FORECASTING MONTHLY INDUSTRIAL PRODUCTION IN REAL-TIME: FROM SINGLE EQUATIONS TO FACTOR-BASED MODELS

by Guido Bulligan*, Roberto Golinelli and Giuseppe Parigi***

Abstract

The aim of this paper is to analyze the performance of alternative forecasting methods to predict the index of industrial production in Italy from one- to three-months ahead. We use twelve different models, from simple ARIMA to dynamic factor models exploiting the timely information of up to 110 short-term indicators, both qualitative and quantitative. This allows to assess the relevance for the forecasting practice of alternative combinations of types of data (real-time and latest available), estimation methods and periods. Out-of-sample predictive ability tests stress the relevance of more indicators in disaggregate models over sample periods covering a complete business cycle (about 7 years in Italy). Our findings downgrade the emphasis on both the estimation method and data revision issues. In line with the classical "average puzzle", the use of simple averages of alternative forecasts often improves the predictive ability of their single components, mainly over short horizons. Finally, bridge and factor-based models always perform significantly better than ARIMA specifications, suggesting that the short-run indicator signal always dominates the noise component. On this regard, bridge models can further increase the amount of signal extracted to improve up to 30-40% the short-run predictive ability of factor-based models and to forecast-encompass them.

Index

Version: April, 21st 2008

JEL classification: C22, C53, E01, E37

Keywords: real-time forecasting; real-time data sets; aggregate, disaggregate and factor models, combining alternative forecasts.

* Research Department, Bank of Italy (guido.bulligan@bancaditalia.it and giuseppe.parigi@bancaditalia.it)

** Department of Economics, University of Bologna, Italy (roberto.golinelli@unibo.it)

1. Introduction[1]

One of the main tasks of the economy watcher is to extract reliable signals from high frequency indicators and to use them to provide the decision-maker with early pictures of the state of the economic situation in the short-run.

The number of indicators on the evolution of economic activity has gradually become very large, spanning from quantitative, or hard, variables (such as the industrial production index) to qualitative, or soft, ones (such as survey data). Although nowadays the statistical agencies provide each month plenty of high frequency indicators, their use entails a trade-off between timeliness and statistical noise: quickly available data may provide signals of difficult interpretation.

The index of industrial production (IPI, henceforth) is probably the most important and widely analyzed high-frequency indicator, given the relevance of the manufacturing activity as a driver of the whole business cycle. As soon as the index is published, extensive comments and reactions of the business analysts witness its relevance. Indeed, the IPI is one of the most important indicators used to forecast the short-run evolution of the GDP in most countries (see for instance Golinelli and Parigi, 2007, for an application to the G7 countries).

However, the IPI itself is characterized by a significant publication delay, which limits its usefulness and motivates the great efforts to compute reliable and updated forecasts.[2] Shortening the delay of the first release is possible but at the expense of a greater degree of revision of the early estimates. This leads to the usual problem of assessing the ability of alternative forecasting methods using real-time data (see e.g. Croushore and Stark, 2001, and Stark and Croushore, 2002; Diebold and Rudebusch, 1991).

The aim of this paper is to explore the real-time performance of alternative ways to forecast the monthly dynamics of the Italian IPI, i.e. different "*forecasting methods, which include the models as well as the estimation procedures and the possible choices of estimation*

---

[1] Paper presented at Euro Area Business Cycle Network workshop: "Using Euro Area Data: Issues and Implications for Economic Analysis", and at the 28th Annual International Symposium on Forecasting of the International Institute of Forecasters. We are grateful to Matteo Ciccarelli, Marco Lippi, Luigi Federico Signorini and workshop and conference participants for helpful comments and suggestions. The views contained here are those of the authors only and do not necessarily reflect those of the institutions for which they work. PRIN financing is grateful acknowledged (R. Golinelli).

[2] Following the legislation introduced in 1998, the countries in the euro area are required to deliver IPI data with a delay no longer than 45 days, necessary to collect information from a large number of production plants (see Ladiray and O'Brien, 2003).

windows" (Giacomini and White, 2006, p. 1549). Our forecasting methods are defined through combinations of the following three sets of options.

First, the degree of model complexity. If randomness is the main feature of the indicator information content, simpler models may be more suitable. On the other hand, complex models are to be preferred to reduce the noise due to the partial information of each indicator. In this case, two options are available: (*i*) disaggregate models, entailing forecasting errors that might compensate at a more aggregate level (see e.g. Hendry and Hubrich, 2007); (*ii*) factor-based models, where a few predictors summarise the information content of a large number of indicators (the so-called common factors of the information set; see Stock and Watson, 2006).

Second, the estimation method. We apply both the ordinary least squares (OLS) and, given the disaggregate nature of the models sub (*i*), the seemingly unrelated regression (SUR) procedures, in the attempt to obtain more efficient parameter estimates by accounting for possible simultaneity of the random shocks to different equations. In the context of the factor models sub (*ii*) the choice is between static and dynamic principal components and different ways to select the appropriate number of factors.

Third, the length of the estimation window. The choice of too wide a window may affect the precision of the estimates (and so of the forecasts) because of the likely occurrence of breaks (see e.g. Stock and Watson, 1996, and Clements and Hendry, 2002). As there is no *a priori* indication about the appropriate length of the estimation window we consider three cases: a window of seven years (chosen according to the average length of the Italian industrial cycle), a shorter window of four years and a longer one of more than ten years.

Finally, we assess the relevance of data revisions by comparing the forecasting performance of models, estimation methods and samples with both latest-available and real-time data (see Kozicki, 2002, for a discussion on the consequences of data revisions).

The models used in this paper are listed in Table 1. The order with which they have been reported follows the historical development of the Italian IPI short-term forecasting analysis characterised by an increasing number of indicators. The *old single bridge models* (OSBM)[3] were specified as single equations with only a few indicators: electricity consumption, temperatures and a trend (possibly nonlinear). These specifications were subsequently updated with additional indicators, thanks to their growing availability, resulting in *updated single bridge models* (USBM) and, more recently, the *final single bridge models* (FSBM). All these

[3] The precise meaning of "bridge model" will be given in Section 2. At this stage, "bridge" qualifies models embodying the researcher's specific indicator selections from a large data set.

models are estimated as they were originally proposed (details about the different specifications are reported in the Appendix) to show the changes in their forecasting performance due to the different data sets and time spans.

Table 1

THE MODELS USED IN THIS PAPER [1]

	A. Univariate models
ARIMA	Univariate ARIMA model
	B. Dynamic single-equation models with few indicators
OSBM	Old single bridge model based on electricity consumption, temperature and trend
USBM	Updated single bridge model (with just few more indicators than OSBM)
FSBM	Final single bridge model (with additional indicators with respect to USBM)
	C. Dynamic multiple-equation models disaggregated in subsectors [2]
ASGR	Aggregation of sectoral annual growth rates
ASLWE	Aggregation of sectoral levels excluding the energy sector
ASWL	Aggregation of sectoral levels using the official weights of the statistical agency
	D. Unstructured empirical indicator approaches
FA-ARDL	Average of bivariate autoregressive distributed lags model forecasts [3]
SW-D	Direct h-step (multistep) forecasts of the static factor model
SW-S	Sequential one-step forecasts of the static factor model
FHLR-F	Generalised dynamic factor model with fixed rule to determine the factors number
FHLR-O	Generalised dynamic factor model with optimal criteria to determine the factors number

([1]) Further details are in the Appendix. ([2]) The sectoral forecasts of these models are aggregated to obtain the IPI. ([3]) Each ARDL forecasting model uses the information of only one indicator of the whole data set.

Thanks to the availability of timely information the IPI modelling issue may be tackled at a more disaggregate level, by specifying different equations for different manufacturing sub-sectors, such as those producing consumption, equipment, intermediate and energy goods (sectoral bridge models). The aggregate IPI predictions are then obtained with three different “aggregator” functions (for the details see the Appendix).

Alternatively, the large number of timely time series can be exploited in forecasting through an “unstructured empirical indicator approach” (Klein and Sojo, 1989), where no accounting/economic relationship is postulated between the indicators and the variable to be forecast. We apply two procedures: *(i)* simple averages of n (where n = 110 is the total number of indicators in our data set, see the Appendix) IPI forecasts from n bivariate autoregressive distributed lags (ARDL) models; and *(ii)* approximate common factor models, both static (Stock and Watson, 2002a and 2002b) and dynamic (Forni *et al.*, 2000 and 2005).

The paper is organised as follows. In Section 2 we classify the models in Table 1 according to the quantity of information used. Section 3 illustrates the main methodological issues in modelling and forecasting with indicators the Italian IPI in real-time. In order to list pros and cons of alternative IPI forecasting methods, we generalise the decomposition of the forecast error to account for both real-time data and the timing of the indicator updates. Section 4 is devoted to the analysis of the out-of-sample performance (from one- to three-months ahead) of the models *sub* A, B and C in Table 1 in predicting the IPI. Section 5 deals with the IPI forecasting ability over the same horizons of the unstructured empirical indicator models listed at point D of Table 1. Section 6 discusses and summarises main findings.

2. Alternative modelling approaches exploiting information sets of different size

The problem of extracting reliable signals from high frequency indicators is not new. Klein and Sojo (1989) suggest two alternative ways to classify the literature.

The “bridge model” approach (BM henceforth; see points B and C in Table 1) estimates the dynamic relationship between the IPI (either aggregate or disaggregate) and a number of pre-selected indicators. In this analysis both the IPI and the indicators are characterised by the same monthly frequency and the “bridge” qualification refers to the link between the dependent variable and the host of indicators used for forecasting.[4] The out-of-sample bridge forecasts are obtained by filling right-hand side explanatory indicators with their values (if known), or with indicator extrapolations when such data are not available. With disaggregate BM, the aggregate IPI forecasts are obtained by aggregating (in alternative ways) the predictions for each sub-sector.

The “empirical indicator model” approach (point D in Table 1) is analyzed in two alternative ways. With the first approach each of the n indicators in the data set is used in a bivariate ARDL regression and the IPI forecast is obtained as the average of the n forecasts. With the second, the IPI forecasts are computed through the principal components of the n indicators in the data set (factor-based models, FM henceforth). The main advantage of this approach is to exploit not only the information content of the single variables but their

[4] The BM approach has been mainly used to forecast GDP and the term “bridge” means the link between the quarterly GDP and the indicators, which are generally available at higher frequency; see Trehan (1989, 1992), Parigi and Schlitzer (1995), Kitchen and Monaco (2003), Sedillot and Pain (2003), Runstler and Sedillot (2003), Baffigi *et al.* (2004), Klein and Ozmucur (2005) and Golinelli and Parigi (2007, 2008).

covariance as well, without incurring in the "curse of dimensionality" as in unrestricted vector autoregressive models (see Stock and Watson, 2006, for an updated survey).[5]

It is important to realise that in the BM approach the choice of the most suitable indicators depends on a lot of statistical testing procedures and on the researcher's skill and experience (see Golinelli and Parigi, 2007), while FM automatically weights each indicator according to their signal-to-noise ratio[6]. In other words, BM allows more flexibility in the specification strategy at the expense of lower automation (see Golinelli and Parigi, 2008 for early attempts to automate BM specifications). In this context both approaches (bridge and empirical indicator models) require a number of modelling settings (e.g. about the estimation method or the sample size) which, in turn, imply alternative choices with different effects on the forecast error.

Let y_t^v denote the v^{th} vintage of the IPI, i.e. the time series to be forecast, where the time period t = 1, 2, ..., $T+v$ is measured in months (the sub-period from 1 to T is common to all the vintages).[7] Let h denote the forecast horizon (h = 1, 2, 3 in our case) and I_t^v the information set to forecast the IPI vintage v given by indicators and lags available at time t.

The h-step ahead forecast of the IPI vintage v, obtained from the use of an estimated model, is defined as $mod_{T+v+h}^{v}(I_{T+v+m}^{v+m}, \hat{\vartheta}^{v})$, where $\hat{\vartheta}^{v}$ is the vector of the estimated parameters of the specification $mod_{T+v+h}^{v}(.)$; $T+v$ is the starting point of the forecast (the last observation of the vintage v); the index m means that some monthly indicator observations belong to the IPI forecasting horizon, as they are are available earlier than the IPI. When $h \leq m$, all indicator observations are already known (the so-called nowcast case); when $h > m$, the indicator observations have to be forecast from h to m through auxiliary models.[8]

$\hat{\varepsilon}_{T+v+h}^{v+h}$ (the forecast error of the h-step ahead forecast of y_{T+v+h}^{v+h}) can be decomposed into three components in order to explain the role played by the modelling choices to account for

[5] Though originally used to extrapolate cyclical conditions (see Zarnowitz, 1992, and Altissimo *et al.*, 2001), FM may also be used to forecast single variables, such as the GDP or the inflation rate (see Marcellino *et al.*, 2003, Cristadoro *et al.*, 2005, Schumacher and Breitung, 2006, Altissimo *et al.*, 2007, and Cristadoro *et al.*, 2007).

[6] The claim of neutrality and generality of factor-based models is questioned by Boivin and Ng (2005 and 2006), who stress the relevance for the assessment of the model forecasting ability of both the size/composition of the data set and the way the factor-based forecasts are formulated.

[7] Each vintage v gives the modeller a snapshot of the data available at $T+v$.

[8] With ARMA or VAR models the IPI forecasts are always "pure forecast", as these models exploit only lagged information (i.e. m = 0 by definition).

both the real-time nature of data and the partial availability of monthly indicators over the forecasting horizon:

$$\begin{aligned}\hat{\varepsilon}^{v+h}_{T+v+h} &= [y^{v+h}_{T+v+h} - E(y^{v}_{T+v+h} \mid I^{v+m}_{T+v+m})] + [E(y^{v}_{T+v+h} \mid I^{v+m}_{T+v+m}) - mod^{v}_{T+v+h}(I^{v+m}_{T+v+m}, \vartheta^{v})] + \\ &+ [mod^{v}_{T+v+h}(I^{v+m}_{T+v+m}, \vartheta^{v}) - mod^{v}_{T+v+h}(I^{v+m}_{T+v+m}, \hat{\vartheta}^{v})]\end{aligned} \tag{1}$$

The first component in squared brackets on the right-hand side of (1) – i.e. the deviation between the actual realisation of the future vintage $v+h$ and the (unknown) conditional expectation function for vintage v – includes idiosyncratic elements, such as future random shocks and data revisions, that cannot be forecast.

The second component arises from the use of $mod^{v}_{T+v+h}(.)$ to approximate the unknown conditional expectation function. This misspecification bias may be reduced by modelling the conditional expectation function with complex models (i.e. with many parameters).

The third component is due to the difference between population and estimated parameters. Its size is related to the number of parameters to be estimated and to the length of the estimation sample.

The forecast error due to the last two components - model misspecification and estimation - is linked to a double trade-off: a) complex *versus* simple models; b) long *versus* short estimation samples. On one side, the use of long samples may bias estimates and forecasts, because of the likely presence of heterogeneity and structural change; on the other, the low number of degrees of freedom – either because of too few data or too many parameters or both - may affect the precision of estimates and forecasts (see e.g. Pesaran and Timmermann, 2007).

In this paper we approximate the conditional expectation function with alternative mixes of the trade-offs discussed above. Regarding the issue of the simple/complex trade-off, we use the univariate ARIMA model (point A in Table 1), alternative single-equation BM for the aggregate IPI (with only a few indicators, point B), multiple-equation disaggregate BM (with a large number of indicators, point C) and the unstructured empirical indicator approach, based on a very large data set of indicators (point D). BM parameters are estimated with OLS and SUR techniques. In the case of FM we use both static and dynamic estimators. The possible bias arising from the non-constancy of the ϑ^{v} parameters is dealt with rolling estimates with different windows.

The forecast errors obtained from different forecasting methods, i.e. combinations of different models with alternative estimators over rolling samples of different size, are compared by using the Giacomini and White test (2006, GW henceforth). The null hypothesis implies that alternative forecasting methods are equally accurate at the forecast target date $T+v+h$ using the available information set at time $T+v+m$.[9] The GW test is particularly suitable for our analysis because is valid under very general data assumptions (including non-constant data generating processes, as typical in the context of forecasting with indicators) and for both nested and non-nested models (e.g. single-equation BM clearly nest ARIMA specifications) estimated with different techniques, over different samples, with revised and unrevised data.[10]

3. The traditional bridge framework and its *out-of-sample* performance

In order to set up an empirical framework to mimic as close as possible the IPI forecasting activity in real-time using the BM, we need three main ingredients. First, a real-time data set representing the data availability at any given date in the past; second, alternative modelling strategies covering the wide range of the trade-offs described above; third, a realistic forecasting practice merging timely availability of indicators with the empirical models.

The real-time data set. The real-time data set includes monthly data for the Italian IPI and its sub-sectors based on the economic destination of the products as well as quantitative and qualitative indicators selected on the basis of their reliability and timeliness (see the Appendix for more details). 60 IPI vintages cover the period 1980-2007; the data set of the 30 indicators selected for our BM is not organised in vintages because Italian raw indicators are never revised. All the models are estimated in levels[11] of raw data; quantitative variables have been mechanically adjusted for trading day variations.[12] Besides the motivations of Wallis (1974) seminal work, we have decided to use raw data also because they are directly available and avoid filtering problems that are exaggerated in real-time data sets. [13]

---

[9] Under the null hypothesis the GW test is distributed as a χ^2 with q degrees of freedom, where q is the dimension of the test function. With q=1, as in our paper, only a constant is considered; with q>1 other variables are used in order to help distinguishing between the forecast performance of the two methods.

[10] Other similar tests, such as Diebold and Mariano (1995), are not normally distributed for nested models (see West, 1996) and in presence of data revisions (on this, see Clark and McCracken, 2007).

[11] The choice of levels instead of logarithms follows from the results in Marchetti and Parigi (2000) and has been confirmed by pre-processing with the program TRAMO as in Osborn *et al.* (1999).

[12] More specifically, if x_r is the raw variable the adjustment is given by: $x_a = x_r \, td_{base} / td_t$, where td_{base} is the average monthly number of trading days in the base year (2000 in our case) and td_t is the number of trading days in month *t* (see Bodo and Signorini, 1987, and the appendix in Bodo *et al*., 1991, for more details).

[13] With monthly USA series, Ghysels *et al.* (2002), find that monetary policy rules based on raw data have more predictive power than those based on seasonally adjusted data. Swanson and van Djik (2007) note that the seasonal adjustment process - highly non-linear - weakens the linkage between first available and final data.

The set up of the econometric models. The BM are specified following the general-to-specific modelling approach over the period 1996.2-2003.1 (details about the specifications can be found in the Appendix). The specifications are kept fixed in order to avoid the arbitrariness of *ex post* searches with data that are known to the modeller (for an alternative approach, see Golinelli and Parigi, 2008). The models are re-estimated at each step over the 2003.2-2007.12 forecasting horizon by considering three windows of different size: 48, 84 and 138 months.

The out-of-sample exercises emulating the real-time forecasting practice. The out-of-sample performance is assessed through a number of forecasting exercises designed to mimic the typical behaviour of practitioners. Forecasting ability is often analyzed with *pseudo ex-ante* exercises using latest available data, while a feature of our analysis is the use of real-time data in both estimation and forecasting. In this, we follow the common practice - motivated by the assumption that the statistical information improves through revisions - of using fully updated data-set prior to each prediction being made (see Corradi *et al.*, 2007, for a different point of view). IPI predictions one- two- and three-months ahead from vintage v are compared with the corresponding first IPI releases published, respectively, one, two and three months later: the historical data to compute the forecast errors are provided by vintages v+1, v+2 and v+3, as in equation (1).[14] Alternatively, these experiments may be conducted by using only the latest vintage of data, i.e. the best-quality data as they embody all the revisions.

Given the publication calendar of the Italian IPI, the one month-ahead forecast of the index is computed by assuming that indicator observations are completely known (nowcast case). The two-step ahead IPI forecasts require auxiliary models to extrapolate the indicators; more specifically, univariate ARIMA models are used to forecast all indicators but the electricity consumption and the temperatures, for which we use as a proxy the data in the first fortnight of the month to be forecast (already known at the time of the prediction). Finally, for the three-month ahead IPI predictions, all indicators should be extrapolated, one month ahead for electricity consumption and temperatures, two months ahead for the others (this is similar to a pure forecast exercise). Overall, the IPI forecasts exploit a decreasing amount of known indicators, from the case (one month ahead) when all indicators are known to the one (three months ahead) when all have to be extrapolated to some degree.

[14] For example, the first regression based on the T=84 rolling window uses the first IPI vintage to estimate the parameters over the period 1996.2-2003.1. The estimated models are then used to forecast the second, third and fourth IPI vintage, respectively one- two- and three-step ahead.

It is usually found that the average of different model predictions is associated with a better performance - as measured by a lower root mean squared forecast error (RMSE) - than the single predictions, probably because of some form of misspecification in models (see Hendry and Clements, 2002, and Stock and Watson, 2004). We therefore consider three simple averages of: the three single-equation BM forecasts; the three aggregate multiple-equation sub-sector forecasts; and the six previous alternative IPI forecasts.[15] This raises to nine the total number of IPI forecasts with BM (one, two and three months ahead).

Although the estimates refer to raw data, the results on the different forecasting performances are presented on the basis of seasonal adjusted data (SA, henceforth), to avoid uninformative seasonal volatility (the results based on raw data are however qualitatively similar and are available upon requests). More specifically, the SA forecasts have been obtained by applying at each step of the forecasting horizon the TRAMO-SEATS[16] procedure (with the same parameter set used by Istat, the Italian national statistical institute) to the IPI raw series augmented by the corresponding one, two or three-month ahead forecasts.

Tables 2-4 of this section report the main results about the forecasting ability of the different models on the basis of the RMSE[17]. Different columns in the tables refer to different combinations of alternative modelling options: rolling window sizes, estimation methods and types of data (real-time or latest available); different rows refer to the models listed in Table 1 sub points A-C. The lower part of each table reports the outcomes of the GW test on the equality between the predictive ability of the best performing model (chosen on the basis of the lowest RMSE in each column) and that of the other models. Across columns, the GW test compares the RMSE of forecasting methods based on the same model but with different samples and estimators, while it is not applicable to compare methods forecasting different types of data (i.e. real-time *vs* latest-available), as the IPI measures vary.

Table 2 about here

[15] More complex forecast combination techniques are presented in Altavilla and Ciccarelli (2007).

[16] TRAMO (Time Series Regression with ARIMA Noise, Missing Observations, and Outliers) performs estimation, forecasting and interpolation of models with missing observations and ARIMA errors, in the presence of outliers. SEATS (Signal Extraction in ARIMA Time Series) performs an ARIMA-based decomposition of an observed time series into unobserved components (see Gomez and Maravall, 2001a and 2001b, for references on TRAMO and SEATS).

[17] Other typical statistics (i.e. the mean error; the mean absolute error; the percentage of times that the sign of the IPI growth rate is correctly predicted; the percentage of times that the forecast error of each model is smaller in absolute value than that of benchmark time series model) deliver qualitatively similar results.

The forecasting performance one month ahead is assessed in Table 2. In the first column (baseline) the modelling activity assumes the traditional settings of the literature: a window size of about seven years, OLS estimates and the latest available data. The main results are three. First, a better performance of the IPI forecasting methods based on coincident indicator observations, which halves the RMSE of the ARIMA model. Second, a better performance of the aggregation of sectoral forecasts over the single-equation BM (the RMSE improvement is in the 20-90% range). Both these findings are statistically significant according to the GW test. Thirdly, the "average puzzle" at work: simple averages tend to outperform the best of their components, as shown by the lowest RMSE of the overall average in the last row of the table.

Different window sizes (both shorter and longer) imply a worse forecasting performance for almost all models, though not uniformly. The reduction in the window size damages more the aggregation of sectoral forecasts, presumably because there are more parameters to be estimated than in single-equation BM. In this context, the overall average is still the best performer. On the other hand, the increase in the window size induces a bias, presumably due to parameter shifts in single-equation BMs (e.g. longer time spans negatively affect the explanatory power of *ad hoc* trend components). It is worth noting that when changes in the window size markedly worsen the BM performance, the results of the overall averages also significantly worsen, i.e. the average puzzle no longer works.

The effect of SUR estimates is a significant worsening of the performance of the three simple averages, maybe because the statistical compensation effect of the errors is somehow already taken into account at the estimation level.

Finally, the "vintage effect" (last column of Table 2) does not appear to affect the forecasting methods ranking that emerges from of the baseline results, where the latest available data are used. This suggests the virtual irrelevance of the real-time issues in the context of monthly IPI revisions when all indicators are not revised. Consistently with the structure of the forecasting models, only the RMSE of the aggregation of sectoral forecasts worsen by more than 5% probably because the amount of the revisions is more pronounced at disaggregate level.

Tables 3 and 4 about here

The extension of the forecasting horizon to two (Table 3) and three months (Table 4) implies (not surprisingly) an increase of the RMSE. The shorter window and the data revision effects seem to disappear, probably because they tend to be overshadowed by the unpredictable shocks occurring over the forecasting horizon.

Two main changes with respect to the one-step ahead case deserve some comments. The best performer is no longer the simple average of the different forecasts. At longer horizons, the predictions from single equation models deteriorate sharply, thus raising the RMSE of the overall average. The performance of the disaggregate models worsens less, suggesting that, among the large number of indicators used in this modelling approach, there might be some with leading properties. The implication is that when indicators are forecast with auxiliary models, it could be advisable to exploit a larger information set. Next section will deepen this point.

Another interesting finding is that the worsening of the baseline RMSE related to the expansion of the window size is permanent and not related to the chosen forecasting horizon; the GW tests always reject the null of equal performance of the models using the baseline settings, independently of the forecasting horizon. This is possibly due to the increasing risk of forecast biases because of the presence of trending levels without any structural relationships (excluded by definition in the BM approach).

4. The construction of factor-based models and assessing their IPI forecasts

Forecasting with approximate factor models is usually carried out in a two step process: firstly, factors are estimated from all available indicators; secondly, the variable of interest is forecast by projecting it onto the space spanned by the estimated factors.

In order to define more precisely the first step, let x_i be a stationary process from the *n* dimensional co-stationary vector X = $[x_{1t},\dots x_{Nt}]$. After suitable transformation and standardisation, x_i can be seen as the sum of two orthogonal and non-observable components:

$$x_{it} = \chi_{it} + \varepsilon_{it} \tag{2}$$

χ_{it} is driven by *q* shocks, common to all variables while ε_{it} is a variable-specific shock, only mildly correlated across the *x*'s. After imposing a factor structure to the common component, equation (2) can be re-written as follows:

$$x_{it} = B_i(L)\cdot U_t + \varepsilon_{it} = \left[b_i^1(L),\dots,b_i^q(L)\right]\cdot\begin{bmatrix} u_t^1 \\ \dots \\ u_t^q \end{bmatrix} + \varepsilon_{it} \tag{3}$$

where U_t is the ($q\times1$) vector of staked shocks and the elements of $B_i(L)$ are polynomials of order *s* in the lag operator *L*.[18] Equation (3) is called the dynamic representation of the factor model. Considering each lag as a distinct factor, equation (3) can be written as (static representation):

[18] Some of the coefficient of the j-th polynomial $b_i^j(L)$ may be zero.

$$x_{it} = b_{i0}^{1} \cdot u_{t}^{1} + \ldots + b_{is}^{1} \cdot u_{t-s}^{1} + \ldots.. + b_{i0}^{q} \cdot u_{t}^{q} \ldots + b_{is}^{q} \cdot u_{t-s}^{q} + \varepsilon_{it} = \left[b_{i0}^{1}, \ldots, b_{is}^{q}\right] \cdot \begin{bmatrix} u_{t}^{1} \\ \ldots \\ u_{t-s}^{q} \end{bmatrix} =$$

$$= \Lambda_{i} \cdot F_{t} + \varepsilon_{it} \qquad (4)$$

where Λ_i and F_t are r-dimensional vectors. The relationship between the two representations is given by $r = q\ (s+1)$, where r is the number of static factors, q is the number of dynamic factors and s is the number of lags of the latter.

The factors F may be estimated according to two main strategies based respectively on representations (4) and (3) of the model (2): the static and the dynamic approaches. With the former, the estimates of the factors are the static principal components of the covariance matrix of the data (see Stock and Watson, 2002a and 2002b, SW henceforth); with the latter, the factors are estimated by dynamic principal components (see Forni *et al.*,2005, FHLR henceforth). While both approaches provide a consistent estimate of the space spanned by the factors, the dynamic one is more efficient, as it explicitly considers the dynamic nature of the shocks.

The second step of the process consists in forecasting the variable of interest h-steps ahead. Predictions can be obtained through either a parametric regression (static approach) or a non-parametric procedure (dynamic approach) or a mixture of the two (see Boivin and Ng, 2005, and D'Agostino and Giannone, 2006).

Factors are estimated with a data set of up to 110 quantitative and qualitative indicators covering different sectors of the economy (more details are in Table A2 of the Appendix). All series have been pre-processed, so as the data set conforms to the theoretical conditions (such as stationarity) needed for estimation of factor models. Given the staggered nature of the releases of different series, all indicators with missing observations for the latest months have been shifted in time so as to have a balanced panel.[19]

Although the "unstructured" nature of FM seems appealing because it reduces the risk of misspecification, the appropriate number of factors has to be chosen. In early empirical applications, the number of factors was usually selected on a judgmental basis by considering the percentage of the cumulated variance of the matrix of observations explained by the first n eigenvalues, and/or the marginal increase in the percentage of total variance explained by the n^{th} eigenvalue. More recently, formal optimality criteria have been proposed, which are based on the trade-off between goodness of fit and over-parameterisation (see Bai and Ng, 2002 and

[19] Let k_i be the number of missing observations of x_i , at the end of the sample so that at time T the latest available observation is $x_{i;T-ki}$; then the variable z_i , used in estimation, is related to x_i by the relationship: $z_{it} = x_{it-k_i}$.

2007, Amengual and Watson, 2007, and Hallin and Liska, 2007). As it is not clear which criterion delivers the best results in forecasting, in this paper the models are estimated according to both the early judgmental and the optimal criteria.

All the models are first estimated over the period 1996.2-2003.1 (84 months, the same window size as for the BM baseline) and are used to compute from 1 to 3 step-ahead forecasts starting from 2003.2. At each step the estimation window is shifted by one observation at a time, selection criteria are re-computed, parameters and factors are re-estimated and new forecasts are computed.[20] As for the BM case above, we also run the same forecasting exercise with shorter and longer window sizes (48 and 138 months respectively) and with real time data (in this case results are basically the same as with latest available data as in the previous section).

Table 5 reports the main results about the FM forecasting ability in terms of the RMSE. Along the rows, we compare the (benchmark) ARIMA model, two forecast-averaging models (FA-ARDL) and four different FMs. The two FA-ARDL forecasts have been obtained by the simple average of forecasts from the bivariate autoregressive distributed lags models of the IPI on all the indicators (110) and excluding the indexes of the IP sub-sectors (80)[21]; for a similar approach, see Stock and Watson (1999) and Kitchen and Monaco (2003).

Table 5 about here

The static factor models differ according to the forecasting equation. In one case, the h-step ahead forecasts, IPI_{t+h}, are expressed as a direct function of both the factors and the IPI in *t* and earlier (SW-D); in the other the h-step forecast is obtained by iterating sequential one-step forecasts (SW-S). Two generalised dynamic factor-model forecasts are reported, according to the selection criterion adopted to determine the number of factors: fixed “cut-off” rule (FHLR-F) and optimal criteria (FHLR-O), respectively. The forecasts from the FHLR-F procedure have been computed by considering alternative estimation window sizes.

In line with the previous empirical literature, the forecasting performance of all the models in Table 5 is better than that of the benchmark univariate ARIMA model at all horizons: the accuracy gains range between 5-18% and 22-35% for FA-ARDL and FM respectively. FM forecast are always associated with statistically significant accuracy gains, without relevant differences between the static and the dynamic approaches. Moreover, within both classes of FM, the outcomes appears to be independent from both the specification of the forecasting

[20]The SA levels of the historical-plus-forecast IPI raw data are obtained with the TRAMO-SEATS procedure.

[21] In the case of the 110 bivariate models, the first 30 indicators in Table A2 are released at the same time as the IPI and so are not used simultaneously but lagged one period.

equation and the optimality criteria. The average puzzle however works even in the FM case: the effects of using alternative estimation window sizes are less pronounced and asymmetric with respect to the BM case. The FM forecasting ability worsens with shorter windows (because of less efficient factor estimates), while it is basically unchanged in longer ones, suggesting that the joint use of many indicators should help preventing the insurgence of breaks. These findings remain broadly unchanged when the data set is restricted to the 30 indicators included in the BM of the previous section (results are not reported but available upon request), thus showing the irrelevance in this case of the Boivin and Ng (2006) *caveats* about oversampling indicators from specific groups.

5. Discussion

The alternative IPI forecasts from bridge to factor-based models are compared in Table 6. Along the rows five alternative forecasting methods are reported for each prediction horizon (from one- to 3-months ahead): ARIMA, the average of single-equation BM, the average of multiple-equation BM, the average of FM and the overall average of bridge and factor models. The first two columns report the RMSEs and their ratios with respect to the ARIMA model.

Table 6 about here

The picture is quite clear-cut: short-term information always matters. Both bridge and factor-based models always perform significantly better over ARIMA, suggesting that in both approaches the short-run indicators signal dominates the noise, independently of the different methods used to extract it. In this context, BM significantly outperform in efficiency FM. In the first case the researcher can increase the amount of signal extracted from the available indicators and improve up to 30-40% the factor-based model RMSE. Though FM are appealing because of their ability to cope with many variables and to capture the business cycle component of the target variable, they fail in fully anticipating the highly idiosyncratic part characterising the short-term dynamics.

Table 6 also reports the Fair and Shiller (1990; FS) t-statistics for the null that the forecast of the model in the row contains no information relevant to future IPI realisations already not in the model in the column (i.e. the model on the row is encompassed by the model in the column). According to this test ARIMA forecasts are generally encompassed by models based on indicators: the FS t-statistics in the “ARIMA” row are never significant, contrary to those in the “ARIMA” column. The parsimonious use of indicators leads BM to outperform all other forecasting approaches, as their FS t-statistics always reject the null against all other forecasts. Among alternative BM, the multiple-equation approach is the best performing in terms of the FS test: the parsimonious (i.e. with restrictions) exploitation of 30 indicators allows the IPI predictions to contain all relevant information.

Overall, the findings in Table 6 lend support to the superiority of the BM approach in terms of forecasting performance. This seems at odds with the results reported in a similar, recent paper by Angelini et al. (2008) which show that FM outperform BM in predicting the euro area GDP short-term evolution. Angelini *et al*. exercise is however quite different from the one presented in this paper. In their quarterly forecasting analysis a crucial factor is the accrual of monthly information over the reference quarter, which is exploited by FM, but not by BM. More specifically, in Angelini *et al*. the BM include only a few quantitative indicators, usually available with a longer delay, and, contrary to FM, cannot exploit more timely information, such as that contained in survey variables. The only way such information could be exploited in the BM approach is through the (monthly) auxiliary models used to extrapolate the quantitative indicators. Given that in the nowcast case the BM forecasting performance becomes similar to the FM one, the comparison in Angelini et al. may be seen as a test of the forecasting performance of the auxiliary models. On this regard, our paper may be seen as an analysis of alternative auxiliary models for the prediction of quarterly GDP, which may be used during the reference quarter, from the beginning when no information is available on monthly indicators (pure forecast case), till when all indictors are known (nowcast case), considering the intermediate cases when only some monthly indicators are released.

Overall, “horse races” and more general comparisons of different forecasting methods may lead to a better understanding of the advantages and disadvantages of the alternative approaches. Bridge models generally provide very precise forecasts which are also very easy to interpret. Indicators that appear to be not linked or only loosely linked to the target variable are ignored. This has two positive implications: *(i)* BM predictions enable to “tell the story” of the forecast on the basis of the evolution of the explanatory indicators. This is a very important feature in periods characterised by deep and rapid changes, when it is needed not only to quantify the relevance of specific events, but to understand their origin as well (recent advances on this topic in the field of FM are shown in Banbura and Runstler, 2007). (*ii*) BM data sets are smaller and less costly to update. The claim that with FM all relevant information is used because nothing is a priori discarded implies that the data sets behind FM are potentially very large, including indicators from many sources, with very different characteristics and quality standards.

However, building BM is more difficult and arbitrary, requiring a number of subjective choices – entailing crucial trade-offs – about the model specification and the size of the estimation sample. We have shown that the IPI forecasts from multiple-equation models are often significantly better than those from single-equation models, especially at longer forecast horizons, suggesting the likely presence of some leading indicators in the information set used by multiple equations. In addition, BM forecasting performance appears to be sensitive to the size of the estimation sample, as their RMSE worsens when both shorter and longer estimation samples are used. Instead, FM appears to be fairly unaffected by the choice of the length of the estimation sample. This confirms the results in the literature about the greater stability of FM forecasts, probably due to the large information set.

In the light of the discussion above, it emerges that the two broad classes of approaches appear to be fundamentally complementary, as the strengths of the one correspond to the weaknesses of the other. Factor-forecasting performance is less efficient because it cannot pre-select the "best" indicators from large data sets and it is less interpretable. This weakness however reduces the risk of omitting important predictors, allows to exploit new information as soon as it becomes available, prevents uncertainty about modeller's skill/experience and delivers forecasts that are less prone to regime-shift biases. Thus, the true challenge for future research is to find how such pros and cons of the two approaches may be fruitfully merged by the forecasting practice.

References

Altavilla, C. and M. Ciccarelli (2007), "Information Combination and Forecast (St)Ability: Evidence From Vintages of Time-Series Data", *Working Paper*, No. 846, European Central Bank.

Altissimo, F. A. Bassanetti, R. Cristadoro. M. Forni, M. Hallin, M. Lippi, L. Reichlin and G. Veronese (2001), "EuroCOIN: A Real Time Coincident Indicator of the Euro Area Business Cycle ", *CEPR Discussion Paper*, No. 3108.

Altissimo, F. R. Cristadoro. M. Forni, M. Lippi and G. Veronese (2007), "New Eurocoin: Tracking Economic Growth in Real Time", *Temi di Discussione*, No. 631, Banca d'Italia.

Amengual, D. and M. W. Watson (2007), "Consistent Estimation of the Number of Dynamic Factors in a Large N and T Panel", *Journal of Business and Economic Statistics*, Vol. 25, No. 1, pp. 91-96.

Angelini, E., G. Camba-Mendez, D. Giannone and L. Reichlin (2008), "Short Term Forecasts of the Euro Area GDP Growth", *CEPR Discussion Paper*, No. 6746.

Baffigi, A., R. Golinelli and G. Parigi (2004), "Bridge Models to Forecast the Euro Area GDP", *International Journal of Forecasting*, Vol. 20, pp. 447-460.

Bai, J. and S. Ng (2002), "Determining the Number of Factors in Approximate Factor Models", *Econometrica*, Vol. 70, pp. 191-221.

Bai, J. and S. Ng (2007), "Determining the Number of Primitive Shocks in Factor Models", *Journal of Business and Economic Statistics*, Vol. 25, No. 1, pp. 52-60.

Banbura, M. and G. Runstler (2007), "A Look into the Factor Model Black Box – Publication Lags and the Role of Hard and Soft Data in Forecasting GDP", *Working Paper*, No. 751, European Central Bank.

Bodo, G., A. Cividini and L. F. Signorini (1991), "Forecasting the Industrial Production Index in Real Time", *Journal of Forecasting*, Vol. 10, pp. 285-299.

Bodo, G., R. Golinelli and G. Parigi (2000), "Forecasting Industrial Production in the Euro Area", *Empirical economics*, Vol. 25, pp. 541-561.

Bodo, G. and L. F. Signorini (1987), "Short-Term Forecasting of the Industrial Production Index", *International Journal of Forecasting*, Vol. 3, pp. 245-259.

Boivin, J. and S. Ng (2005), "Understanding and Comparing Factor-Based Forecasts", *International Journal of Central Banking*, Vol. 1, No. 3, pp. 117-151.

Boivin, J. and S. Ng (2006), "Are More Data Always Better for Factor Analysis?", *Journal of Econometrics*, Vol. 132, pp. 169-194.

Bruno, G. and C. Lupi (2004), "Forecasting Industrial Production and the Early Detection of Turning Points", *Empirical economics*, Vol. 29, pp. 647-671.

Clark, T. E. and M. W. McCracken (2007), "Test of Equal Predictive Ability With Real-Time Data", paper presented at the conference: *Real-Time Analysis and Methods in Economics*, Federal Reserve Bank of Philadelphia, April.

Clements, M. P. and D. F. Hendry (2002), "Modelling Methodology and Forecast Failure", *Econometrics Journal*, Vol. 5, pp. 319-344.

Corradi, V., A. Fernandez and N. R. Swanson (2007), “Information in the Revision Process of Real-Time Datasets”, paper presented at the conference: *Real-Time Analysis and Methods in Economics*, Federal Reserve Bank of Philadelphia, April.

Cristadoro, R., M. Forni, L. Reichlin and G. Veronese (2005), “A Core Inflation Indicator for the Euro Area”, *Journal of Money, Credit, and Banking*, Vol. 37, pp. 539-560.

Cristadoro, R., G. Saporito and F. Venditti (2007), “Forecasting Inflation and Tracking Monetary Policy in the Euro Area: Does National Information Help?”, *Temi di Discussione*, forthcoming, Banca d'Italia.

Croushore, D. and T. Stark (2001), “A Real-Time Data Set for Macroeconomists”, *Journal of Econometrics*, Vol. 105, pp. 111-130.

D’Agostino, A. and D. Giannone (2006), “Comparing Alternative Predictors Based on Large Panel Factor Models”, *Working Paper*, No. 680, European Central Bank.

Diebold, F. X. and R. S. Mariano (1995), “Comparing Predictive Accuracy”, *Journal of Business and Economic Statistics*, Vol. 13, pp. 253-263.

Diebold, F. X. and G. D. Rudebusch (1991), “Forecasting Output with the Composite Leading Index: a Real Time Analysis”, *Journal of the American Statistical Association*, Vol. 86, pp. 603-10.

Fair, R. C. and R. J. Shiller (1990), “Comparing Information in Forecasts from Econometric Models”, *The American Economic Review*, Vol. 80, No 3, pp. 375-389.

Forni, M., M. Hallin, M. Lippi and L. Reichlin, (2000), “The Generalized Dynamic Factor Model: Identification and Estimation”, *Review of Economics and Statistics*, Vol. 82, No. 4, pp. 540-554.

Forni, M., M. Hallin, M. Lippi and L. Reichlin, (2005), “The Generalized Factor Model: One Sided Estimation and Forecasting ”, *Journal of the American Statistical Association*, Vol. 100, No. 471, pp. 830-840.

Ghysels, E., N. R. Swanson and M. Callan (2002), “Monetary Policy Rules with Model and Data Uncertainty”, *Southern Economic Journal*, Vol. 69, No. 2, pp. 239-265.

Giacomini, R. and H. White (2006), “Tests of Conditional Predictive Ability”, *Econometrica*, Vol. 74, pp. 1545-1578.

Golinelli, R. and G. Parigi (2007), “The Use of Monthly Indicators to Forecast Quarterly GDP in the Short Run: An Application to the G7 Countries”, *Journal of Forecasting*, Vol. 26, pp. 77-94.

Golinelli, R. and G. Parigi (2008), “Real-Time Squared: A Real Time Data Set for Real-Time GDP Forecasting”, *International Journal of Forecasting*, forthcoming

Gomez, V. and A. Maravall (2001a), “Automated Modelling Methods for Univariate Series”, in Pena D., Tiao G. C. and Tsay R. S. (eds.) *A Course in Time Series Analysis*, Ch. 7, New York, Wiley and Sons.

Gomez, V. and A. Maravall (2001b), “Seasonal Adjustment and Signal Extraction in Economic Time Series”, in Pena D., Tiao G. C. and Tsay R. S. (eds.) *A Course in Time Series Analysis*, Ch. 8, New York, Wiley and Sons.

Hallin, M. and R. Liska (2007), “Determining the Number of Factors in the General Dynamic Factor Model”, *Journal of the American Statistical Association*, Vol. 102, No. 478, pp. 603-617.

Hendry, D. F. and M. P. Clements (2002), “Pooling of Forecasts”, *Econometrics Journal*, Vol. 5, pp. 1-26.

Hendry, D. F. and K. Hubrich (2007), “Combining Disaggregate Forecast or Combining Disaggregate Information to Forecast an Aggregate”, *mimeo*.

Kitchen, J. and R. Monaco (2003), “Real-Time Forecasting in Practice”, *Business Economics*, October, pp. 10-19.

Klein, L. R. and S. Ozmucur (2005), “The University of Pennsylvania Models for High-Frequency Macroeconomic Modelling”, paper presented at the Project LINK Spring Meeting, UNAM, Mexico City, Mexico.

Klein, L. R., and E. Sojo (1989), “Combinations of High and Low Frequency Data in Macroeconometric Models”, in L. R. Klein and J. Marquez (eds.), *Economics in Theory and Practice: An Eclectic Approach*, Kluwer Academic Publishers, pp. 3-16.

Kozicki, S. (2002), “Comments on Forecasting with a Real-Time Data Set for Macro-economists”, *Journal of Macroeconomics*, Vol. 24, pp. 541-557.

Ladiray, D. and D. O’Brien (2003), “Nowcasting Eurozone Industrial Production”, *Working Papers and Studies*, European Commission.

Marcellino, M., J. H. Stock and M. W. Watson (2003), “Macroeconomic Forecasting in the Euro Area: Country Specific Versus Area-Wide Information”, *European Economic Review*, Vol. 47, pp. 1-18.

Marchetti, D. J. and G. Parigi (2000), “Energy Consumption, Survey Data and the Prediction of Industrial Production in Italy: A Comparison and Combination of Different Models”, *Journal of Forecasting*, Vol. 19, pp. 419-440.

Osborn, D. R., S. Heravi and C. R. Birchenhall (1999), “Seasonal Unit Roots and Forecasts of Two-Digit European Industrial Production”, *International Journal of Forecasting*, Vol. 15, pp. 27-47.

Parigi, G. (2005), “New Estimates of the Italian Industrial Production Index”, Banca d’Italia, *mimeo*.

Parigi, G. and G. Schlitzer (1995), “Quarterly Forecasts of the Italian Business Cycle by Means of Monthly Economic Indicators”, *Journal of Forecasting*, Vol. 14, pp. 117-141.

Pesaran, H. and A. Timmermann (2007), “Selection of Estimation Window in the Presence of Breaks”, *Journal of Econometrics*, Vol. 137, pp. 134-161.

Runstler, G. and F. Sedillot (2003), “Short-Term Estimates of Euro Area Real GDP by Means of Monthly Data”, *Working Paper*, No. 276, European Central Bank.

Schumacher, C. and J. Breitung (2006), “Real-Time Forecasting of GDP Based on a Large Factor Model with Monthly and Quarterly Data”, *Discussion Paper*, No. 33, Deutsche Bundesbank.

Sedillot, F. and N. Pain (2003), “Indicator Models of Real GDP Growth in Selected OECD Countries”, Economics Department, *OECD Working paper*, No 364.

Stark, T. and D. Croushore (2002), “Forecasting With a Real-Time Data Set For Macroeconomists”, *Journal of Macroeconomics*, Vol. 24, No. 4, pp. 507-568.

Stock, J. H. and M. W. Watson (1996), “Evidence on Structural Instability in Macroeconomic Time Series Relations”, *Journal of Business Economics and Statistics*, Vol. 14, p. 11-30.

Stock, J. H. and M. W. Watson (1999), “Forecasting Inflation”, *Journal of Monetary Economics*, Vol. 44, p. 293-335.

Stock, J. H. and M. W. Watson (2002a), “Macroeconomic Forecasting Using Diffusion Indexes”, *Journal of Business and Economic Statistics*, Vol. 20, pp. 147-162.

Stock, J. H. and M. W. Watson (2002b), “Forecasting Using Principal Components from a Large Number of Predictors ”, *Journal of the American Statistical Association*, Vol. 97, pp. 1167-1179.

Stock, J. H. and M. W. Watson (2004), “Combination Forecasts of Output Growth in a Seven-Country Data Set”, *Journal of Forecasting*, Vol. 23, pp. 405-430.

Stock, J. H. and M. W. Watson (2006), “Forecasting with Many Predictors”, in G. Elliott, C. W. J. Granger and A. Timmermann (eds.), *Handbook of Economic Forecasting*, North Holland, Chapter 10, pp. 515-554.

Swanson, N. R. and D. van Dijk (2007), “Are Statistical Agencies Getting it Right? Data Rationality and Business Cycle Asymmetry”, *Journal of Business and Economic Statistics*, Vol. 24, No. 1, pp. 24-42.

Trehan, B. (1989), “Forecasting Growth in Current Quarter Real GNP”, Federal Reserve Bank of S. Francisco, *Economic Review*, Winter, pp. 39-52.

Trehan, B. (1992), “Predicting Contemporaneous Output”, Federal Reserve Bank of S. Francisco, *Economic Review*, No 2, pp. 3-11.

Wallis, K. F. (1974), “Seasonal Adjustment and Relation Between Variables”, *Journal of the American Statistical Association*, Vol. 69, pp. 18-31.

West, K. D. (2006), “Forecast Evaluation”, in G. Elliott, C. W. J. Granger and A. Timmermann (eds.), *Handbook of Economic Forecasting*, North Holland, pp. 99-134.

White, H. (1980), “A heteroskedasticity-Consistent Covariance Matrix and a Direct Test for Heteroskedasticity”, *Econometrica*, Vol. 48, pp. 817-838.

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

Zizza, R. (2002), “Forecasting the Industrial Production Index for the Euro Area Through Forecasts for the Main Countries”, *Temi di Discussione*, No. 441, Banca d'Italia.

Appendix – The Italian data set and bridge models specification

IPI vintages. IPI data are issued by the Italian Statistical Agency (Istat). The first available vintage of aggregate IPI spans over the 1953.1-2001.1 period. However, since the first vintage of disaggregate IPI sub-sector data spans over the 1980.1-2003.1 period, our real-time data-set for both aggregate and disaggregate IPI data is made of 60 monthly vintages: from v = 1 (1980.1-2003.1) to v = 60 (1980.1-2007.12). During the 2003-2007 period, Italian IPI did not experience any benchmark revision, as the year 2000 is the same base year of all our vintages. Therefore, statistical revisions are the only source of real-time changes affecting IPI updates in our exercises. In this paper we will use the following labels:

IPI index of industrial production (total manufacturing sector, i.e. aggregate IPI)
ConPI index of industrial production (consumption goods sub-sector)
IntPI index of industrial production (intermediate goods sub-sector)
InvPI index of industrial production (investment goods sub-sector)
EnePI index of industrial production (energy sub-sector)

The data set of BM indicators. The 30 indicators in our data set are not available at the same initial period. However, all of them span the common 1991.8-2007.12 period. In order to prevent the loss of few useful indicators, namely *ITEIMG, ITNOMG*, and *ITPGMG* from the purchasing managers index survey (see Table A1), they have been "backcasted" from 1996.6 to 1991.8 with other proxy-indicators. The Italian real-time data set is limited to the indexes of industrial production, both aggregate and by sub-sector, because the whole set of indicators is not subject to revision. In order to guard against future information creeping into past vintage data through potentially heterogeneous seasonal adjustment procedures, all the series in the data-set are not seasonally adjusted. Table A1 is devoted to merge the list of 30 variables belonging to our indicator information set with the corresponding seven econometric specifications where they are included.

Table A1 about here

Four single-equation BM directly forecast the IPI: ARIMA, OSBM, USBM and FSBM (see Table 1 in the text). The IPI can also be predicted by aggregating the forecasts of the four IPI sub-sectors (ConPI, IntPI, InvPI and EnePI) using alternative aggregator functions (ASGR, ASLWE and ASWL, see Table 1). In turn, each sub-sector production is predicted by a single equation.

The ARIMA specification is obtained with the usual time series techniques. The selected specification is given by a parsimonious AR(3) model for the 12-month growth rate of the IPI.

In the first bridge model, defined as OSBM (see e.g. Bodo and Signorini, 1987, Bodo *et al.*, 1991 and Marchetti, and Parigi, 2000), the main driving force of the IPI (adjusted for trading day variation) is consumption of electricity, along with meteorological data to take account of non industrial use of electricity. Successive, minor, refinements of this specification lead to the following version (see Bodo *et al.*, 2000, Zizza, 2002, Ladiray and O'Brien, 2003 and Bruno and Lupi, 2004):

$$IPI_t = \alpha + \beta log(ENELN_t) + \sum_{i=1}^{2} \gamma_i \; TEMPN_t^{\,i} + \eta \frac{\sum_{i=1}^{3} IPI_{t-i}}{3} + \sum_{i=1}^{2} \delta_i T^{\,i} + Sdum \; + \varepsilon_t$$

Where: T is a deterministic linear trend; *Sdum* is the set of seasonal dummy variables (to account for possible cyclical shifts in the seasonal patterns the dummy set also includes two seasonal dummy variables, for August and December, which are multiplied by the three months growth rate of orders received by manufacturing firms, *ORDISTAT*) and ε_t is the usual disturbance term. Hereafter, the label of the variables is the same as it appears in the table A1, unless explicitly stated; *mav(x,y)* represents the *y*-months uncentered moving average of *x*.

The above specification has been successively amended by Parigi (2005), by changing the trend specification and including some indicators from survey variables, leading to the USBM model:

$$IPI_t = \alpha + \beta \log\left(ENEL2_t + ENEL4_t\right) + \sum_{i=1}^{2} \gamma_i \, TEMPN^{\,i} + \lambda \, Ord_t + \\ + \mu \frac{\sum_{i=1}^{11} \Delta IPI_{t-i}}{11} + \sum_{i=1}^{2} \delta_i \frac{1}{T^i} + Sdum + \varepsilon_t$$

where *Ord=NLOINV/mav(NLOINV,36).*

More recently, given the availability of the indicators from the purchasing managers index survey, the specification has been further changed, giving rise to the FSBM:

$$IPI_t = \alpha + \beta \; ENELN_t + \delta \; \Delta ENEL9_t + \\ + \sum_{i=1}^{2} \gamma_i TEMPN^{\,i} + \phi \; Pmi_{t-1} + \eta \; Lprod_t + \iota \, Retcon_t + \varphi \frac{1}{T^4} + Sdum + \varepsilon_t$$

where: *Pmi=ITNOMG–mav(ITPGMG,2)*; *Lprod=NLPIMP/NSCIMP*; *Retcon=mav(TOIMPCN,15).* The presence of *ENEL9*, with a negative sign, is supposed to improve the correction for the consumption of electricity from the services sector. Again, the seasonal dummy variables for August and December are specified as a multiple of the trend.

The richness of the information set, apparent from the increasing complexity of the specifications above, is at the roots of the disaggregate approach. In this case, the specifications for the 4 sub-sectors of the IPI are the following.

Consumption goods:

$$ConPI_t = \alpha + \beta \frac{1}{ENEL\,2_t} + \sum_{i=1}^{2} \gamma_i \; TEMPN^{\,i} + \phi \; ITNOMG_{t-1} + \\ + \eta \; Rbus_{t-7} + \delta \; \Delta Rsto_{t-1} + \phi \Delta CLIMA_{t-3} + \iota T + \varphi \frac{1}{T^4} + Sdum + \varepsilon_t$$

where *ConPI* is the index of industrial production in the consumption goods sector; *Rsto=mav(SCORTE, 6)*; *Rbus =AFFEXA+AFFEXP*.

Investment goods:

$$InvPI_t = \alpha + \beta\left(ENEL2_t + ENEL4_t\right) + \phi \frac{1}{ENELS_t} + \sum_{i=1}^{2} \gamma_i \; TEMPN^i + \delta \; InvPI_{t-3} + \\ + \eta \; Conf_{t-1} + \phi \; \Delta Conf_{t-1} + \delta \; \Delta TASSO_{t-1} + \iota \, T + Sdum + \varepsilon_t$$

where *InvPI* is the index of industrial production in the investment goods sector; *ENELS,* the electricity consumption in the south of Italy, which is supposed to capture the adjustment for the use of electricity other than for industrial purpose; *Conf = FIDINV+CLIMA+CLICOM*. *Sdum* includes seasonal dummy variables for August and December expressed as multiple of the trend.

Intermediate goods:

$$IntPI_t = \alpha + \beta\ ENELN_t + \phi \frac{1}{ENEL\,5_t} + \sum_{i=1}^{2} \gamma_i\ TEMPN^i + \delta\ IntPI_{t-3} + +\phi\ \Delta(\ MIntPI\ )_{t-1} +$$
$$+ \iota\ FIDINTN_{t-1} + \delta\ \Delta FIDINTN_{t-1} + \mu\ ITEIMG_t + Sdum + \varepsilon_t$$

where *IntPI* is the index of industrial production in the intermediate goods sector; *MIntPI=mav(IntPI,9)*.

Energy goods:

$$EnePI_t = \alpha + \beta\ ENELS_t + \phi\left(ENEL3_t + ENEL9_t\right) + \gamma\ \Delta ENELS_t + \iota\ TEMPN_t +$$
$$+ \kappa\ TEMPS_t + \eta\ EnePI_{t-1} + \mu\ \Delta CLIMA_{t-2} + \nu\ ITPGMG_t + \rho\ \Delta Infl_t + Sdum + \varepsilon_t$$

where *EnePI* is the index of industrial production in the energy goods sector; *Infl=mav(ΔCCPNZ,5)* and *ΔCCPNZ* is the month-on-month consumer inflation rate.

As a final step, the forecasts of the above sub-sectors are aggregated into the IPI forecast by the following 3 aggregator functions.

ASGR:

$$\Delta IPI_t = \alpha + \beta\ \Delta ConPI_t + \gamma\ \Delta InvPI_t + \delta\ \Delta IntPI_t + \mu\ \Delta EnePI_t + \varepsilon_t$$

where the Δ operator refer to the 12-month growth rate of the variable.

ASLWE:

$$IPI_t = \alpha + \beta\ ConPI_t + \gamma\ InvPI_t + \delta\ IntPI_t + \varepsilon_t$$

ASWL:

$$IPI_t = a\ ConPI_t + b\ InvPI_t + c\ IntPI_t + d\ EnePI_t$$

where *a*, *b*, *c*, and *d* are the official weights underlying the computation of the aggregate IPI.

Table A2 about here

Table 2

1-MONTH AHEAD PREDICTIVE ABILITY OF ALTERNATIVE FORECASTING METHODS OVER THE 2003.2 - 2007.12 PERIOD (59 MONTHS) [1]

	Baseline	Change in the: window size				estimator		data definition	
Size of the rolling window [2]:	84	48		138		84		84	
Estimation method:	OLS	OLS		OLS		SUR		OLS	
Type of data [3]:	LA	LA		LA		LA		RT	
Models [4]:	*RMSE*	*RMSE*	*ratios [5, 6]*	*RMSE*	*ratios [5, 6]*	*RMSE*	*ratios [5, 6]*	*RMSE*	*ratios [5, 6]*
ARIMA	1.39	1.43	1.029	1.40	1.010	1.39	1.000	1.32	0.950
OSBM	0.74	0.89	1.198**	0.79	1.070	0.70	0.941	0.77	1.034
USBM	0.61	0.73	1.196**	1.04	1.700**	0.68	1.120*	0.61	0.995
FSBM	1.00	0.71	0.714	1.31	1.313**	1.02	1.025	0.98	0.986
Average of Single BM	0.57	0.68	1.206**	0.70	1.231**	0.81	1.427	0.57	1.004
ASGR	0.54	0.75	1.390***	0.61	1.134**	0.65	1.212***	0.56	1.048
ASLWE	0.55	0.72	1.301***	0.57	1.028	0.62	1.116	0.59	1.069
ASWL	0.52	0.74	1.413***	0.59	1.128**	0.62	1.189**	0.56	1.070
Average of the Aggregation of Sectors	0.53	0.73	1.376***	0.58	1.099	0.62	1.179**	0.56	1.066
Average of all models (excl. ARIMA)	0.52	0.68	1.308***	0.62	1.186**	0.64	1.227***	0.54	1.029

RMSE ratios with respect to the best performance in each column [6]

	Baseline	window size (48)	window size (138)	estimator	data definition
ARIMA	2.666 ***	2.097 ***	2.470 ***	2.251 ***	2.462 ***
OSBM	1.424 ***	1.304 ***	1.397 **	1.131	1.430 ***
USBM	1.173 *	1.074	1.830 **	1.110	1.135
FSBM	1.917	1.048	2.310 ***	1.659	1.837
Average of Single BM	1.089 *	1.005	1.230 **	1.313	1.063
ASGR	1.031	1.096 **	1.073	1.055	1.050
ASLWE	1.061	1.055	1.000	1.000	1.102
ASWL	1.001	1.081 **	1.023	1.005	1.040
Average of the Aggregation of Sectors	1.015	1.068 *	1.026	1.010	1.050
Average of all models (excl. ARIMA)	1.000	1.000	1.088	1.036	1.000

(1) The performance is measured by the root mean squared forecasting error (RMSE) of the growth rates with respect to the previous month (if not otherwise indicated). (2) Number of months. (3) LA (latest available vintage) or RT (real-time vintages). (4) For a description, see Table 1. (5) Ratio of the RMSE in the previous column with respect to the baseline. In this context, for the real-time case the Giacomini and White test of equal predicting ability cannot be applied. (6) *, ** and *** respectively reject at 10%, 5% and 1% the null of the GW test.

Table 3

2-MONTHS AHEAD PREDICTIVE ABILITY OF ALTERNATIVE FORECASTING METHODS OVER THE 2003.2 – 2007.12 PERIOD (59 MONTHS) [1]

	Baseline	Change in the: window size				estimator		data definition	
Size of the rolling window [2]:	84	48		138		84		84	
Estimation method:	OLS	OLS		OLS		SUR		OLS	
Type of data [3]:	LA	LA		LA		LA		RT	
Models [4]:	*RMSE*	*RMSE*	*ratios [5, 6]*	*RMSE*	*ratios [5, 6]*	*RMSE*	*ratios [5, 6]*	*RMSE*	*ratios [5, 6]*
ARIMA	1.88	1.90	1.007	1.93	1.026*	1.88	0.996	1.78	0.943
OSBM	1.17	1.51	1.290***	1.30	1.105	1.22	1.040	1.17	0.998
USBM	1.27	1.48	1.165**	1.90	1.499**	1.34	1.056	1.24	0.981
FSBM	1.36	1.39	1.024	1.81	1.327***	1.38	1.015	1.35	0.989
Average of Single BM	1.10	1.32	1.210**	1.26	1.153*	1.27	1.156***	1.07	0.979
ASGR	1.01	1.30	1.283***	1.15	1.132**	1.19	1.176***	1.01	0.995
ASLWE	0.98	1.18	1.203**	1.05	1.074	1.12	1.143**	1.00	1.017
ASWL	1.00	1.27	1.277***	1.13	1.133**	1.17	1.168***	1.01	1.010
Average of the Aggregation of Sectors	0.99	1.24	1.258***	1.10	1.117*	1.15	1.167***	1.00	1.008
Average of all models (excl. ARIMA)	1.02	1.26	1.237***	1.16	1.138*	1.19	1.175***	1.00	0.989

RMSE ratios with respect to the best performance of the column [6]

	Baseline	48	138	SUR	RT
ARIMA	1.922 ***	1.610 ***	1.838 ***	1.676 **	1.782 ***
OSBM	1.198 *	1.284 ***	1.233 *	1.091	1.176
USBM	1.292 ***	1.252 ***	1.805 ***	1.194 **	1.246 **
FSBM	1.390 ***	1.183 *	1.718 ***	1.235 ***	1.352 ***
Average of Single BM	1.119	1.126 **	1.201 **	1.131 **	1.077
ASGR	1.036	1.105 ***	1.092 **	1.066 **	1.014
ASLWE	1.000	1.000	1.000	1.000	1.000
ASWL	1.018	1.080 **	1.074 **	1.041	1.010
Average of the Aggregation of Sectors	1.037	1.054 **	1.049 *	1.030 *	1.000
Average of all models (excl. ARIMA)	1.039	1.066 *	1.099 *	1.067 **	1.008

([1]) The performance is measured by the root mean squared forecasting error (RMSE) of the growth rates with respect to the previous month (if not otherwise indicated). ([2]) Number of months. ([3]) LA (latest available vintage) or RT (real-time vintages). ([4]) For a description, see Table 1. ([5]) Ratio of the RMSE in the previous column with respect to the baseline. In this context, for the real-time case the Giacomini and White test of equal predicting ability cannot be applied. ([6]) *, ** and *** respectively reject at 10%, 5% and 1% the null of the GW test.

Table 4

3-MONTHS AHEAD PREDICTIVE ABILITY OF ALTERNATIVE FORECASTING METHODS OVER THE 2003.2 – 2007.12 PERIOD (59 MONTHS) [1]

	Baseline	Change in the: window size				estimator		data definition	
Size of the rolling window [2]:	84	48		138		84		84	
Estimation method:	OLS	OLS		OLS		SUR		OLS	
Type of data [3]:	LA	LA		LA		LA		RT	
Models [4]:	*RMSE*	*RMSE*	*ratios [5, 6]*	*RMSE*	*ratios [5, 6]*	*RMSE*	*ratios [5, 6]*	*RMSE*	*ratios [5,6]*
ARIMA	2.09	2.15	1.030	2.14	1.024*	2.08	0.996	1.99	0.951
OSBM	1.44	1.84	1.277***	1.62	1.123**	1.54	1.067	1.43	0.993
USBM	1.68	1.78	1.059	2.53	1.505***	1.76	1.047	1.64	0.975
FSBM	1.80	1.78	0.988	2.23	1.239***	1.87	1.041	1.78	0.991
Average of Single BM	1.45	1.60	1.106	1.67	1.158*	1.65	1.143***	1.41	0.974
ASGR	1.30	1.48	1.140**	1.48	1.138**	1.57	1.204***	1.28	0.982
ASLWE	1.21	1.32	1.098	1.35	1.115**	1.45	1.199***	1.21	1.002
ASWL	1.27	1.44	1.136**	1.45	1.145**	1.53	1.210***	1.26	0.996
Average of the Aggregation of Sectors	1.24	1.40	1.128*	1.42	1.140**	1.51	1.210***	1.24	0.993
Average of all models (excl. ARIMA)	1.32	1.46	1.107	1.51	1.147**	1.57	1.188***	1.29	0.980

RMSE ratios with respect to the best performance of the column [6]

	Baseline	48	138	SUR	RT
ARIMA	1.732 ***	1.625 ***	1.590 ***	1.439 **	1.645 ***
OSBM	1.193 **	1.387 ***	1.201	1.061	1.182 *
USBM	1.391 ***	1.342 ***	1.877 ***	1.214 ***	1.354 ***
FSBM	1.490 ***	1.342 ***	1.655 ***	1.293 ***	1.475 ***
Average of Single BM	1.198 ***	1.207 ***	1.243 **	1.142 ***	1.165 **
ASGR	1.078 **	1.119 **	1.099 ***	1.082 ***	1.057
ASLWE	1.000	1.000	1.000	1.000	1.000
ASWL	1.050 *	1.087 **	1.078 ***	1.060 **	1.044
Average of the Aggregation of Sectors	1.031	1.059 **	1.053 ***	1.040 **	1.022
Average of all models (excl. ARIMA)	1.091 **	1.101 **	1.123 **	1.081 **	1.067

([1]) The performance is measured by the RMSE of the growth rates with respect to the previous month (if not otherwise indicated). ([2]) Number of months. ([3]) LA (latest available vintage) or RT (real-time vintages). ([4]) For a description, see Table 1. ([5]) Ratio of the RMSE in the previous column with respect to the baseline. For the real-time case the GW test cannot be applied. In this context, for the real-time case the Giacomini and White test of equal predicting ability cannot be applied. ([6]) *, ** and *** respectively reject at 10%, 5% and 1% the null of the GW test.

Table 5

PREDICTIVE ABILITY OF ALTERNATIVE FACTOR-BASED FORECASTING METHODS OVER THE 2003.2 - 2007.12 PERIOD (59 MONTHS); WINDOW SIZE OF 84 MONTHS [1]

Forecast horizon:	1-month		2-months		3-months	
Models:[2]	*RMSE*	*Ratios* [3]	*RMSE*	*ratios* [3]	*RMSE*	*ratios* [3]
ARIMA	1.39	1.000	1.88	1.000	2.09	1.000
FA-ARDL (110 bivariate models)	1.34	0.965	1.73	0.918	1.71	0.819
FA-ARDL (80 bivariate models)	1.31	0.944	1.72	0.913	1.71	0.817
SW-D	0.94	0.681**	1.38	0.730*	1.63	0.778*
SW-S	0.94	0.681**	1.46	0.773	1.55	0.740**
FHLR-F	0.93	0.667**	1.30	0.690**	1.64	0.784**
FHLR-O	0.92	0.666**	1.31	0.697**	1.62	0.776**
Average of Factor-based Models	0.91	0.658**	1.30	0.689**	1.58	0.754**
FHLR-F (window size = 48)	0.97	0.701**	1.38	0.735*	1.67	0.800
FHLR-F (window size = 138)	0.96	0.692**	1.26	0.670**	1.56	0.745**

([1]) The performance is measured by the root mean squared forecasting error (RMSE) of the seasonally adjusted growth rates with respect to the previous month. Estimation window of 84 months if not otherwise indicated. ([2]) For a description, see Table 1. ([3]) Ratio of the RMSE with respect to ARIMA model; *, ** and *** respectively reject at 10%, 5% and 1% the null of equal predicting ability according to the Giacomini and White test.

Table 6

THE COMPARISON OF ALTERNATIVE FORECASTING APPROACHES

				Fair and Shiller (1990) test outcomes [2]				
1-month ahead	*RMSE*	*Ratio* [1]		*ARIMA*	*BM single*	*BM multiple*	*Factor-based*	*Avg (excl. ARIMA)*
ARIMA	1.39	1.000		-	0.27	0.15	1.49	0.83
BM single-equation	0.57	0.408	***	7.04	-	1.30	7.33	5.15
BM multiple-equations	0.53	0.381	***	8.79	3.23	-	9.52	6.10
Factor-based	0.91	0.658	**	2.04	-0.56	-0.32	-	-7.80
Average (excl. ARIMA)	0.71	0.510	***	3.65	-0.03	-0.12	10.30	-
2-months ahead								
ARIMA	1.88	1.000		-	1.08	0.93	1.39	1.18
BM single-equation	1.10	0.582	***	3.65	-	-0.74	2.48	1.54
BM multiple-equations	0.99	0.524	***	5.67	2.76	-	4.12	3.14
Factor-based	1.30	0.689	**	1.48	-0.01	-0.48	-	-2.42
Average (excl. ARIMA)	1.14	0.604	**	2.60	0.31	-0.56	3.30	-
3-months ahead								
ARIMA	2.09	1.000		-	0.72	0.46	1.09	0.96
BM single-equation	1.45	0.691	**	2.77	-	-1.41	2.23	1.48
BM multiple-equations	1.24	0.595	***	4.23	3.28	-	3.44	2.93
Factor-based	1.58	0.754	*	1.37	-0.13	-0.79	-	-2.21
Average (excl. ARIMA)	1.41	0.674	**	2.15	0.19	-0.94	2.82	-

([1]) *, ** and *** respectively reject at 10%, 5% and 1% the null of equal predicting ability according to the Giacomini and White test. ([2]) We report the White-consistent t-statistics of the estimates of β_R and β_C parameters in the regression:

$$\frac{y_t - y_{t-h}}{y_{t-h}} = \alpha + \beta_R \frac{\hat{y}_{R,t} - y_{t-h}}{y_{t-h}} + \beta_C \frac{\hat{y}_{C,t} - y_{t-h}}{y_{t-h}}$$

, where $\hat{y}_{R,t}$ and $\hat{y}_{C,t}$ are the forecasts of the two models being compared (and respectively listed along the rows and the columns), and h is the forecast horizon (h = 1, 2, 3 months ahead). The null hypothesis is that the forecast of the model *R* (in the row) contains no information relevant to IPI forecast not in the model *C* (in the column). Results are robust to the use of Newey-West t-statistics.

Table A1

THE BM MONTHLY INDICATORS DATA-SET AND THE SPECIFIC MODELS THEY ENTER

Label	Description	Aggregate IPI single-equation BM [1]			Disaggregate BM equations by IPI sub-sector [1]			
AFFEXA	ISAE retail trade survey; current business situation (balances; index, 2000=100)				ConPI			
AFFEXP	ISAE retail trade survey; expectation on business evolution (balances; index, 2000=100)				ConPI			
CLICOM	ISAE retail trade survey; confidence index (2000=100)						InvPI	
CLIMA	ISAE Consumer sentiment index (2000=100)				ConPI		InvPI	EnePI
CPCCNZ	Consumer price index (national definition; ISTAT)							EnePI
FIDINTN	ISAE Industrial survey; confidence index, intermediate goods (2000=100)					IntPI		
FIDINVN	ISAE Industrial survey; confidence index, equipment goods (2000=100)						InvPI	
IENEL2	Electricity consumption in sector 2 (Milan), index number (ENEL)		USBM		ConPI		InvPI	
IENEL3	Electricity consumption in sector 3 (Venice), index number (ENEL)							EnePI
IENEL4	Electricity consumption in sector 4 (Florence), index number (ENEL)		USBM				InvPI	
IENEL5	Electricity consumption in sector 5 (Rome), index number (ENEL)					IntPI		
IENEL6	Electricity consumption in sector 6 (Naples), index number (ENEL)						InvPI	
IENEL7	Electricity consumption in sector 7 (Palermo, Sicily), index number (ENEL)						InvPI	
IENEL8	Electricity consumption in sector 8 (Cagliari, Sardinia), index number (ENEL)						InvPI	
IENEL9	Electricity consumption in sector 9 (railroad transport.), index number (ENEL)			FSBM				EnePI
IENELN	Northern electricity consumption (sum over the sector s1 to 4)	OSBM		FSBM		IntPI		
IENELS	Southern electricity consumption, (sum over the sectors 6 to 8)							EnePI
ITEIMG	Employment index (purchasing managers index; NTC-REUTERS)					IntPI		
ITNOMG	New orders index (purchasing managers index; NTC-REUTERS)			FSBM	ConPI			
ITPGMG	Stock of purchased goods (purchasing managers index; NTC-REUTERS)			FSBM				EnePI
NLOINV	ISAE Industrial survey; orders actual tendency, equipment goods (balances, 2000=100)		USBM					
NLPIMP	ISAE Industrial survey; production actual tendency (balances, 2000=100)			FSBM				
NSCIMP	ISAE Industrial survey; stocks levels (balances, 2000=100)			FSBM				
ORDISTAT	Orders level in real terms, manufacturing sector (ISTAT)		USBM					
RETTGEN	Coefficient for trading day adjustment	OSBM	USBM	FSBM	ConPI	IntPI	InvPI	EnePI
SCORTE	ISAE retail trade survey; finished goods stocks (balances; index, 2000=100)				ConPI			
TASSO	Short-term interest rate on bank loans, total economy (Bank of Italy)						InvPI	
TEMPN	Temperature, monthly average in Turin, Milan, Venice and Florence	OSBM	USBM	FSBM	ConPI	IntPI	InvPI	EnePI
TEMPS	Temperature, monthly average in Naples, Bari, Palermo and Cagliari							EnePI
TOIMPCN	ISAE; average of confidence indexes in the construction and retail trade surveys (indices, 2000=100)			FSBM				

([1]) Details are in Table 1, points B and C respectively.

Tab. A2

FACTOR-BASED MODEL INDICATORS

No.	Description	tcode
(1) IPI and orders, ISTAT, [31]		
1	*Consumer goods*	2
2	*Consumer Durable*	2
3	*Consumer Non-durable*	2
4	*Capital Goods*	2
5	*Intermediate Goods*	2
6	*Energy*	2
7-29	*A0215 –A0240 (2-digit Ateco branches)*	2
30	*New Order to manufacturing*	2
31	*Total orders (volume),* BI est. on ISTAT	2
(2) Domestic and Intern. Prices, varia, [4]		
32	*Consumer Prices, ISTAT*	2
33	*Producer Prices, total industry,* ISTAT	2
34	*Oil price index,* IMF	2
35	*Competitiveness indicator for Italy,* BI	1
(3) Rates, money and stock market ind., BI, [9]		
36	*Share price (MIB) index*	2
37	*Treasury bill rate*	3
38	*Money market rate*	3
39	*Govt bond yield (long term)*	3
40	*Govt bond yield (medium term)*	3
41	*Bond yield secondary market*	3
42	*Money supply M1*	2
43	*Money supply M2*	2
44	*Money supply M3*	2
(4) Consumer Survey, ISAE, [6]		
45	*- Consumer Confidence indicator*	1
46	*- Situation next 12 months*	1
47	*- Unemployment next 12 months*	1
48	*- Prices next 12 months*	1
49	*- Prices last 12 months*	1
50	*- Savings at present*	1
(5) Retail Survey, ISAE, [5]		
51	*- Retail Confidence indicator*	1
52	*- Future business situation*	1
53	*- Current business situation*	1
54	*- Inventory level*	1
55	*- Retail sales volume, BI est. on ISTAT*	2
(6) Construction survey, ISAE, [3]		
56	*- Construction confidence indicator*	1
57	*- Order book position*	1
58	*- Price expectations*	1
(7) Manufacturing survey, ISAE, [40]		
59	*Confidence indicator (total)*	1
60	*Total order (total)*	1
61	*Domestic order (total)*	1
62	*Order from abroad (total)*	1
63	*Production levels (total)*	1
64	*Order expectations (total)*	1
65	*Production expectations (total)*	1
66	*Price expectation (total)*	1
67	*General ec. situation expect. (total)*	1
68	*Inventory levels (total)*	1
69	*Confidence indicator (capital)*	1
70	*Total order (capital)*	1
71	*Domestic order (capital)*	1
72	*Order from abroad (capital)*	1
73	*Production levels (capital)*	1
74	*Order expectations (capital)*	1
75	*Production expectations (capital)*	1
76	*Price expectations (capital)*	1
77	*General ec. situation expect. (capital)*	1
78	*Inventory levels (capital)*	1
79	*Confidence indicator (consumer)*	1
80	*Total order (consumer)*	1
81	*Domestic order (consumer)*	1
82	*Order from abroad (consumer)*	1
83	*Production levels (consumer)*	1
84	*Order expectations (consumer)*	1
85	*Production expectations (consumer)*	1
86	*Price expectations (consumer)*	1
87	*General ec. situation expect. (consumer)*	1
88	*Inventory levels (consumer)*	1
89	*Confidence indicator (intermediate)*	1
90	*Total order (intermediate)*	1
91	*Domestic order (intermediate)*	1
92	*Order from abroad (intermediate)*	1
93	*Production levels (intermediate)*	1
94	*Order expectations (intermediate)*	1
95	*Production expect. (intermediate)*	1
96	*Price expectations (intermediate)*	1
97	*Gen. ec. situation exp. (intermediate)*	1
98	*Inventory levels (intermediate)*	1
(8) Electricity Consumption, ENEL, [9]		
99	*Turin district*	2
100	*Milan district*	2
101	*Venice district*	2
102	*Rome district*	2
103	*Florence district*	2
104	*Neaples district*	2
105	*Cagliari district*	2
106	*Palermo district*	2
107	*Rail-network consumption*	2
(9) Purchasing Manuf. Index, NTC-Reuters, [3]		
108	*Employment Index*	1
109	*New Orders Index*	1
110	*Stock of Purchased goods Index*	1

Note: Indicators are classified in 9 groups; bold characters indicate, respectively, the description of each group, the data providers (see below), and (in squared brackets) the number of indicators in each group. Data providers: ISTAT = Italian Statistical Agency, BI = Bank of Italy, ISAE = Institute for Studies and Economic Analysis, IMF = International Monetary Fund, ENEL = Italian Electric Power Company. *tcode* is the transformation code: 1 = no transformation, 2 = growth rate over the same month of the previous year (i.e. Δlog), 3 = monthly first difference. All series have been cleaned from outlier and quantitative indicators have been adjusted for trading day variations.