Forecasting Euro-Area Industrial Production Using (Mostly) Business Surveys Data

Bruno, Giancarlo and Lupi, Claudio

ISAE (Institute for studies and economic analisys) Rome (Italy)

October 2003

ISAE Istituto di Studi e Analisi Economica

FORECASTING EURO-AREA INDUSTRIAL PRODUCTION USING (MOSTLY)\ BUSINESS SURVEYS DATA

by

Giancarlo Bruno, Claudio Lupi

ISAE

Rome

March, 2003

ABSTRACT*

In this paper we propose a relatively simple procedure to predict Euro-zone industrial production using mostly data derived from the business surveys of the three major economies within the European Monetary Union (France, Germany, and Italy). The basic idea is that of estimating business cyclical indicators to be used as predictors for the industrial production in France and Germany; as far as Italy is concerned, forecasts are produced using a model that in the recent past proved to be able to produce accurate forecasts up to six months ahead. In order to derive quantitative predictors from the business surveys data and to aggregate the nation-wide forecast into the Euro-zone forecast, we propose using an approach based on dynamic factors and unobserved components models. The resulting forecasts are accurate up to six steps ahead.

JEL Classification: C53, C32, E32

Key Words: Forecasting, VAR models, Industrial production, Cyclical Analysis

----

We would like to thank, without implicating, Sergio de Nardis and Fabien Toutlemonde for comments and discussion. The opinions expressed in this paper are solely the responsibility of the authors and should not be interpreted as reflecting the views of ISAE or its staff.. Thanks to Gianna Foschi for her excellent editorial assistance.

NON TECHNICAL SUMMARY

The forecasting of euro-zone main economic aggregates can be done aggregating the national forecasts or forecasting the aggregate. In this paper we focus our attention of forecasting the monthly industrial production index in both ways, using national leading information coming from qualitative business surveys.
The disaggregate approach, besides giving an added value consisting in the knowledge of national forecasts, gives better results also on the aggregate euro area. The forecasting experiment is carried out on raw data, so to take into accounts revisions stemming from seasonal adjustment. Useful forecasts can be produced up to six months ahead.

UN INDICATORE DI FIDUCIA PER L'ECONOMIA ITALIANA

SINTESI

IN questo lavoro viene proposta una procedura di previsione dell'indice della produzione industriale dell'area euro utilizzando essenzialmente i risultati delle inchieste congiunturali coordinate dalla Commissione europea per i tre maggiori paesi dell'unione monetaria: Francia, Germania e Italia. Le inchieste congiunturali vengono utilizzate per costruire un indicatore composito per ogni paese con un modello a fattori dinamici. Per l'Italia vengono utilizzati i risultati di un modello, sviluppato presso l'ISAE, capace di produrre previsioni affidabili fino a sei passi in avanti. Infine, un modello strutturale di serie storiche viene utilizzato per aggregare i risultati nazionali e costruire le previsioni per l'area dell'euro. Tali previsioni sono valide fino a sei passi in avanti.

Classificazione JEL: C53, C32, E32

Parole chiave: Previsioni, modelli VAR, produzione industriale, analisi del ciclo economico

INDEX

1. INTRODUCTION

A great deal of effort has been devoted in the last few years to obtain accurate forecasts of the industrial production index (IPI) in many countries (see, among others, Bruno and Lupi, 2001; Huh, 1998; Marchetti and Parigi, 2000; Osborn *et al.*, 1999; Simpson *et al.*, 2001). More recently, as a result of the increased European economic integration, interest has been partly shifted towards predicting Euro-zone macroeconomic aggregates rather than national ones (see *e.g.*, Baffigi *et al.*, 2002; Ballabriga and Castillo, 2000; Bodo *et al.*, 2000; Buffeteau and Mora, 2000; Marcellino *et al.*, 2001; Zizza, 2002). The interest in this new research area is justified by the need, on the part of public and private economic agents, of monitoring the European economy for more accurate decision making.

In this note we investigate the possibility of predicting the Euro-zone IPI using mainly data derived from the business surveys carried out in three major countries in the European Union, namely France, Germany, and Italy.

The remainder of the paper is organized as follows: section 2 is devoted to the derivation of the main predictors for France and Germany from the business surveys data. The description of the forecasting model for France and Germany follows. The last part of the section is dedicated to a brief discussion of the model used to forecast the Italian industrial production. In section 3 the aggregation of the national predictions into the Euro-zone forecast is illustrated and the results are evaluated in terms of forecast errors statistics, compared to the direct forecast of aggregate Euro-zone IPI using a univariate model and a model using survey data. A final section draws some concluding remarks.

2. FORECASTING NATIONAL INDUSTRIAL PRODUCTION INDEXES

We are interested in forecasting the Euro-zone industrial production index. The definition we use for industrial production covers NACE Rev. 1 Sections C (products from mining and quarrying), D (manufactured products), and E (electrical energy, gas, steam, and hot water) and excludes constructions. We consider forecasting official industrial production series a priority and for this reason we refer to Eurostat data downloaded from the New Cronos database for France and Germany and to Istat (the Italian National Statistical Institute) series for Italy. The main idea we pursue in this paper is to forecast the Euro-zone industrial production by using the forecasts of the three major countries as indicators. In this sense, the approach is not far from the one followed in Zizza (2002): however, the way this goal is pursued is rather different. The reason for aggregating national forecasts, rather than forecasting the aggregate, is threefold. First, possibly asymmetric national shocks may be important in explaining aggregate variations in industrial activity. Second, better national forecasts can be substituted in the forecasting model as soon as they become available, no matter the method they are based on: indeed, we develop a stylized preliminary model to forecast industrial production in France and Germany while we use an already existing, well consolidated model for Italian industrial activity. Third, this way of building aggregate predictions allows to enlarge the set of countries considered in the analysis in a rather straightforward manner. Given that we want to obtain multi-step forecasts, we need a technology that allows reaching this goal. At this stage of our research we rely on the VAR methodology. As usual, some compromises and simplifications are necessary: while we would have preferred using raw data, Eurostat's official series are either seasonally adjusted and corrected for trading days, or at least corrected for trading days: in this work we focus on the latter version.

2.1 Estimation of the business indicators for France and Germany

Qualitative business surveys represent well known sources of information about the manufacturing sector. They provide early statements about variables which have quantitative counterparts (*e.g.* recent production trends) as well as information about variables which, although not directly observed (such as expectations), are widely used for assessing the current situation of the industrial sector. Their importance is self-explaining, as they are unique sources of information about businessmen's "mood". However, also qualitative variables which in principle have observable quantitative counterparts are increasingly considered by economists, whether because the empirical counterpart is not available at all, or because it is published with too much delay with respect to the reference period.

Business surveys are generally conducted asking businessmen particular questions (*e.g.* past production, future production, selling prices) which imply threefold answers such as "high", "low", "same". Quantification is generally obtained by taking the balance between the positive ("high") and negative ("low") percentage of answers.[1] The set of indicators obtained can then be used as a descriptive device about the current situation of the industrial sector as well as of its future prospects. Moreover, their features make them particularly suited to be used in models for forecasting industrial production and detecting turning points (see, for example, Bruno and Lupi, 2001; García-Ferrer and Bujosa-Brun, 2000; Kauppi *et al.*, 1996). In addition, synthetic indicators can be built, which summarize the information contained in the surveys, on the basis of the hypothesis that answers to

[1] Of course, there are other, more sophisticated, methods which, however, are seldom used in practice: see, among others, Batchelor (1981, 1986), Carlson an Parkin (1975), Foster and Gregory (1977), Pesaran (1987, especially chapter 8). A recent reference is Mitchell *et al.* (2002).

different questions are all influenced by a (some) common unobservable driving variable(s), as is done for example in Doz and Lenglart (1999) and European Commission (2000).

In this section we illustrate the estimation of such two synthetic indicators for France and Germany, building upon the business surveys data on the manufacturing sector produced by the European Commission (European Commission, 1997). The following balances are used in developing the indicators:

1. production trend observed in recent months;
2. assessment of order-book level;
3. assessment of export order-book level;
4. assessment of stocks of finished products;
5. production expectations for the months ahead;
6. selling price expectations for the months ahead.

All the data are provided by the European Commission seasonally adjusted. Balances are also standardized prior to modelling them.

The maintained hypothesis in the derivations below is that the balance related to each of the n questions (x_{it}) can be represented as being composed of two parts: a component common to all the questions (C_t) and another one proper of the particular question considered (u_{it}), such that

$$x_{it} = \gamma_i C_t + u_{it} \qquad i = 1, ..., n. \tag{1}$$

The model which will be used here is a *dynamic factor model*. It is a factor model in that the n variables can be adequately represented by one single factor; it is dynamic because the common factor as well as the specific (or idiosyncratic) component can be represented by dynamic stochastic processes. The γ_i are factor loadings specific for every question.

In particular, we specify the common component to follow a stochastic cycle:

$$\begin{bmatrix} C_t \\ C_t^* \end{bmatrix} = \rho \begin{bmatrix} \cos\lambda_c & \sin\lambda_c \\ -\sin\lambda_c & \cos\lambda_c \end{bmatrix} \begin{bmatrix} C_t \\ C_t^* \end{bmatrix} + \begin{bmatrix} \kappa_t \\ \kappa_t^* \end{bmatrix} \tag{2}$$

where λ_c is the frequency in radians, ρ is a damping factor bounded between 0 and 1, and κ_t, and κ_t^* are independent, zero mean, normal random variables with variance σ_κ^2 and $\sigma_{\kappa^*}^2$. C_t^* can be considered an auxiliary variable to our aims.

The specific component is allowed to be a first order autoregressive (stationary)

Table 1: Dynamic factor model: estimated coefficients (1985:01-2002:03)

	France	S.E.	Germany	S.E.
γ_1	0.164	0.011	0.132	0.010
γ_2	0.132	0.010	0.083	0.007
γ_3	0.150	0.013	0.070	0.009
γ_4	-0.134	0.015	-0.110	0.012
γ_5	0.151	0.013	0.137	0.012
γ_6	0.069	0.020	0.107	0.016
ϕ_1	0.176	0.121	-0.293	0.080
ϕ_2	0.962	0.021	0.987	0.011
ϕ_3	0.872	0.037	0.981	0.013
ϕ_4	0.694	0.056	0.935	0.026
ϕ_5	0.657	0.056	0.724	0.060
ϕ_6	0.931	0.022	0.920	0.027
$\ln \sigma_{\varepsilon_1}$	-1.680	0.073	-0.838	0.059
$\ln \sigma_{\varepsilon_2}$	-2.378	0.095	-2.323	0.064
$\ln \sigma_{\varepsilon_3}$	-1.749	0.059	-1.992	0.054
$\ln \sigma_{\varepsilon_4}$	-1.062	0.051	-1.776	0.059
$\ln \sigma_{\varepsilon_5}$	-1.175	0.052	-1.483	0.057
$\ln \sigma_{\varepsilon_6}$	-1.156	0.050	-1.365	0.053
ψ_ρ	3.978	0.101	3.834	0.078
ψ_λ	-5.625	1.659	-7.759	3.781
ρ	0.985		0.992	
λ	0.113		0.130	
$period\ (months)$	55.4		48.3	
Log-likelihood	-46.554		123.231	

As the parameters $\sigma^2_{\varepsilon_i}$, ρ and λ are bounded, in order to avoid solving a constrained optimization problem, transformed parameters $\ln \sigma_{\varepsilon_i}$, ψ_ρ and ψ_λ are used. In order to recover the original parameters of model composed by equations (2) and (3), the following back-transformations are needed: $\sigma^2_{\varepsilon_i} = \exp(2 \cdot \ln \sigma_{\varepsilon_i})$; $\rho = \frac{|\psi_\rho|}{\sqrt{1+\psi 2_\rho}}$; $\lambda = \frac{2\pi}{2+\exp(\psi_\lambda)}$.

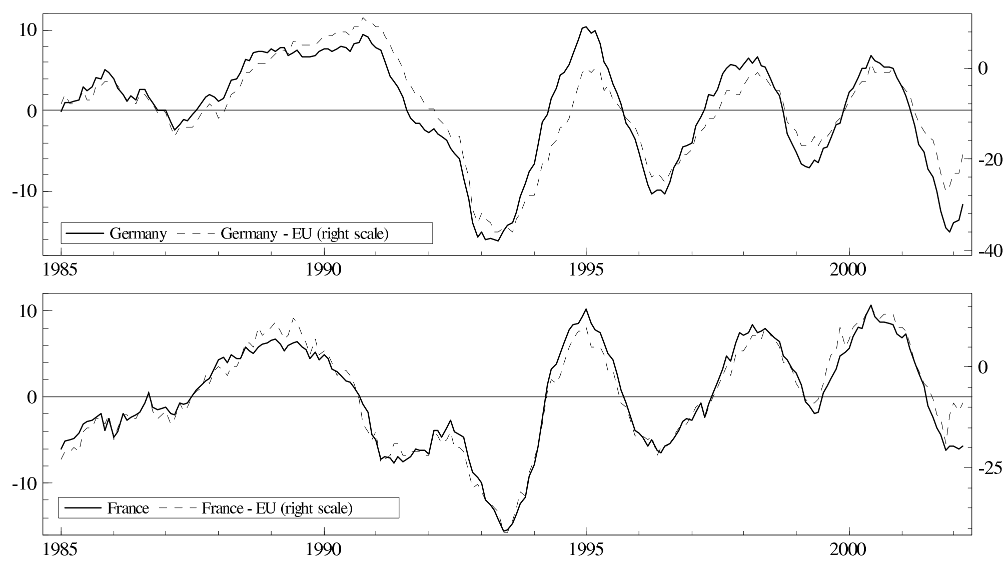
process:

$$u_{it} = \phi_i u_{i,t-1} + \varepsilon_{it} \tag{3}$$

The specification (2) is particularly suited to our case because it can accommodate various cyclical behaviours in a natural way; for example, when $\sigma^2_\kappa = 0$ the cycle is deterministic; when $\sigma^2_\kappa > 0$ and the damping factor $\rho = 1$, the cycle is nonstationary. In our case, as a further identifying assumption we fixed $\sigma^2_\kappa = \sigma^2_{\kappa^*}$ to be equal to 1, thus ruling out the deterministic case.

The reduced form for model in (2) is an ARMA(2,1); the restrictions in the parameter space of the AR polynomial needed to generate a cyclical behaviour of the common component are automatically enforced in the cyclical model. Following Harvey (1989) we see this as an advantage over the use of a reduced form, rather than a disadvantage.

Figure 1: Extracted common factors (solid lines) and European Commission's manufacturing climate indices (dashed lines).



The main results of the estimation process over the period 1985.1-2002.3 are listed in table 1.

Figure 1 displays the estimated common factors for Germany and France, together with the business climate indicators calculated by the European Commission for the same countries. The latter are calculated with a simple average of the assessments concerning the order-book level, the stocks of finished products (with sign changed) and the production expectations for the months ahead.

In what follows we will refer to the common components as the "business indicators".

2.2 The VAR for France and Germany

In order to forecast the industrial production indices for France and Germany, we propose building a single VAR for both countries that models jointly the industrial production indices and the business indicators of France and Germany. The approach is, to some extent, similar to the route followed in Buffeteau and Mora (2000) to predict Euro-zone GDP. Note that, strictly speaking, the use of the business indicators described in the previous sub-section gives rise to a "generated regressors" problem that might somewhat complicate the inference; however, we will not consider this aspect explicitly here.[2]

Throughout the rest of the paper $y_{g,t}$ with $g \in \{EUR, D, F, I\}$ indicates the working days adjusted industrial production index of country g (Euro-zone, Germany, France, Italy) as of time t; $C_{g,t}$ stands for the business indicator of country g at time t; τ_t are the number of working days in month t; $\Delta_\tau \equiv (1 - L^\tau)$ denotes the usual difference operator such that $\Delta_\tau z_t = z_t - z_{t-\tau}$.

Given that the business indicators are intended to sinthesyze information that refer to the "distance" from a "normal" situation,[3] they can be thought of as representing year-on-year variations of the overall business climate for the manufacturing sector. For this reason we model annual differences of the industrial production indices. This choice has the additional advantage of largely solving seasonality problems.[4] In order to allow for the existence of any residual seasonality (or for over-correction induced by seasonal differences), an overparameterized VAR with 12 lags and seasonal dummies is initially specified: furthermore, in order to consider possible residual calendar effects, the deterministic part of the VAR includes, beside the constant (c) and the seasonal dummies (s_l), also a term for trading days (τ_t). More precisely, denoting by $\mathbf{x}_t = (\Delta_{12} \log y_{D,t}, \Delta_{12} \log y_{F,t}, C_{D,t}, C_{F,t})'$, the unrestricted VAR is

$$\mathbf{x}_t = \mathbf{c} + \sum_{j=1}^{12} \boldsymbol{\pi}'_j \mathbf{x}_{t-j} + \sum_{j=0}^{1} \vartheta_j \Delta_{12} \tau_{t-j} + \sum_{l=1}^{11} \kappa_l s_l + \boldsymbol{\nu}_t. \tag{4}$$

The model is estimated over the period 1992:01-2000:09, leaving 18 observations (up to 2002:03) for out of sample forecast evaluation.[5] The final (parsimonious)

[2] Note that this problem is *a fortiori* present in all those models that use national accounts quarterly data (see e.g. Cainelli and Lupi 1999; Bruno et al., 2002; Eurostat, 1997).

[3] This depends on the very nature of the business surveys.

[4] See Osborn (2001) on this aspect.

[5] The analysis has been carried out using PcGive 10.0 (Doornik and Hendry, 2001).

Table 2: **France and Germany: VAR diagnostics**

Equation	Test			p-value
$\Delta_{12} \log y_D$	AR(1-12)	F(12,75)	1.2194	0.2861
$\Delta_{12} \log y_F$	AR(1-12)	F(12,75)	1.2971	0.2383
C_D	AR(1-12)	F(12,75)	0.7476	0.7008
C_F	AR(1-12)	F(12,75)	1.1775	0.3147
$\Delta_{12} \log y_D$	Normality	$\chi^2(2)$	1.1746	0.5558
$\Delta_{12} \log y_F$	Normality	$\chi^2(2)$	3.3506	0.1873
C_D	Normality	$\chi^2(2)$	2.0633	0.3564
C_F	Normality	$\chi^2(2)$	5.0902	0.0785
$\Delta_{12} \log y_D$	ARCH(1-12)	F(12,63)	0.3775	0.9669
$\Delta_{12} \log y_F$	ARCH(1-12)	F(12,63)	1.4500	0.1678
C_D	ARCH(1-12)	F(12,63)	0.9819	0.4757
C_F	ARCH(1-12)	F(12,63)	0.6739	0.7697
$\Delta_{12} \log y_D$	Hetero	F(34,52)	0.8457	0.6946
$\Delta_{12} \log y_F$	Hetero	F(34,52)	0.5515	0.9659
C_D	Hetero	F(34,52)	0.8532	0.6848
C_F	Hetero	F(34,52)	0.7488	0.8131
VAR	AR(1-12)	F(192,146)	0.9436	0.6484
VAR	Normality	$\chi^2(8)$	14.0740	0.0799
VAR	Hetero	F(340,441)	0.6633	1.0000

AR(1-12) is the LM-test for residuals autocorrelation up to order 12. Normality is the Doornik-Hansen (1994) test for residuals normality. ARCH(1-12) is the LM-test for residuals autoregressive conditional heteroscedasticity up to order 12. Hetero is the White (1980) test for unconditional heteroscedasticity of the residuals.

parameterization is selected on the basis of sequential simplification, and includes lags 1 to 3, lag 6, no seasonal dummies, one period lagged trading days term. The p-value of the restrictions implied by this reduction is 0.832, suggesting the validity of the reduction. The main diagnostics of the restricted VAR are reported in table 2.

In order to evaluate the validity of this VAR as a forecasting tool, a historical simulation in which the business indicators are substituted by their concurrent estimates for each month in the forecasting sample is required. This procedure exactly replicates a "real world" situation in which the model and the indicators are re-estimated each month on the base of the available information. The results reported in table 3 indicate that the forecasts produced in this way are adequate up to six months ahead for France, while for Germany they seem to be reliable enough only up to three steps ahead.

Table 3: Industrial production indices: forecast errors statistics using concurrent estimates of the business indicators (2000:10-2002:03)

	ME	MAE	RMSE
		1-step	
France	-0.0001	0.0120	0.0170
Germany	-0.0012	0.0116	0.0153
		3-step	
France	-0.0027	0.0165	0.0221
Germany	-0.0028	0.0175	0.0215
		6-step	
France	-0.0114	0.0160	0.0209
Germany	-0.0074	0.0256	0.0331

The forecast errors statitics refer to $\Delta_{12} \log y_{i,t}$ ($i \in \{D, F\}$). ME is the mean error; MAE is the mean absolute error; RMSE is the root mean square error.

2.3 Italy

In order to forecast the Italian industrial production index, we use a quite different approach, which is described in detail in Bruno and Lupi (2001). The variable to be forecast in this case is raw industrial production index (sections C to E). Predictions are derived using a trivariate VAR in (log) industrial production index, business surveys production prospects, and (log) tons of goods transported by railways. The second variable represents the entrepreneurs' sentiment about future production; the third one is a leading variable of industrial production, since goods transported by railways in Italy are mainly intermediate goods and raw materials. The stochastic part of the VAR is complemented with an accurate parameterization of the deterministic components (see Bruno and Lupi, 2001). We decided to rely on this model rather than estimating a new one on trading days adjusted data, given the good performance that we are able to obtain. Indeed, in Bruno and Lupi (2001) we show that our VAR predictions outperform those obtained on the basis of a robust ARIMA model used as a benchmark, are on average at least as good as the survey-based projections elaborated by CSC (the research department of the Confederation of Italian Industry), and more accurate than those deriving from the IRS (Istituto per la Ricerca Sociale) econometric model. The properties of the forecasts over the period 2000:10-2002:03 are reported in table 4.

Table 4: Italian industrial production index: forecast errors statistics (2000:10-2002:03)

ME	MAE	RMSE
	1-step	
-0.0014	0.0114	0.0138
	3-step	
-0.0045	0.0131	0.0155
	6-step	
-0.0171	0.0196	0.0228

The forecast errors statitics refer to $\Delta_{12} \log y_{I,t}$. ME is the mean error; MAE is the mean absolute error; RMSE is the root mean square error.

Table 5: Estimates of parameters of model (4)

Parameter	Value	S.E.
α_0	0.0002	0.0001
α_1	1.0017	0.0033
α_2	-0.8130	0.0037
θ	-0.7069	0.0473
Eq. Standard Error	0.0023	

In order to produce forecasts that are consistent with those derived for France and Germany, the raw series augmented with the forecasts is finally adjusted for calendar effects. The official Istat series is adjusted for trading days on a very disaggregate basis, using a proportional approach (Istat, 1996). Because of the disaggregation required, we are not able to reproduce this procedure: instead, in this paper we use an approximation which, however, proved to work quite well. A bridge equation with MA residuals is estimated which links the working days adjusted figures ($y_{I,t}$) to the raw index ($\tilde{y}_{I,t}$) and to trading days (τ_t):

$$\Delta_{12} \log y_{I,t} = \alpha_0 + \alpha_1 \Delta_{12} \log \tilde{y}_{I,t} + \alpha_2 \Delta_{12} \log \tau_t + \left(1 + \theta L^{12}\right) \varepsilon_t. \qquad (5)$$

Again, a real time exercise has been carried out estimating (5) for every time period in the evaluation sample. The estimates of the parameters in (5) are listed in table 5.

3. AGGREGATING THE NATIONAL FORECASTS

The aggregate trading days corrected industrial production for the Euro area is obtained by Eurostat by aggregating the national indices with a fixed weighting scheme of the kind

$$y_{EUR,t} = \sum_{i=1}^{12} \delta_i y_{i,t} \qquad (6)$$

where $y_{i,t}$ is the industrial production index for the i-th country and δ_i is the attached weight (with $\sum_{i=1}^{12} \delta_i = 1$). The aggregate seasonally adjusted figure is then obtained applying the adjustment procedure TRAMO-SEATS (Gómez and Maravall, 1998) to the aggregate trading days corrected Euro-zone index.

Had we available forecasts for all the Euro-zone countries, (6) would be the most obvious way to get the working-days adjusted aggregate index. Nevertheless, even if the three countries we consider here account for most of the production realized in the Euro-zone, in order to get $y_{EUR,t}$ we have to modify model (6) in such a way to deal with the indices of the remaining countries. A fairly general unobserved components model offers a natural framework to accomplish this task:

$$\begin{cases} y_{EUR,t} = \alpha_1 y_{D,t} + \alpha_2 y_{F,t} + \alpha_3 y_{I,t} + \mu_t + \beta_t \tau_t + \sum_{j=1}^{6} \gamma_{jt} + \varepsilon_t \\ \mu_t = \mu_{t-1} + \eta_t \\ \beta_t = \beta_{t-1} + \xi_t \\ \gamma_{jt} = \gamma_{j,t-1} \cos \lambda_j + \gamma^*_{j,t-1} \sin \lambda_j + \omega_{jt} \\ \gamma^*_{jt} = -\gamma_{j,t-1} \sin \lambda_j + \gamma^*_{j,t-1} \cos \lambda_j + \omega^*_{jt} \end{cases} \qquad (7)$$

with $\varepsilon_t \sim \mathsf{IID}(0, \sigma^2_\varepsilon)$, $\eta_t \sim \mathsf{IID}(0, \sigma^2_\eta)$, $\xi_t \sim \mathsf{IID}(0, \sigma^2_\xi)$, $\omega_{jt} \sim \mathsf{IID}(0, \sigma^2_\omega)$, $\omega^*_{jt} \sim \mathsf{IID}(0, \sigma^2_\omega)$, and $\lambda_j = 2\pi j/12$. The term μ_t represents the unobserved non-seasonal component of the industrial production due to the excluded countries; τ_t stands for the number of trading days; $\sum_{j=1}^{6} \gamma_{jt}$ models the (trigonometric) seasonality of the series.

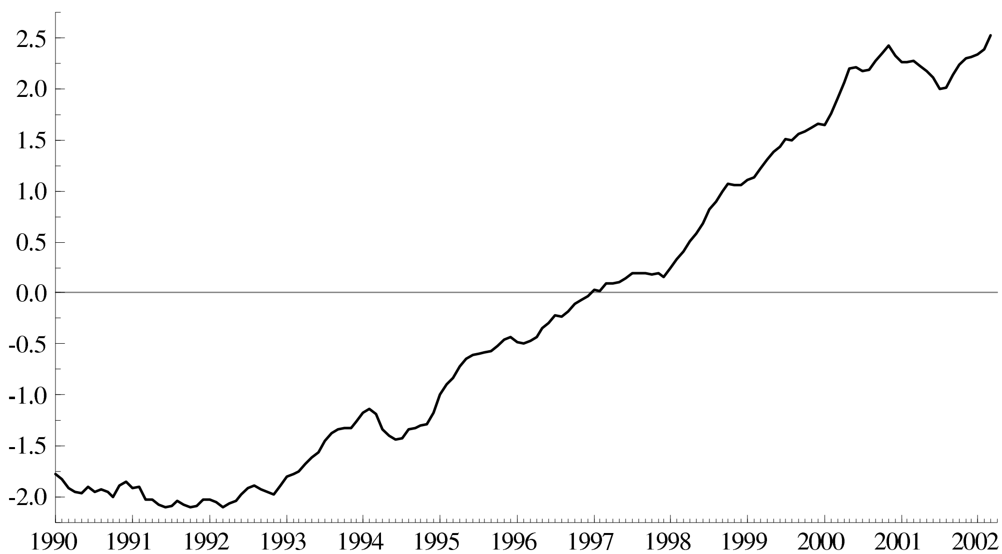
The general model (7) is simplified by imposing $\beta_t = 0$ and $\sigma^2_\omega = 0$: indeed cal-

endar effects are not significant in (7), while seasonality is well reproduced using a deterministic representation. The main estimation results are reported in table 6.
Figure 2 graphs the estimate of μ_t which, except for a scale factor, loosely repre-

Table 6: Estimates of parameters of model (6) (1985:01-2002:03)

Parameter	Value	S.E.
α_1	0.419	0.0003
α_2	0.348	0.0006
α_3	0.240	0.0002
$\ln \sigma_\eta$	-2.012	0.0138
$\ln \sigma_\varepsilon$	-1.344	0.0124

Figure 2: Estimate of μ_t.



sents the weight of the 9 excluded European countries in the Euro-zone industrial production index: as expected, this weight has been increasing, in particular over the period 1992-2000. The simple fixed-weight aggregation (that could have been carried out also by regression methods, as in Zizza, 2002), would result in poorer forecasts.

We are now in the position of aggregating the forecasts obtained along the lines exposed in sections 2.2 and 2.3 and the results in terms of forecast errors statistics are reported in table 7. The aggregate forecasts appear to be fairly informative even over the six-step ahead horizon, the mean absolute error being about 1.5%. We also produce two alternative forecasts. The first using an ARMA model on

Table 7: Euro-zone industrial production index: forecast errors statistics (2000:10-2002:03)

	ME	MAE	RMSE
		1-step	
Multi-country	-0.0006	0.0088	0.0111
Direct VAR	-0.0061	0.0134	0.0156
Direct-ARMA	0.0142	0.0165	0.0202
		3-step	
Multi-country	-0.0026	0.0112	0.0149
Direct VAR	-0.0070	0.0163	0.0206
Direct-ARMA	0.0108	0.0194	0.0221
		6-step	
Multi-country	-0.0103	0.0150	0.0194
Direct VAR	-0.0164	0.0210	0.0239
Direct-ARMA	0.0027	0.0126	0.0166

The forecast errors statitics refer to $\Delta_{12} \log y_{EUR,t}$. "Multi-country" refers to the forecasts derived by aggregating the predictions for France, Germany, and Italy; "Direct VAR" refers to the forecasts derived from applying a VAR directly to the euro-zone IPI and business indicator; "Direct-ARMA" refers to the forecasts derived from direct application of an ARMA model to the Euro-zone data. ME is the mean error; MAE is the mean absolute error; RMSE is the root mean square error.

$\Delta_{12} \log y_{EUR,t}$. An ARMA(4,4) model is selected on the basis of the BIC criterion over the same sample used for identifying the VAR. The ARMA model is then sequentially re-estimated by MLE and one to six-step ahead forecasts are produced over the period 2000:10-2002:03 to be compared with those obtained by applying the procedure described before. The second set of alternative forecasts are obtained from a VAR model between the Euro-zone IPI and a Euro-zone business indicator calculated in the same way as those of Germany and France. Both The ARMA forecasts ("Direct-ARMA")and the VAR model ones ("Direc-VAR") are reported in table 7. The comparison between the three sets of forecasts shows that our procedure produce predictions that are on average more accurate than those stemming from the ARMA model. However, while the MAE and the RMSE of our forecasts are monotonically increasing over the forecasting horizon, those of the ARMA decrease drastically in correspondence with the six-step ahead horizon; it should be said, however, that this effect looks very much like a spurious one tied to the small forecasts sample.Moreover, the aggregation of national forecasts appears superior, in this context, to the aggregate approach, at all forecasting horizons.

The model of the working days adjusted index as illustrated in the previous section has revealed successful in improving the forecast errors with respect to an ARIMA model used as a benchmark. Of course, it is interesting to forecast not only the annual variations of the raw index, but also the month-to-month growth of the seasonally adjusted figures. The circumstance that, in order to produce the seasonally adjusted Euro-zone index, Eurostat adjusts the Euro-area aggregate index directly allow us to reproduce the same procedure as the one described in the previous sections adding at the end of every vintage of the working days adjusted index either our model forecasts or the ARIMA forecasts. Seasonal adjustment is then carried out with TRAMO-SEATS. Forecasting errors are compared with reference to the latest seasonally adjusted figures released by Eurostat.[6] (See. table 8.)

[6] The multi-country procedure has been followed also by specifying appropriate models directly on seasonally adjusted data. However, the forecasting performance is significantly worse than the one we can obtain from using non seasonally adjusted data, particularly so at three and six steps ahead: in particular, the three-step ahead MAE is about 30% bigger and the six-step ahead MAE is about 60% bigger when using seasonally adjusted rather than raw data.

Table 8: Euro-zone seasonally adjusted industrial production index:forecast errors statistics (2000:10-2002:03)

	ME	MAE	RMSE
		1-step	
Multi-country	-0.0009	0.0072	0.0084
Direct VAR	-0.0056	0.0098	0.0119
Direct-ARMA	0.0117	0.0140	0.0168
		3-step	
Multi-country	-0.0029	0.0095	0.0126
Direct VAR	-0.0069	0.0137	0.0172
Direct-ARMA	0.0093	0.0169	0.0190
		6-step	
Multi-country	-0.0110	0.0132	0.0171
Direct VAR	-0.0171	0.0193	0.0224
Direct-ARMA	0.0022	0.0117	0.0141

The forecast errors statitics refer to $\Delta_{12} \log y_{EUR,t}$. "Multi-country" refers to the forecasts derived by aggregating the predictions for France, Germany, and Italy; "Direct VAR" refers to the forecasts derived from applying a VAR directly to the euro-zone IPI and business indicator; "Direct-ARMA" refers to the forecasts derived from direct application of an ARMA model to the Euro-zone data. ME is the mean error; MAE is the mean absolute error; RMSE is the root mean square error.

4. CONCLUDING REMARKS

In this paper we propose a relatively simple procedure to predict Euro-zone industrial production using mostly data derived from the business surveys of the three major economies within the European Monetary Union. The basic idea is that of estimating business cyclical indicators to be used as predictors for the industrial production in France and Germany; as far as Italy is concerned, forecasts are produced using a model that in the recent past proved to be able to furnish accurate forecasts up to six months ahead. In order to derive quantitative predictors from the business surveys data and to aggregate the nation-wide forecast into the Euro-zone forecast, we propose using an approach based on dynamic factors and unobserved components models.

The overall results show fairly good forecasts performance up to six steps ahead.

REFERENCES

Baffigi, A., Golinelli, R. & Parigi, G. (2002). Real-time GDP forecasting in the Euro area, Bank of Italy, mimeo.

Ballabriga, F.C. & Castillo, S. (2000). BBVA-ARIES: a forecasting and simulation model for the EMU economy, mimeo.

Batchelor, R.A. (1981). Aggregate expectations under the stable laws, *Journal of Econometrics*, 2, 199-210.

Batchelor, R.A. (1986). Qualitative vs quantitative measures of inflation expectations, *Oxford Bulletin of Economics and Statistics*, 2, 99-120.

Bergström, R. (1995). The relationship between manufacturing production and different business survey series in Sweden 1968–1992, *International Journal of Forecasting,* 11, 379-393.

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

Buffeteau, S. & Mora, V. (2000). Predicting the national accounts of the euro zone using business surveys, *Insee Conjoncture*, December 2000, 10-18.

Brunello, G., Lupi, C. & Ordine, P. (2000). Regional disparities and the Italian NAIRU, *Oxford Economic Papers*, 52, 146-177.

Bruno, G., Cubadda, G., Giovannini, E. & Lupi, C. (2002). The flash estimate of the Italian real gross domestic product, in R.Barcellan and G.L.Mazzi (Eds.), *Workshop on Quarterly National Accounts*. Luxembourg, Eurostat.

Bruno, G. & Lupi, C. (2001). Forecasting industrial production and the early detection of turning points, ISAE Documenti di Lavoro n. 20/01: improved and updated version presented at the 22nd International Symposium on Forecasting (ISF), Dublin, June 2002.

Cainelli, G. & Lupi, C. (1999). Aggregazione contemporanea e specificazione econometrica nella stima trimestrale dei conti economici nazionali, *Statistica*, 49, 239-267.

Carlson, J.A. and Parkin, J.M. (1975). Inflation expectations, *Economica*, 123-138.

Doornik, J.A. & Hansen, H. (1994). A practical test for univariate and multivariate normality, Nuffield College Discussion Paper

Doornik, J.A. & Hendry, D.F. (2001). *Modelling Dynamic Systems Using PcGive, Vol II.* London: Timberlake Consultants Ltd.

Doz, C. & F. Lenglart (1999). Analyse factorielle dynamique: test du nombre de facteurs, estimation et application à l'enquête de conjoncture dans l'industrie. *Annales d'économie et de statistique*, 54, 91-127.

Eurostat (1997). *Handbook on Quarterly National Accounts.* Luxembourg, Eurostat.

European Commission (1997). The joint harmonised EU programme of business and consumer surveys, *European Economy*, 6.

European Commission (2000). Business Climate Indicator for the Euro Area, (`http://europa.eu.int/comm/economy_finance/indicators/business_climate/2001/presentation_climate.pdf`).

Foster, J. and Gregory, M. (1977). Inflation expectations: The use of qualitative survey data, *Applied Economics*, 9, 319-329.

García-Ferrer, A. & Bujosa-Brun, M. (2000). Forecasting OECD industrial turning points using unobserved components models with business survey data, *International Journal of Forecasting*, 16, 207-227.

Gómez, V., & Maravall, A. (1998). *Programs TRAMO and SEATS: Instructions for the User,* Madrid: Banco de España. (`http://www.bde.es/servicio/software/econome.htm`).

Harvey, A.C. (1989). *Forecasting, Structural Time Series Models and the Kalman Filter.* Cambridge: Cambridge University Press.

Huh, C. (1998). Forecasting industrial production using models with business cycle asymmetry, *Federal Reserve Bank of San Francisco Economic Review*, n. 1, 29-41.

Istat (1996). "Numeri indici della produzione industriale", *Metodi e Norme*.

Kauppi, E., Lassila, J. & Teräsvirta, T. (1996). Short-term forecasting of industrial production with business survey data: experience from Finland's great depression 1990–1993, *International Journal of Forecasting,* 12, 373-381.

Marcellino, M., Stock, J.H. & Watson, M.W. (2001). Macroeconomic forecasting in the Euro Area: country specific versus area-wide information, IGIER Working Paper n. 201.

Marchetti, D.J. & 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, 419-440.

Mitchell, J., Smith, R.J. & Weale, M.R. (2002). Quantification of qualitative firm-level survey data, *Economic Journal*, 112, C117-C135.

Pesaran, M.H. (1987). *The Limits to Rational Expectations*. Oxford, Basil Blackwell.

Osborn, D.R. (2001). Unit-Root versus deterministic representations of seasonality for forecasting, Ch. 18 in Clements M.P. & Hendry, D.F. (Eds.) *A Companion to Economic Forecasting*. Oxford: Basil Blackwell.

Osborn, D.R., Heravi, S. & Birchenhall, C.R. (1999). Seasonal unit roots and forecasts of two-digit European industrial production, *International Journal of Forecasting*, 15, 27-47.

Simpson, P.W., Osborn, D.R. & Sensier, M. (2001). Forecasting UK Industrial Production over the Business Cycle, *Journal of Forecasting*, 20, 405-424.

White, H. (1980). A heteroskedastic-consistent covariance matrix estimator and a direct test for heteroskedasticity, *Econometrica*, 48, 817-838.

Zizza, R. (2002): Forecasting the industrial production index for the Euro area through forecasts for the main countries, Bank of Italy Temi di Discussione No. 411.