figure 2a

FRANCE: BUSINESS CONFIDENCE INDEX VERSUS EUINDEX



Source: Based on European Commission data.

The final formulation is the following:

$$y_t = a + b y_{t-1} + g y_{t-2} + d \log(elecons_t) + l\, euindex_t + seasonals + u_t$$

where $euindex_t = \left(prodex_{t-1} - \sum_{i=1}^{12} \frac{pricex_{t-i-1}}{12} \right)$

2.3 Germany

For Germany, it was necessary to deal with the consequences of the reunification in 1991, which plausibly caused a structural change in the series during the subsequent months.[10] First, I tried to describe its behaviour by using a smooth transition regression

[10] The Chow test, performed by considering as a changepoint several months during mid-1992, always provided evidence for the hypothesis of structural break (the F-statistic was never below 2.7, corresponding to p-values under the standard confidence level of 5 per cent).

model *à la* Terasvirta (see, for example, Granger and Terasvirta (1993)), taking time as a transition variable. However, results obtained from the estimation were not encouraging and suggested shifting the estimation interval upwards starting from 1993.

In contrast with France and Italy, electricity consumption actually did not play an important role in the specification of the model, possibly owing to the lesser dependence of German industry on electricity as a source of energy.[11] Another difference with respect to French and Italian models is the lack of autoregressive terms. In choosing the regressors I kept in mind Germany's well-known export orientation, and hence included a competitiveness indicator (*comp*) as well as the US industrial production index (*ipus*) on the right-hand side of the model. Several qualitative variables extracted from the European Commission survey were also tried; the most significant one was the assessment of order-book levels (*ord*), commonly taken to be a leading variable. Moreover, *euindex* was introduced as an explanatory variable.[12]

The definitive form of the model for the log-level of the index can be written as

$$y_t = a + b \sum_{i=1}^{12} \frac{ord_{t-i-1}}{12} + g \Delta \log\left(ipus_t\right) + d\,euindex_t + \\ + l \sum_{i=1}^{12} \frac{\log\left(comp_{t-i}\right)}{12} + m \cdot trend^3 + seasonals + u_t .$$

The main findings of the parameter estimation using the ordinary least squares procedure and the results of the set of econometric tests performed on the single-country models are reported in Tables 1a-1c.[13]

---

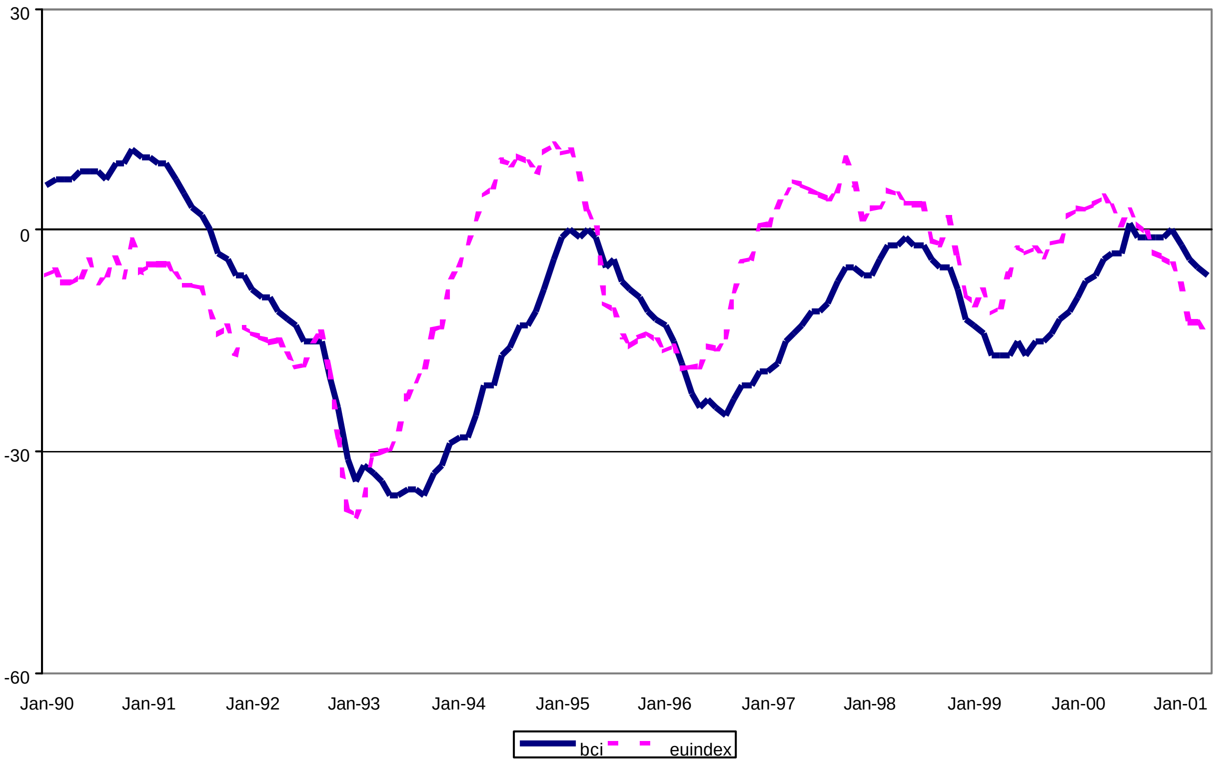
[11]Among the different sources of energy used by the industrial sector in Germany, the share of electricity consumption is about 22 per cent (reference year: 2000; source: *International Energy Agency*); the corresponding figure for France is 54 per cent (reference year: 1999; source: *Observatoire de l'Energie - Ministére de l'économie, des finances et de l'industrie*); for Italy, it is approximately 30 per cent (reference year: 1999; source: *International Energy Agency*).

[12] As in the case of France, the behaviour of *euindex* seems to lead that of BCI (Figure 2b), although the correlation coefficients are lower (0.47 between *euindex* and BCI, 0.54 between *euindex* and BCI one lag ahead).

[13] The estimation range, 1993.01 – 1999.12, was chosen in order to leave a sufficient sample to verify the forecasting ability of the different models. Coefficient estimates of the seasonal dummies, the time trends and the constant are omitted; m.a. = moving average. Results for Italy may differ from those reported in Marchetti and Parigi (2000) owing to the different range used for the estimation. I also tried to estimate simple ARIMA models for the three countries, but the models presented here always behaved better.

Figure 2b

GERMANY: BUSINESS CONFIDENCE INDEX VERSUS EUINDEX



Source: Based on European Commission data.

The three models fit the data quite well: the adjusted R^2 never goes below 0.963 (Germany), the standard errors of the equations range from 0.89 (France) to 1.49 (Germany). The battery of tests on residuals confirms the validity of the models, since the errors seem to be free of autocorrelation, approximately normal and not heteroschedastic.[14] The models seem to be well specified, since both the RESET test for functional form and the Chow test for stability analysis provide satisfactory values for all countries.

[14] Only the p-value of the Ljung-Box statistic test for France falls below the five per cent significance level, but the value of the Lagrange Multiplier statistic – more suitable if there are lagged dependent variables among the regressors – does not signal autocorrelation in the residuals.

Table 1a

SINGLE-COUNTRY MODELS: ESTIMATION AND TESTING

France: dependent variable $y_t = \log(ipfrr)$

specification	coefficient estimate	t	diagnostic tests	statistic	p-value
y_{t-1}	0.330	3.441	Adjusted R^2	0.990	
y_{t-2}	0.504	5.677	Durbin-Watson	2.058	
$\log(elecons_t)$	0.126	3.875	S.E. of regression (%)	0.89	
$euindex_t$	0.00042	4.352	Ljung-Box (12)/(24)	22.8 /32.2	0.03 / 0.12
			Lagrange Multiplier (1-12)	1.75	0.08
			Skewness / kurtosis	0.271 / 2.798	
			Jarque-Bera	1.17	0.56
			ARCH(12)	13.867	0.31
			ADF (5 % critical value: -5.23)	-7.69	
			RESET	2.726	0.07
			CHOW (stability test)	0.935	0.54

Table 1b

SINGLE-COUNTRY MODELS: ESTIMATION AND TESTING

Germany: dependent variable $y_t = \log(ipder)$

specification	coefficient estimate	t	diagnostic tests	statistic	p-value
$\text{m.a.}(ord_{t-1},12)$	0.00108	7.948	Adjusted R^2	0.963	
$\Delta\log(ipus_t)$	1.087	3.686	Durbin-Watson	2.017	
$euindex_t$	0.00135	7.327	S.E. of regression (%)	1.49	
$\text{m.a.}(\log(comp_t),12)$	-0.328	-2.903	Ljung-Box (12)/(24)	20.1 / 31.9	0.07 / 0.13
			Lagrange Multiplier (1-12)	1.40	0.20
			Skewness / kurtosis	-0.301 / 3.398	
			Jarque-Bera	1.83	0.40
			ARCH(12)	9.221	0.68
			ADF (5 % critical value: -5.23)	-5.96	
			RESET	1.484	0.23
			CHOW (stability test)	1.229	0.28

Table 1c

SINGLE-COUNTRY MODELS: ESTIMATION AND TESTING

Italy: dependent variable $y_t = ipitr$

specification	coefficient estimate	t	diagnostic tests	statistic	p-value
m.a.$(y_{t-1}, 2)$	0.395	6.218	Adjusted R^2	0.995	
log($elecons_t$)	76.368	10.881	Durbin-Watson	1.973	
weather	1.333	4.567	S.E. of regression	1.07	
weather^2	-0.0519	-5.297	Ljung-Box (12)/(24)	5.1 / 21.1	0.95 / 0.63
			Lagrange Multiplier (1-12)	0.37	0.97
			Skewness / kurtosis	-0.402 / 3.364	
			Jarque-Bera	2.73	0.26
			ARCH(12)	6.423	0.89
			ADF (5 % critical value: -5.23)	-7.09	
			RESET	2.780	0.07
			CHOW (stability test)	1.467	0.14

In order to evaluate the prediction performance of the models, the estimation procedure was reiterated using a rolling regression technique. Starting from the sample January 1993 – December 1997, thus using a window width of five years, I derived the one-step-ahead forecasts for the period ranging from January 1998 to November 2000 (the last month in which French electricity consumption was available); I then checked for both in-sample and out-of-sample forecasting ability. The values of the conventional root mean square error statistic are very similar for France and Italy (respectively 1.30 and 1.37; see Table 2, first three lines), higher for Germany (1.96); the same ranking holds also if we analyze the results from the mean absolute error statistic.

In Figure 3 the comparison of the patterns of the actual series with those of the fitted values seems to confirm the overall promising forecasting performance. As in Bruno and Lupi (2001), I looked at the directional *forecasts*: the proportion of forecasts whose sign was correct[15] was 88 per cent (30 out of 34) for Germany and Italy and 97 per cent (33 out of 34) for France.

Table 2

ALL MODELS: PREDICTION PERFORMANCE

Model	Mean error (ME)	Mean absolute error (MAE)	Root mean square error (RMSE)	Theil U	Fraction of RMSE due to resid. var.	diff. slope from 1	bias
France	0.20	1.01	1.30	0.011	0.96	0.02	0.02
Germany	0.16	1.68	1.96	0.017	0.92	0.07	0.01
Italy	-0.06	1.07	1.37	0.013	0.99	0.00	0.01
Separate-countries	0.14	0.89	1.07	0.009	0.96	0.02	0.02
Sub-index	0.13	0.90	1.07	0.009	0.96	0.02	0.02
BGP	0.43	1.01	1.23	0.011	0.84	0.04	0.12
All-countries	0.26	0.91	1.10	0.010	0.93	0.01	0.06
Separate-cou.+US	0.18	0.89	1.07	0.009	0.96	0.01	0.03

Notes: lines 4-8 in the table refer to models that will be introduced later.

3. Models for the euro area

3.1 Taxonomy and choice of the models

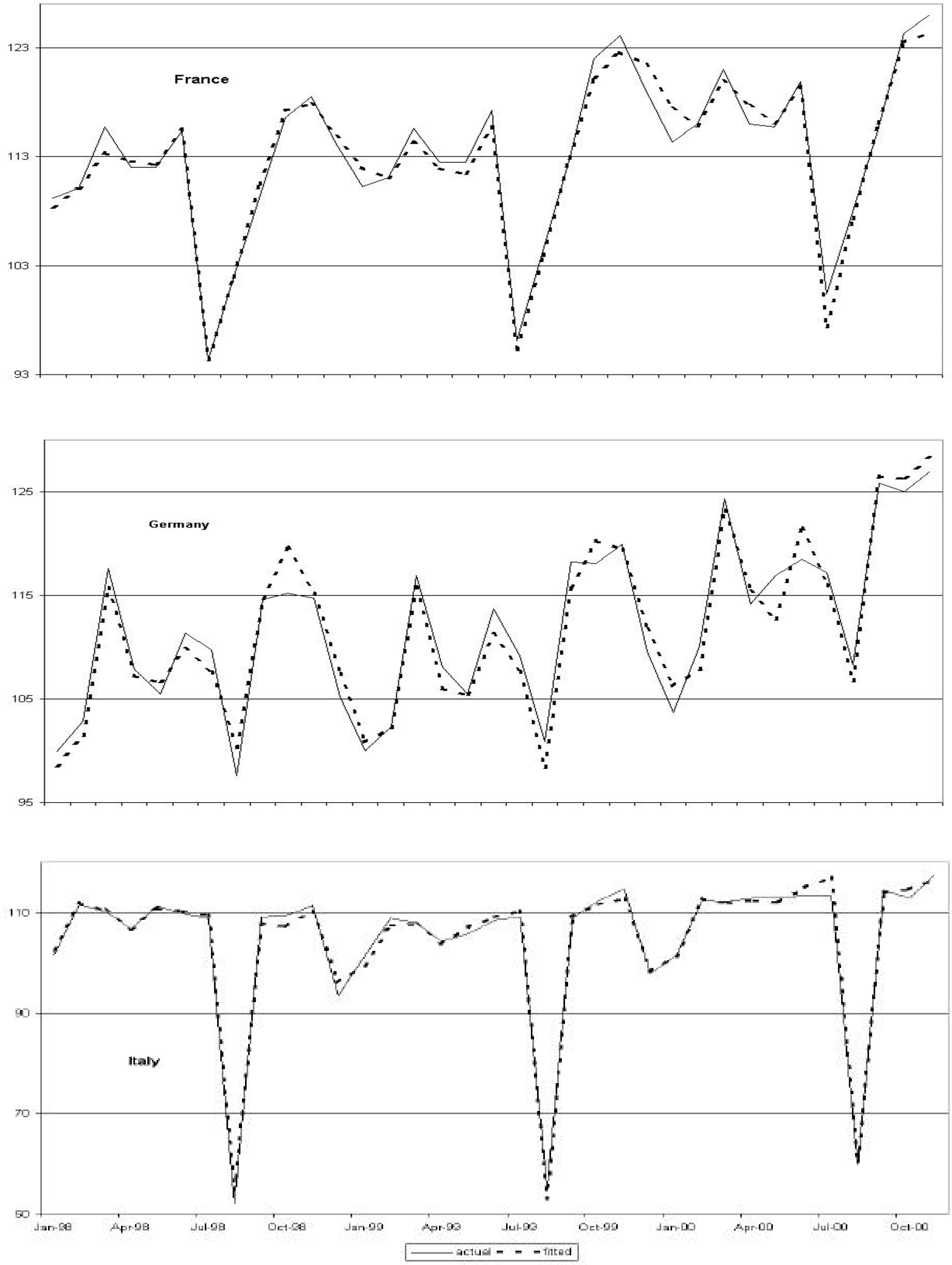
So far we have obtained - and tested - separate models for the major EMU nations. We now face the problem of predicting the evolution of manufacturing activity for the euro area as a whole. The following strategies were used:

a. **considering the past behaviour of the aggregate index**, hence identifying an autoregressive model or, more generally, a model within the ARIMA class. This approach has been successful in some studies, but, as Bodo, Golinelli and Parigi (2000) argue, the task is actually not too hard since this composite measure shows quite a smooth evolution as

[15] Actually, the authors relaxed the concept of "correct sign" of a forecast by saying that "if a prediction has wrong sign, but the difference with the actual growth rate is less than one percentage point, the sign of the forecast is correct" (Bruno and Lupi (2001), p. 15). I adopt a narrower definition and say that the sign of the forecast is correct *only if* the actual growth rate has the same sign.

Figure 3

INDUSTRIAL PRODUCTION INDEX: ACTUAL VERSUS FITTED VALUES



asymmetries among national cycles tend to be “averaged out”.[16] Nevertheless, in case of extreme smoothness, problems may arise in the detection of the turning points. Within this strategy, I built some autoregressive models (with different lag structures) but the results – available on request from the author – suggested it would be better to proceed in other directions;

b. **employing indicators for the aggregate index**, such as production indices for countries not in the Monetary Union or variables from qualitative surveys. One model of this type is that constructed by Bodo, Golinelli and Parigi (2000), whose attractiveness lies in its connecting the euro-area index with the BCI for the whole area (derived from the European Commission survey) and with the US index of industrial production. As the authors point out, the latter variable is famous for its timeliness. It can be taken here as a proxy of the evolution of demand outside the euro area;

c. **using a sub-index as a regressor by combining forecasts for the French, German and Italian indices** with adequate weights. A natural choice is to adopt the same weights used by Eurostat for calculating the official index;[17]

d. **regressing the aggregate index against the forecasts of the three countries' production indices** kept separately (borrowing the concept from Parigi and Schlitzer (1995), in a sort of 'bridge model');

e. **introducing in *d*, as a further explanatory variable, the forecast obtained from a model for all other countries in the euro area jointly considered.** In this model the dependent variable – the linear combination, using Eurostat weights, of the production indices of all EMU countries except Germany, France and Italy – was regressed against its own lag and the coincident value of the US production index.

---

[16] Analyzing the stability of European money demand, Arnold (1994) argued that a stable aggregated variable may derive from unstable national components owing to the desynchronisation of the sources of instability in the different countries. On the other hand, if the factors of instability act synchronically, the aggregate is not necessarily more stable than the single components. However, in the case of industrial production, where there is no "centralization" of the sources of instability among countries, aggregation should reduce the volatility of the series.

[17] The weights provided by Eurostat (in February 2001) are 0.373 for Germany, 0.191 for France and 0.185 for Italy.

Differences emerging from the assessment of approaches *c* and *d* should provide useful insights into the matter of convergence among nations. Method *d* should in theory prove to be better, the lower the comovement the countries show; if the three series exhibit a high correlation, it seems advisable to collapse them into a single regressor, at least to avoid problems of multicollinearity that could affect the quality of the estimated coefficients. The results of some studies of the correlation between the production indices for the main countries[18] suggest at this stage that aggregating the three countries is not likely to result in an improvement, since each of the indices still conveys its own information content.

Therefore, we want to assess the validity of the "sub-index" model, belonging to class c:

$$\log(ipeuro_t) = a + b\,\log(ipeuro_{t-1}) + g\,\log(subindex_t) + l \cdot trend + seasonals + u_t,$$

of the "separate-countries" model, falling in category d:

$$\begin{aligned}\log(ipeuro_t) &= a + b\,\log(ipeuro_{t-1}) + g\,\log(ipfra_t) + d\,\log(ipger_t) + z\,\log(ipita_t) \\ &+ l \cdot trend + seasonals + u_t\end{aligned}$$

and of the "all-countries" model, classified in e:

$$\log(ipeuro_t) = a + b\,\log(ipother_t) + g\,\log(ipfra_t) + d\,\log(ipger_t) + z\,\log(ipita_t) + seasonals + u_t,$$

where *ipeuro* is the industrial production index for the whole area, *ipfra*, *ipger*, *ipita* and *ipother* indicate respectively the forecasts of the production index for France, Germany, Italy and all other countries, and *subindex* is the weighted average of *ipfra*, *ipger* and *ipita.*

Relevant evidence will be also derived by comparing the performance of the above

[18] As a preliminary check I looked at the correlation coefficients over different periods (1992-2000,1992-1994,1995-1997,1998-2000), obtaining respectively the following coefficients: France/Germany: 0.792, 0.692, 0.669, 0.617; Germany/Italy: 0.607, 0.562, 0.628, 0.563; Italy/France: 0.648, 0.646, 0.614, 0.611. As time passes, I observed a decrease – or at most a substantial stationarity - in the correlation for all pairs of countries: the relation among countries tends to become weaker. Taking the determinant of the correlation matrix $R = \{r_{ij}\}$ as a measure of collinearity (in presence of multicollinearity, the matrix is almost singular), I found further evidence against an increase in the comovement among countries. In fact, instead of diminishing, the value of the determinant rises from 0.29 in the first period to 0.35 in the last. However, these results should be considered with caution, because of the different volatility exhibited by the indices in the above ranges.

models with that of the model by Bodo, Golinelli and Parigi (2000, hereafter referred to as "BGP" model).

3.2 Empirical results

As shown by the coefficient estimates and the results from the usual battery of tests, the models reported in Tables 3a-3c[19] (i.e. separate-countries, sub-index and BGP) are in all satisfactory.[20] The same is not true for the all-countries model (shown in Table 3d): the RESET test for functional form strongly rejects the hypothesis of linearity, and the residuals are definitely skewed and nonnormal.

The lowest standard error of the regression is provided by the all-countries model (0.11); at the same time, the standard error also improves with respect to BGP by using indifferently the separate-countries or the sub-index models, since the statistic for the BGP model is 0.84 against values of about 0.24 for the other two.

Also the comparison of the RMSE provides evidence indifferently in favour of the separate-countries or the sub-index models against the BGP model (1.07 for the first two, 1.23 for the third; see Table 2). The RMSE for the all-countries model is 1.10 and thus does not mark an improvement in forecasting performance with respect to both separate-countries and sub-index models. Further, it can be observed that the forecasting ability shown by all models for the euro area sharply outperforms that for the single-country indices (the RMSE ranges from 1.07 to 1.23 for the euro-area models and from 1.30 to 1.96 for single-country models). The aggregation process smoothes out sharp variations in the single indices, with an offsetting (and not a cumulation) of the forecast errors.

[19] Estimation range: 1993.01 – 1999.12. Coefficient estimates of the seasonal dummies, the time trends and the constant are omitted. Results for the BGP model may differ from those reported in Bodo, Golinelli and Parigi (2000) due to the different range used for the estimation.

[20] The only test that is not accepted at a 5 per cent level of confidence is the Ljung-Box test for the separate-countries model, which has a significance level of 4 per cent. However, the t-statistic for the autocorrelation coefficients (from lag 1 to lag 24) is always - in modulus - below 2, and the Lagrange Multiplier test signals no autocorrelation in the residuals.

Table 3a

EURO-AREA MODELS: ESTIMATION AND TESTING

Separate-countries model: dependent variable $\log(ipeuro_t)$

specification	coefficient estimate	t	Diagnostic tests	statistic	p-value
$\log(ipeuro_{t-1})$	0.0932	3.051	Adjusted R^2	0.999	
$\log(ipfra_t)$	0.238	7.233	Durbin-Watson	1.600	
$\log(ipger_t)$	0.441	21.630	S.E. of regression (%)	0.245	
$\log(ipita_t)$	0.183	19.251	Ljung-Box (12)/(24)	21.8 / 32.8	0.04 / 0.11
			Lagrange Multiplier (1-12)	1.21	0.30
			Skewness / kurtosis	0.291 / 2.780	
			Jarque-Bera	1.36	0.51
			ARCH(12)	13.617	0.33
			ADF (5 % critical value: -5.23)	-5.99	
			RESET	1.349	0.27
			CHOW (stability test)	0.669	0.82

Table 3b

EURO-AREA MODELS: ESTIMATION AND TESTING

Sub-index model: dependent variable $\log(ipeuro_t)$

specification	coefficient estimate	t	diagnostic tests	statistic	p-value
$\log(ipeuro_{t-1})$	0.0887	3.720	Adjusted R^2	0.999	
$\log(subindex_t)$	0.877	38.882	Durbin-Watson	1.590	
			S.E. of regression (%)	0.244	
			Ljung-Box (12)/(24)	14.5 / 24.2	0.27/0.45
			Lagrange Multiplier (1-12)	0.58	0.85
			Skewness / kurtosis	0.063 / 2.891	
			Jarque-Bera	0.10	0.95
			ARCH(12)	13.169	0.36
			ADF (5 % critical value: -5.23)	-5.41	
			RESET	1.758	0.18
			CHOW (stability test)	1.031	0.44

Table 3c

EURO-AREA MODELS: ESTIMATION AND TESTING

BGP model: dependent variable $\lfloor \log(\, ipeuro_t)$

specification	coefficient estimate	t	diagnostic tests	statistic	p-value
$\log(\, ipeuro_{t-1})$	-0.522	-5.669	Adjusted R^2	0.993	
$\Delta_{11} \log(\, ipeuro_t)_{-1}$	-0.290	-3.792	Durbin-Watson	2.024	
$\log(\, ipus_{t-1})$	0.283	5.529	S.E. of regression (%)	0.844	
$\lfloor \log(\, ipus_t)$	0.645	3.462	Ljung-Box (12)/(24)	7.6 / 11.2	0.81 / 0.99
$\lfloor \log(\, ipus_t)_{-1}$	-0.370	-1.849	Lagrange Multiplier (1-12)	0.57	0.86
BCI_{t-1}	0.00137	5.031	Skewness / kurtosis	-0.374 / 3.002	
$(\Delta BCI)_{-2}$	0.00179	2.539	Jarque-Bera	1.96	0.38
			ARCH(12)	4.967	0.96
			ADF (5 % critical value: -5.23)	-5.49	
			RESET	1.413	0.25
			CHOW (stability test)	1.685	0.08

Table 3d

EURO-AREA MODELS: ESTIMATION AND TESTING

All-countries model: dependent variable $\log(\, ipeuro_t)$

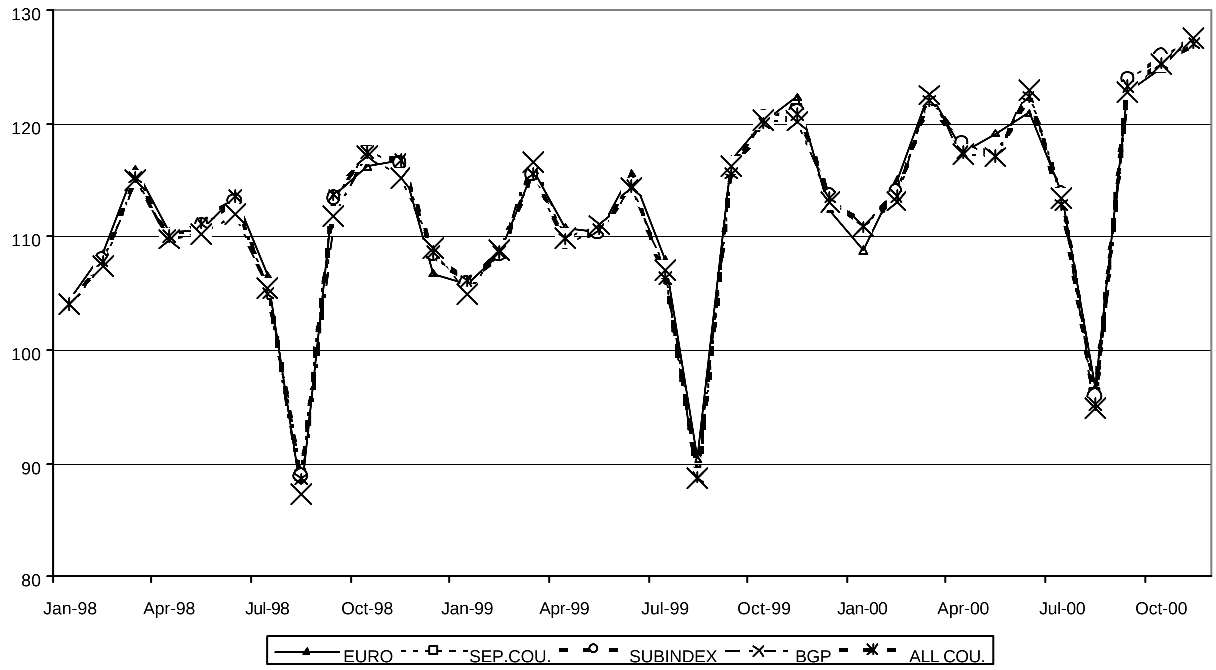
specification	coefficient estimate	t	diagnostic tests	statistic	p-value
$\log(\, ipother_t)$	0.263	33.24	Adjusted R^2	0.999	
$\log(\, ipfra_t)$	0.199	15.83	Durbin-Watson	1.882	
$\log(\, ipger_t)$	0.376	41.53	S.E. of regression (%)	0.111	
$\log(\, ipita_t)$	0.151	37.09	Ljung-Box (12)/(24)	8.4 / 20.6	0.75 / 0.66
			Lagrange Multiplier (1-12)	0.431	0.94
			Skewness / kurtosis	1.409 / 11.64	
			Jarque-Bera	289.1	0.0
			ARCH(12)	1.523	0.99
			ADF (5 % critical value: -5.23)	-6.23	
			RESET	18.42	0.0
			CHOW (stability test)	0.456	0.96

This holds only partially taking the Theil U statistic as a measure of the accuracy of the forecasts: this statistic, which does not have scaling problems and hence is not affected by the different degree of smoothness of the series, still takes the same value for the separate-countries and sub-index models, and a slightly higher one for the all-countries and BGP models; at the same time, however, the U-values for the euro-area models are not very different from those for the single-country models (see Table 2).

Figure 4a plots the observed series versus those predicted by the four models for the period from January 1998 to November 2000. As in the single-country models, the graph shows a close relationship between the series for all models. Only in very few episodes is there a perceptible deviation of the fitted values from the actual ones; in particular, the sign (or direction) of the forecast was inconsistent with the data in only one case (out of 34) for the separate-countries, the sub-index and the all-countries models, and in four cases for the BGP model.

Figure 4a

EURO-AREA INDEX: ACTUAL VERSUS FITTED VALUES



The results of testing and the lack of timeliness in the data for some of the smaller countries of the area argue against opting for the all-countries model, whose forecasting performance, moreover, has proven to be inferior to that of other models. On the other hand, the econometric analysis has not provided any clear indication of choice between the separate-countries and the sub-index models. However, the latter implies a constant revision of the weights supplied by Eurostat, which updates them as soon as new national data are released. Introducing the operational burden as a selection criterion, we can at this stage also exclude the sub-index model from further analysis and award preference to the model in which the forecasts of the production indices for France, Germany and Italy are separately considered.

4. To combine or not to combine?

When a wide range of explanatory variables is available and many different models can be built, it is natural to try to summarize all this information in order to obtain a better forecast.

I shall consider two types of combination. In the first, I will try to merge the information sets that up to now were separate in the models described. I will then tackle the problem of combining forecasts coming from the different models.

4.1 Combination of information sets

A natural extension of the analysis so far performed is the joint use of all the information available - the single-country indices, the BCI and the US index – on the hope that different variables will concur, and not conflict, in explaining the behaviour of the euro-area index.

The inclusion of the BCI index in the separate-countries model was not successful, since the coefficients attached to this variable - at several lags – were never significant. Results obtained by considering the US index as an additional predictor (the relative model will be referred to as “separate-countries+US”) are better than those for the BGP and broadly similar to those for the simple separate-countries model (in terms of both standard error of the regression and RMSE, respectively equal to 0.24 and 1.07; see Table 2 and Tables 3a-3e[21]). However, a substantial improvement in the specification of the model comes from replacing the time trend – previously included as a regressor and whose coefficient was highly significant – with a 12-term moving average for the US index.

Table 3e

EURO-AREA MODELS: ESTIMATION AND TESTING

Separate-countries+US model: dependent variable $\log(ipeuro_t)$

specification	coefficient estimate	t	diagnostic tests	statistic	p-value
$\log(ipeuro_{t-1})$	0.0737	2.353	Adjusted R^2	0.999	
$\log(ipfra_t)$	0.228	6.951	Durbin-Watson	1.533	
$\log(ipger_t)$	0.435	21.366	S.E. of regression (%)	0.244	
$\log(ipita_t)$	0.189	19.765	Ljung-Box (12)/(24)	14.3 / 33.7	0.28 / 0.09
m.a $(\log(IPUS_t),12)$	0.135	9.982	Lagrange Multiplier (1-12)	0.93	0.53
			Skewness / kurtosis	0.409 / 2.571	
			Jarque-Bera	2.98	0.23
			ARCH(12)	11.086	0.52
			ADF (5 % critical value: -5.23)	-5.57	
			RESET	1.051	0.36
			CHOW (stability test)	0.987	0.49

Figure 4b shows that during the simulation range January 1998 - November 2000 the separate-countries+US model provides predictions that track the observed euro-area industrial production index quite well. Further, the model fails to forecast in the right direction – that is, the sign of the prediction is incorrect – in only one case (May 2000).

4.2 Combination of forecasts

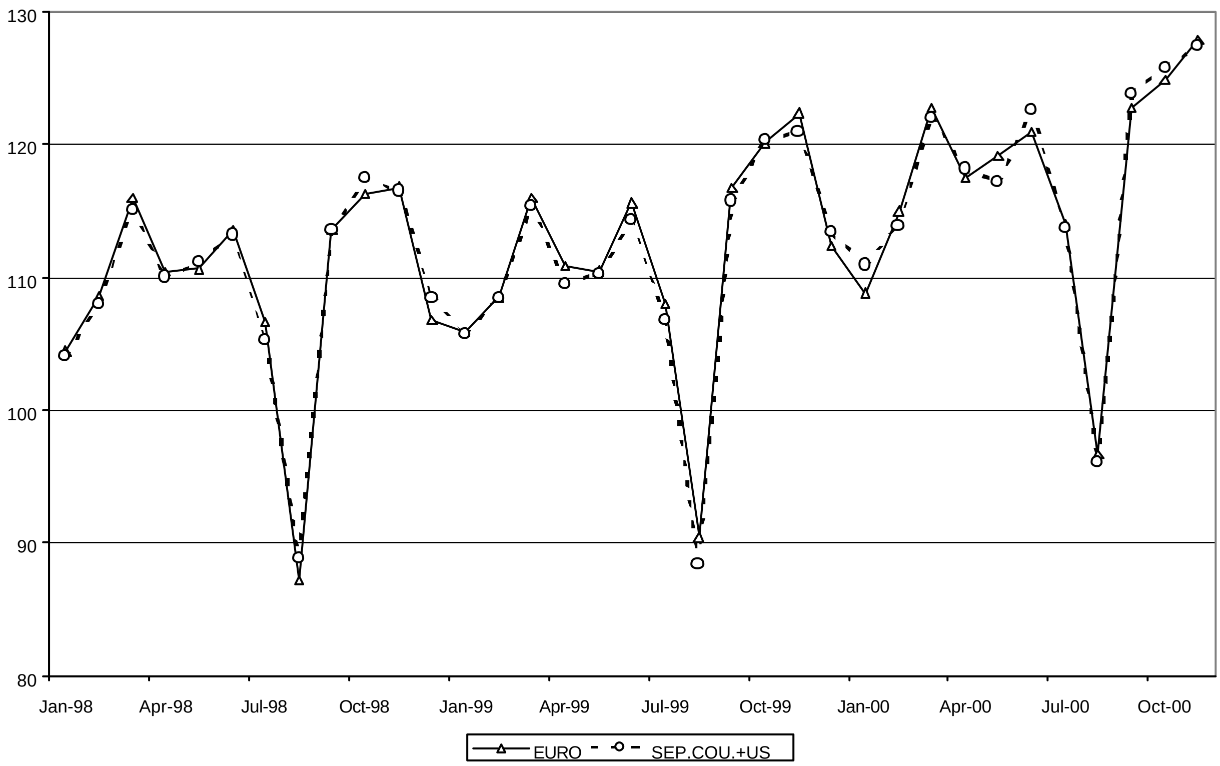
Rather than combining information sets *ex-ante*, we could combine forecasts obtained from two (or more) independent models *ex-post*. Despite the evidence of strong

[21] Estimation range: 1993.01 – 1999.12. Coefficient estimates of the seasonal dummies and of the constant are omitted; m.a. = moving average.

collinearity among individual forecasts, observed by Diebold and Lopez (1996) and Gregory *et al.* (2001), when a set of different forecasts of the same event is available, it is "good practice" to combine them; doing so should produce a more accurate forecast.

Figure 4b

EURO-AREA INDEX: ACTUAL VERSUS FITTED VALUES



The synthesis is usually performed by means of a linear combination of the competing forecasts; very often, it reduces to a simple average (often defined as "consensus forecast"). The composite forecast can be obtained as a convex combination of the individual ones, if the coefficients are constrained to be positive and add up to one, or using a more general system of weights.[22]

[22] Bates and Granger (1969), for example, employ a variance-covariance method: given two unbiased forecasts combined through weight vector between zero and one, they easily showed that the optimal weight (in the sense of minimizing the mean squared prediction error) is $\alpha^* = \dfrac{\sigma_{22}^2 - \sigma_{12}}{\sigma_{22}^2 + \sigma_{11}^2 - 2\sigma_{12}}$, where σ's are unconditional forecast error variances/covariances.

Alternatively, one can pool the forecasts by using regression techniques, as in Granger and Ramanathan (1984). Combinations of forecasts can then be obtained by estimating a restricted or unrestricted OLS regression of realizations on the separate forecasts. Given n different forecasts $F_{1t}, F_{2t}, ..., F_{nt}$, one estimates

$$y_t = b_0 + b_1 F_{1t} + ... + b_n F_{nt} + u_t$$

where y_t is the actual value and u_t is a zero-mean disturbance term. Once we estimate the equation, we can use it post-sample to derive the optimal combined forecast.

Granger and Ramanathan compared different regression-based methods. The "constrained form", where coefficients are imposed to sum up to one, has the advantage of yielding an unbiased composite forecast, provided the individual forecasts have zero average errors. This kind of regression was found to be suboptimal compared with the unconstrained linear regression including a constant term; the optimality of the latter resulted not only from the minimization of the mean squared error, but also from the unbiasedness of the combined forecast regardless of the biasedness of the component ones.

Accordingly, I ran an *unconstrained* regression: I included a constant term, took the actual value of the euro-area index as the response variable and the forecast coming from the BGP model as one of the two regressors (called forecast2). The other regressor was the forecast generated by the separate-countries+US model (called forecast1). The regression equation is[23]

$$ipeuro_t = a + b \cdot forecast1 + g \cdot forecast2 + u_t .$$

The estimation was performed over the sample January 1998 - November 2000; the results from the OLS regression are in Table 4 and can be used to assess whether one forecast encompasses the other.

Surprisingly, the sum of the coefficients is almost equal to one, although the

[23] I also tried the logarithmic version of the above regressions, obtaining roughly the same results.

coefficient attached to the forecast coming from the BGP model is not significant at standard significance level. The constant is insignificant: this is good news, since it means that our primary forecasts do not systematically under – or over – estimate the actual value of the euro-area index.[24]

Table 4

FORECAST COMBINATION. RESULTS FROM OLS REGRESSION

constant		separate-countries+US forecast		BGP forecast	
coeff.	t-statistic	coeff.	t-statistic	coeff.	t-statistic
2.290	[0.96] {0.88}	0.69	[2.48] {2.89}	0.29	[1.07] {1.23}

Notes: t-statistics in brackets; White t-statistics in braces.

If, for example, $\alpha = 0$, $\beta = 1$ and $\gamma = 0$, u_t reduces to the forecast error from the first model; given that forecast errors are in general heteroschedastic, I applied the White procedure to obtain heteroschedastic-consistent estimator of the variance of the OLS estimator. The relative t-statistics are appreciably higher than the corresponding standard t, but not enough to allow acceptance of a coefficient for the BGP forecast different from zero.

Since the coefficient estimate is significant for the first forecast but not for the second, the former is said to "dominate" the latter (see Fair and Shiller (1987)); in other words, the first original forecast completely obscures the other, incorporating all the information relevant to forecasting the euro-area index and thus implying that combining the forecasts is in this case improductive. This exercise confirms the preference given to the separate-countries+US specification.

As in Fair and Shiller (1987; 1990), as a final step I tested the validity of the models by regressing the actual logarithmic changes of the euro-area index (i.e. its growth rates) against

[24] The intercept was not significant even when I regressed the euro-area index against each of the forecasts, one at a time. The coefficient attached to the forecasts was in all cases equal to 0.98 (omitting a priori the constant, equal to 1).

forecast changes from the same two models compared above; in formula

$$\left(\log\left(ipeuro_t\right)-\log\left(ipeuro_{t-1}\right)\right)=a+b\left(\log\left(forecast1\right)-\log\left(ipeuro_{t-1}\right)\right)+$$
$$+g\left(\log\left(forecast2\right)-\log\left(ipeuro_{t-1}\right)\right)+u_t.$$

The estimates are reported in Table 5. Again, the estimated coefficients approximately add up to one (although the regression is still unconstrained); the difference between the standard and the White t-statistics is not as large as before, thus providing evidence for a less heteroschedasticity in the errors. Since the estimate of β (the coefficient attached to the first forecast) is nonzero and that of γ (the coefficient for the second forecast) is not significant, it can be implied that both models individually contain information to forecast the euro-area index, but when the models are jointly considered the BGP model does not provide any additional information with respect to that derived from the separate-countries+US model.

Table 5

FAIR AND SHILLER TEST

constant		separate-countries+US forecast		BGP forecast	
coeff.	t-statistic	coeff.	t-statistic	coeff.	t-statistic
0.003	[1.38] {1.36}	0.66	[2.43] {2.52}	0.33	[1.23] {1.27}

Notes: t-statistics in brackets; White t-statistics in braces.

5. Conclusion

This paper has focused on country-specific, indicator-based, regression models for the industrial production index of France, Germany, Italy and of the euro area. Future efforts could be devoted to investigating a multivariate approach (for example, structural VAR or VECM).

The overall goodness of fit and forecasting performance of all the models can be regarded as satisfactory. Especially encouraging were the results obtained by plugging the forecasts for the major euro-area countries into the model for the area, both separately and aggregated to form a sub-index. The introduction of the forecast for the production index of the minor euro-area countries as a predictor increased the precision of the estimates, but did not enhance the forecasting ability of the model with respect to those based only on the forecasts for the main economies.

Further improvement was achieved by (ex-ante) combining the information sets used in the model proposed by Bodo, Golinelli and Parigi and in the separate-countries model; the result is the separate-countries+US model, in which the forecasts coming from the country-specific models and the actual US production index have a role in explaining the evolution of the euro-area level of activity.

At least from a theoretical point of view, an optimal solution should be the ex-post pooling (through a linear regression with a constant term) of two primary forecasts, in this case the first obtained from the separate-countries+US model and the second from the BGP model. I have shown that linearly combining the forecasts for the euro-area index from this pair of models does not produce the expected outcome, since the aggregate BGP model is "dominated" by the other competing forecast. This provides further evidence in favour of the aggregation of national forecasts.

References

Altissimo, F., Bassanetti, A., Cristadoro, R., Reichlin, L. and Veronese, G. (2001), *"The construction of coincident and leading indicators for the euro area business cycle"*, Banca d'Italia, Tema di discussione No. 434.

Arnold, I.J.M. (1994), "The myth of a stable European money demand", *Open Economies Review*, **5**, pp. 249-259.

Baffigi, A. and Bassanetti, A. (2001), *"Cyclical indicators for the euro area: real time estimates of turning point probabilities"*, Banca d'Italia, manuscript.

Bates, J.M. and Granger, C.W.J. (1969), "The combination of forecasts", *Operations Research Quarterly*, **20**, pp. 451-468.

Bodo, G., Golinelli, R. and Parigi, G. (2000), *"Forecasting industrial production in the euro area"*, Banca d'Italia, Tema di discussione No. 370.

Bodo, G. and Signorini, L.F. (1987), "Short-term forecasting of the industrial production index", *International Journal of Forecasting*, **3**, pp. 245-259.

Bruno, G. and Lupi, C. (2001), *"Forecasting industrial production and the early detection of turning points"*, ISAE, manuscript.

Carnazza, P. and Parigi, G. (2001), *"The evolution of confidence for European consumers and businesses in France, Germany and Italy"*, Banca d'Italia, Tema di discussione No. 406.

Clemen, R.T. (1989), "Combining forecasts: a review and annotated bibliography", *International Journal of Forecasting*, **5**, pp. 559-583.

Darnell, A.C. (1994), *A dictionary of econometrics*, Edward Elgar, Aldershot.

Deutsch, M., Granger, C.W.J. and Terasvirta, T. (1994), "The combination of forecasts using changing weights", *International Journal of Forecasting*, **10**, pp. 47-57.

Diebold, F.X. and Lopez, J.A. (1996), "Forecast evaluation and combination", *NBER Technical Working Paper* 192.

Fair, R.C. and Shiller, R.J. (1987), "Econometric modeling as information aggregation", *NBER Working Paper* 2233.

Fair, R.C. and Shiller, R.J. (1990), "Comparing information in forecasts from econometric models", *American Economic Review*, **3**, pp. 375-389.

Granger, C.W.J. and Ramanathan, R. (1984), "Improved methods of combining forecasts", *Journal of Forecasting*, **3**, pp.197-204.

Granger, C.W.J. and Terasvirta, T. (1993), *"Modelling nonlinear economic relationships"*, Oxford: Oxford University Press.

Greene, W.H. (1997), *"Econometric analysis"*, 3rd Edition, Prentice Hall, New Jersey.

Gregory, A.W., Smith, G.W. and Yetman, J. (2001), “Testing for forecast consensus”, *Journal of Business & Economic Statistics*, **19**, pp. 34-43.

Grunfeld, Y. and Griliches, Z. (1960), “Is aggregation necessarily bad?”, *The Review of Economics and Statistics*, **1**, pp. 1-13.

Harvey, A.C. (1993), *Time series models*, 2nd edition, Harvester Wheatsheaf, New York.

Holden, K. and Thompson, J. (1997), “Combining forecasts, encompassing and the properties of UK macroeconomic forecasts”, *Applied Economics*, **29**, pp. 1447-1458.

Hoover, K.D. and Perez, S.J. (1999), “Data mining reconsidered: encompassing and the general-to-specific approach to specification search”, *Econometrics Journal*, **2**, pp. 167-191.

Lin, C.J. and Terasvirta, T. (1994), “Testing the constancy of regression parameters against continuous structural change” *Journal of Econometrics*, **62**, pp. 211-228.

Marchetti, D.J. and Parigi, G. (2000), “Energy consumption, survey data and the prediction of industrial production in Italy: a comparison and combination of different models”, *Journal of Forecasting*, **19**, pp. 419-440.

McConnell, M.M. and Perez-Quiros, G. (2000), “Output fluctuations in the United States: what has changed since the early 1980’s?”, *American Economic Review*, **5**, pp. 1464-1476.

Palm, F.C. and Zellner, A. (1991), *“To combine or not to combine? Issues of combining forecasts”*, Paper Univ. Catholique Louvain, Center Operations Research and Econometrics.

Parigi, G. and Schlitzer, G. (1995), “Quarterly forecasts of the Italian business cycle by means of monthly economic indicators”, *Journal of Forecasting*, **14**, pp. 117-141.

Terasvirta, T. (1984), “Short-term forecasting of industrial production by means of quick indicators”, *Journal of Forecasting*, **3**, pp. 409-416.

RECENTLY PUBLISHED "TEMI" (*)

No. 418 — *Ingredients for the New Economy: How Much does Finance Matter?*, by M. BUGAMELLI, P. PAGANO, F. PATERNÒ, A.F. POZZOLO, S. ROSSI and F. SCHIVARDI (October 2001).

No. 419 — *ICT Accumulation and Productivity Growth in the United States: an Analysis Based on Industry Data,* by P. CASELLI and F. PATERNÒ (October 2001).

No. 420 — *Barriers to Investment in ICT,* by M. BUGAMELLI and P. PAGANO (October 2001).

No. 421 — *Struttura dell'offerta e divari territoriali nella filiera dell'information and communication technologies in Italia,* by G. IUZZOLINO (October 2001).

No. 422 — *Multifactor Productivity and Labour Quality in Italy, 1981-2000,* by A. BRANDOLINI and P. CIPOLLONE (October 2001).

No. 423 — *Tax reforms to influence corporate financial policy: the case of the Italian business tax reform of 1997-98,* by A. STADERINI (November 2001).

No. 424 — *Labor effort over the business cycle,* by D. J. MARCHETTI and F. NUCCI (November 2001).

No. 425 — *Assessing the effects of monetary and fiscal policy,* by S. NERI (November 2001).

No. 426 — *Consumption and fiscal policies: medium-run non-Keynesian effects,* by G. RODANO and E. SALTARI (November 2001).

No. 427 — *Earnings dispersion, low pay and household poverty in Italy, 1977-1998,* by A. BRANDOLINI, P. CIPOLLONE and P. SESTITO (November 2001).

No. 428 — *Nuove tecnologie e cambiamenti organizzativi: alcune implicazioni per le imprese italiane,* by S. TRENTO and M. WARGLIEN (December 2001).

No. 429 — *Does monetary policy have asymmetric effects? A look at the investment decisions of Italian firms,* by E. GAIOTTI and A. GENERALE (December 2001).

No. 430 — *Bank-specific characteristics and monetary policy transmission: the case of Italy,* by L. GAMBACORTA (December 2001).

No. 431 — *Firm investment and monetary transmission in the euro area,* by J. B. CHATELAIN, A. GENERALE, I. HERNANDO, U. VON KALCKREUTH and P. VERMEULEN (December 2001).

No. 432 — *Financial systems and the role of banks in monetary policy transmission in the euro area,* by M. EHRMANN, L. GAMBACORTA, J. MARTÍNEZ-PAGÉS, P. SEVESTRE and A. WORMS (December 2001).

No. 433 — *Monetary policy transmission in the euro area: what do aggregate and national structural models tell us?,* by P. VAN ELS, A. LOCARNO, J. MORGAN and J.P. VILLETELLE (December 2001).

No. 434 — *The construction of coincident and leading indicators for the euro area business cycle,* by F. ALTISSIMO, A. BASSANETTI, R. CRISTADORO, L. REICHLIN and G. VERONESE (December 2001).

No. 435 — *A core inflation index for the euro area,* by R. CRISTADORO, M. FORNI, L. REICHLIN and G. VERONESE (December 2001).

No. 436 — *A real time coincident indicator of the euro area business cycle,* by F. AITISSIMO, A. BASSANETTI, R. CRISTADORO, M. FORNI, M. LIPPI, L. REICHLIN and G. VERONESE

No. 437 — *The use of preliminary data in econometric forecasting: an application with the Bank of Italy Quarterly Model,* by F. BUSETTI (December 2001).

No. 438 — *Financial crises, moral hazard and the "speciality" of the international interbank market: further evidence from the pricing of syndicated bank loans to emerging markets,* by F. SPADAFORA (March 2002).

No. 439 — *Durable goods, price indexes and quality change: an application to automobile prices in Italy, 1988-1998,* by G. M. TOMAT (March 2002).

No. 440 — *Bootstrap bias-correction procedure in estimating long-run relationships from dynamic panels, with an application to money demand in the euro area,* by D. FOCARELLI (March 2002).

(*) Requests for copies should be sent to:
Banca d'Italia – Servizio Studi – Divisione Biblioteca e pubblicazioni – Via Nazionale, 91 – 00184 Rome (fax 0039 06 47922059). They are available on the Internet at www.bancaditalia.it