Do financial variables help forecasting inflation and real activity in the EURO area? *

Mario Forni
Dipartimento di Economia Politica
Universitá di Modena and CEPR

Marc Hallin
ISRO, ECARES, and Département de Mathématique
Université Libre de Bruxelles

Marco Lippi
Dipartimento di Scienze Economiche
Universitá di Roma La Sapienza

and

Lucrezia Reichlin
ECARES, Université Libre de Bruxelles and CEPR

September 18, 2001

Abstract

This paper uses a large data set of monthly time series both real and nominal for the main countries of the euro area to evaluate the role of financial variables in forecasting aggregate inflation and industrial production. The panel contains 725 variables which we organize in five blocks: sectoral and national industrial production, sectoral and national prices, national money aggregates, financial variables and miscellaneous leading variables. We use the dynamic factor model proposed by Forni, Hallin, Lippi and Reichlin (2000) to establish leading properties of all variables in the panel with respect to inflation and industrial production. We then construct aggregates of leading variables by each block and establish the marginal role of the financial aggregate in the forecasting equations. We find that financial variables help forecast inflation but not industrial production.

JEL subject classification : C13, C33, C43.
Key words and phrases : Dynamic factor models, dynamic principal components, business cycle, forecasting, financial variables.

*Research supported by La Fondation de la Banque de France

1 Introduction

There is a large literature in finance and macroeconomics which suggest that financial variables are good predictors of inflation and real economic activity. Empirical evidence, however, is mixed and results are not robust with respect to model specification, sample choice and forecast horizon (for an excellent review of the empirical literature, see Stock and Watson, 2000). This is clearly a puzzle for economic theory and one that is worth investigating.

Our paper exploits the information from a large panel of monthly time series for the six main economies of the euro area. The panel contains industrial production data (by sectors and nations), prices (by sectors and nations), money aggregates (by nations), a variety of potentially leading variables (survey data and others) and financial variables such as interest rates (nominal and real, for different countries and maturities), spreads and exchange rates.

The key idea of the paper is to evaluate whether by pooling information from a broad group of financial variables we can obtain good predictions for the Euro-area industrial production and consumer price indexes. In other words, instead of evaluating the predictive content of single financial variables, we evaluate the predictive content of averages of many of such variables, suitably selected. Forecasting performances at different time horizons are evaluated through an out-of-sample simulation exercise.

The motivation of our strategy is not far from Stock and Watson (2000), who have recently suggested that, by combining forecasts from poorly performing bivariate models, the predictive power of financial variables is rescued. Here, instead of combining bivariate forecasts obtained with different financial variables as predictors, we directly combine information. By pooling forecasts, poor performances are averaged out; by pooling predictors, as we do, noisy informations are averaged out.

Our reference model is the generalized dynamic factor model proposed and discussed in Forni, Hallin, Lippi and Reichlin (2000, 2001a, 2001b) and Forni and Lippi (2001), which is specifically designed to handle large panels of dynamically related time series.

In this model, each time series in the panel is represented as the sum of two components: a component which captures most of the multivariate correlation (the *common component*) and a component which is poorly cross-sectionally correlated (the *idiosyncratic component*). The common components in the cross section have, so to speak, 'reduced rank', meaning they are all driven by a few common shocks. Such low dimensionality implies that the common components can be consistently estimated and forecasted on the basis of few regressors, i.e. the present and the past of the common shocks, or linear combinations of these shocks.

Unfortunately, the common shocks are not observable. Here we try to capture the relevant information by constructing averages of the variables in the panel. The key idea behind this procedure is simply that averaging kills the idiosyncratic components, which are almost uncorrelated, because of an obvious large-number effect.

Cleaning from the idiosyncratic noise to obtain pure aggregate information, however, is not the only thing we need. By aggregating variables which are too 'lagging' with respect to the target variables we could end up with information which, despite being free from the idiosyncratic component, is not enough up to date to be useful in prediction. To solve this problem we analyze the leading-lagging relations between the common components in the panel by estimating their dynamic covariance structure and select only the leading variables before aggregation.

The whole procedure is then as follows. First, we establish the number of common shocks and estimate the covariance structure of the common components as suggested in Forni, Hallin, Lippi and Reichlin (2000). Second, we use this information to identify and select the 'leading' variables. Third, we take the simple average of these leading variables by category, i.e. the average of leading financial, money, price, industrial production and miscellaneous variables. Finally, we use the present and the past of these aggregates to predict our target variables and evaluate the performance of the financial predictor.

The paper is organized as follows. In Section 2 we introduce the model and provide examples. In Section 3 we briefly illustrate the data set and the data treatment. Section 4 reports detail of the forecasting exercise and the empirical results. Section 5 concludes.

2 Theory

2.1 The model

We assume that our i-th time series, suitably transformed, is a realizations from a zero mean, wide-sense stationary process x_{it}. Each process in the panel is thought of as an element from an infinite sequence, so that $i = 1, \dots, \infty$. Moreover, all of the x's are co-stationary, i.e. stationarity holds for the n-dimensional vector process $\mathbf{x}_{nt} = (x_{1t}, ..., x_{nt})'$, for any n.

Each variable in the panel follows the relation

$$x_{it} = \chi_{it} + \xi_{jt} = \mathbf{b}_i(L)\mathbf{u}_t + \xi_{it} = \sum_{h=1}^{q} b_{ih}(L)u_{it} + \xi_{it} \tag{2.1}$$

where χ_{it} is the common component, $\mathbf{u}_t = (u_{1t}, \dots, u_{qt})'$ is the q-dimensional vector of the common shocks, which are unit variance white noises mutually orthogonal at all leads and lags, $\mathbf{b}_i(L) = b_{i1}(L), \dots, b_{is}(L)$ is a row vector of square-summable functions in the lag operator, and the idiosyncratic component ξ_{it} is orthogonal to $\mathbf{u}_{t-k}$ for any k and i.

Moreover, we assume that (a) the q non-zero eigenvalues of the spectral-density matrix of $\boldsymbol{\chi}_{nt} = (\chi_{1t}, \dots, \chi_{nt})'$, say $\lambda_1^\chi(\theta), \dots, \lambda_q^\chi(\theta)$, go to infinity as $n \to \infty$ a.e. on $[-\pi, \pi)$; (b) the largest eigenvalue of the spectral-density matrix of $\boldsymbol{\xi}_{nt} = (\xi_{1t}, \dots, \xi_{nt})'$, say $\lambda_1^\xi(\theta)$, is smaller than a real number λ a.e. on $[-\pi, \pi)$ for any n.

For detailed comments on the model we refer to Forni, Hallin, Lippi and Reichlin (2000). Here we shall limit ourselves to a few remarks.

First, the model generalizes the traditional dynamic factor model of Sargent and Sims (1977) and Geweke (1977), in that the idiosyncratic components are not necessarily orthogonal to each other. This feature is shared with the static approximate factor model of Chamberlain and Rothschild (1983), for which our model provides a dynamic generalization.

Second, the orthogonality assumption is replaced here by assumptions (a) and (b), which, loosely speaking, impose a minimum amount of cross-correlation for the common components and a maximum amount for the idiosyncratic components. Condition (b), in particular, includes the classic mutual orthogonality as a specific case and guarantees that suitable linear combinations of the idiosyncratic components, like the simple average, vanish as n goes to infinity. This property will be used below.

Finally, notice that the distributed lags in front of the common shocks u_{kt} are quite general. Different variables in the cross-section may react to the same shock with different signs and time delays, giving rise to a wide range of dynamic behaviors. In particular, variables may be 'leading' or 'lagging' in a sense that will be clarified in the sequel.

2.2 A stylized example

To convey the intuition of our forecasting procedure we introduce the highly stylized example

$$x_{it} = \chi_{it} + \xi_{it} = u_{t-s_i} + \xi_{it},$$

where the spectral density matrix of the vector $(\xi_{1t}\ \xi_{2t}\ \cdots\ \xi_{nt})$ is equal to $\mathbf{I}_n$. This is the case in which the variables ξ are strictly idiosyncratic (i.e. mutually orthogonal). We also assume that s_i is equal to zero, one, or two, this being a stylization of real situations in which some of the variables are lagging ($s_i = 2$), some are leading ($s_i = 0$), with respect to a central group of variables ($s_i = 1$), let us call them coincident.

Let us see what happens in this example when taking a simple cross-sectional average of a subset S of the x_{it}'s. We get

$$X_t = au_t + bu_{t-1} + cu_{t-2} + \sum_{i \in S} \xi_{it}/n_S,$$

where a, b and c are respectively the percentages of leading, coincident and lagging variables in S (so that $a + b + c = 1$) and n_S denotes the total number of variables in the set.

The first thing to stress is that the variance of the idiosyncratic component $\sum_{i \in S} \xi_{it}/n_S$ is $n_S/n_S^2 = 1/n_S$, so that the component itself vanishes as $n_S \to \infty$. Different linear combinations like weighted averages would have the same effect. In this paper we do not make any attempt to optimize the weights of the aggregate predictor and use a simple average (for a discussion of the optimal weights see Forni, Hallin, Lippi and Reichlin, 2001b).

Now assume that n_S is large, so that the idiosyncratic term is negligible, and that we want to use $X_t = au_t + bu_{t-1} + cu_{t-2}$ in order to predict, say, x_{1t+1}, which we assume to be coincident. Moreover, let us concentrate on the prediction of the common part $\chi_{1t+1} = u_t$ (we shall come back to the prediction of the idiosyncratic part). Finally, let us assume for illustrative purposes that we do not make use of the lags of X_t in prediction and use instead the static projection

$$\chi_{1t+1} = u_t = AX_t + \Omega_t.$$

Obviously

$$A = \frac{a}{a^2 + b^2 + c^2},$$

while the variance of the residual has a maximum for $a = 0$ and a minimum of zero when $b = c = 0$. In other words, considering the common components $\chi_{it} = u_{t-s_i}$, none of them has any power to predict itself or the common component of leading variables. However, the common components of the leading variables help predicting the common component of the coincident (and lagging) variables.

This discussion suggest to identify the leading variables before aggregation. In the following subsection we explain our procedure to select the leading variables. Here we conclude with two remarks.

First, notice that in the example, the dimension of the common factor space is one, so that a single aggregate is sufficient to capture the relevant information. In a more general model, the number of aggregates should match q. To identify q we follow Forni, Hallin, Lippi and Reichlin (2000), i.e. we compute the principal component series (see the Appendix) and retain only principal components explaining more than a percentage $\bar{p}$ of the total variance in the system. Precisely, with reference to equation (A.1) in the Appendix, we require $p_q > \bar{p}$ and $p_{q+1} < \bar{p}$.

Second, notice that our objective is to forecast a variable which contains an idiosyncratic term. Such term is a white noise in the example, but in general is autocorrelated and therefore can be predicted. Since multivariate information only help to forecast the common component, we specify the forecasting equation as a projection on the present and past of the q leading aggregates *plus lags of the dependent variable.* Such lags, when controlling for the common factor space, should be useful in predicting the idiosyncratic term.

2.3 Identifying the leading variables

In order to identify the leading variables we can follow different strategies. First, we can look at the contemporaneous and lagged covariances between the common components. Second, we can study the phase shift in the polar form of the cross-spectral densities. Third, we can analyze the Granger causality relations. In the stylized example above, all these strategies are theoretically equivalent. In the general case, however, things are not so simple, because we have more than just one shock and the response functions are more complicated, so that the very concept of 'leading variable' is problematic. Different strategies imply different concepts of what is 'leading' and produce different results.

As explained below, we concentrate here on the cross-spectra and the implied time phase lead. The main reason is that we can estimate consistently and quite directly the spectral-density matrix of the common components as explained in the Appendix. An alternative possibility would be testing for Granger-causality of each χ_{it} toward χ_{1t}. However, this would imply estimating the variables χ_{it}, not only their joint spectral density, and perform a test which entails a generated regressors problem.

Let us now describe our criterion in more detail. Consider the estimated cross-spectral density of each common component with respect to χ_{1t}. In general the cross spectral density between two variables a_t and b_t can be expressed, in its 'polar form', as $S_{hj}(\theta) = A_{hj}(\theta)e^{-i\phi_{hj}(\theta)}$ where $A_{hj}(\theta)$ is the 'amplitude' and $\phi_{hj}(\theta)$ is the 'phase'. The phase $\phi_{hj}(\theta)$ measures the angular shift between the cosine waves of a_t and b_t at frequency θ, while $\phi_{hj}(\theta)/\theta$ measures the time shift. Then, let the phase angle shift of χ_{jt} with respect to χ_{1t} be $\phi_j(\theta)$, $-\pi < \theta \leq \pi$. At frequency zero, the phase may be either 0 or π depending on whether long-run correlation is positive or negative. Following Granger and Hatanaka (1964, ch. 12), we interpret $\phi_j(0) = \pi$ as indicating that χ_{jt} is in 'phase opposition' and define the new series of interest as

$$\omega_{jt} = \begin{cases} \chi_{jt} & \text{if } \phi_j(0) = 0 \\ -\chi_{jt} & \text{if } \phi_j(0) = \pi \end{cases}$$

Now we classify the resulting time series as being leading, coincident or lagging according to their phase delay with respect to χ_{1t}. We compute the phase angle shift of ω_{jt}, $j = 1, \ldots, n$, with respect to χ_{1t}, at a typical business cycle frequency, say $\theta^* > 0$, and classify x_{jt} as coincident if $|\phi_j(\theta^*)|$ is smaller than a prespecified value τ, leading if $\phi_j(\theta^*) < -\tau$ and lagging if $\phi_j(\theta^*) > \tau$.

Of course, this is equivalent to computing the 'time delay' $\phi_j(\theta^*)/\theta^*$ and compare it with τ/θ^*. Notably, if θ^* is sufficiently close to 0, the estimate of the time delay $\phi_j(\theta^*)/\theta^*$ can be regarded as an estimate of the derivative of the phase angle at $\theta = 0$. This is interesting in that such derivative is equal to the 'mean lag', which is a well-known time-domain statistic measuring the 'delay' of a time series.

To conclude this section, note that, in the stylized example above, the phase angle shift of the leading variables with respect to χ_{1t} is $-\theta$, corresponding to a mean lag of -1 (mean lead of 1); the same, with opposite signs, for the lagging variables.

3 Data set and data treatment

The database used in this paper has been constructed by the Banca d'Italia research department within the Bank of Italy-CEPR project (a detail description is in Altissimo et al., 2001). We are using 725 monthly time series on key aggregate and sectoral variables for the six main economy of the euro area—Germany, France, Italy, Spain, The Netherlands, Belgium—and for, when available, the euro area as a whole. The time span is 1985:1-2000:6.

For the purpose of this paper we have organized the data in five blocks:

- block 1: 155 financial variables (interest rates, nominal and real, spreads and exchange rates);
- block 2: 18 money aggregates (money stocks for different countries);
- block 3: 138 industrial production variables (indexes for different countries and industrial sectors);
- block 4: 135 price variables (producer price indexes and consumer price indexes);

- block 5: 279 miscellaneous variables (European Commission surveys; national Institutes surveys and other variables typically used for short term analysis).

The variables were selected so as to satisfy two requirements: one concerning the length of the series and the other their homogeneity over time and across countries.

We removed outliers from each series using Tramo, a procedure developed by Gomez and Maravall (1999); in particular we focused on transitory changes, level shifts and additive outliers. The same procedure allowed to adjust for working days effects, whenever requested. We did not remove seasonality. To induce stationarity we took first log difference for industrial productions, financial series, monetary aggregates, prices, business and household survey responses and most interest rates; *real* interest rates and the spreads between long and short term nominal interest rates did not need any transformation.

The series were normalized subtracting their mean and then dividing for their standard deviation. This standardization is necessary to avoid overweighting series with large variance when estimating the spectral density.

4 Empirical results

As the first step of our empirical exercise we identified the number of common factors q as explained in Section 2.2, by requiring a minimal contribution $\bar{p} = 10\%$ to the explained variance. We found $q = 4$.

Then we estimated the spectral-density matrix of the common components as explained in the Appendix and used this matrix to obtain an estimate of the time phase shift of all the common components with respect to European IP and CPI indexes.

Having the time shift we classified the variables as leading, coincident and lagging by following the procedure in Section 2.3. We focused on the frequency $\theta^* = \pi/36$, corresponding to a period of six years, and set $\tau = \pi/36$, so that the critical time phase shift to define the leading variables is one month.

Table 5.1: Leading-lagging relations

	with respect to the IP index		*with respect to the CPI index*	
Blocks	% of leading variables	average time lead	% of leading variables	average time lead
Ind. production	0.57	1.86	0.24	-4.03
Prices	0.59	3.20	0.50	-0.88
Finance	0.15	-6.40	0.83	8.67
Money	0.50	0.59	0.44	-0.52
Other	0.70	3.12	0.36	-1.35

Table 5.1 describes the resulting leading-lagging relations. The first column shows the percentage of variables of each block which are leading with respect to the industrial production index. The second column shows the average time lead. The third and

fourth columns report the same information with respect to the consumer price index. It is worth noting that most of the financial variables are leading with respect to prices and the average time lead is large (about eight months). By contrast, the variables belonging to the other blocks are mostly coincident or slightly lagging. This suggests that the financial variables can be useful in prediction, as in fact is the case (see below). Note also that the behavior of the financial variables with respect to industrial production is puzzling, since they are mostly lagging: a result which seems at odds with the general picture resulting from the table.

We than proceeded to the aggregation step and constructed the following predictors:

- z_{1t}: simple average of financial leading variables (block 1);
- z_{2t}: simple average of money leading variables (block 2);
- z_{3t}: simple average of industrial production leading variables (block 3);
- z_{4t}: simple average of price leading variables (block 4);
- z_{5t}: simple average of miscellaneous leading variables (block 5);
- Z_t: simple average of all but financial leading variables.

Finally we used these predictors to estimate the following models, where x_t denotes the target variable:

$$x_{t+h} = \alpha_0^h(L)x_t + \epsilon_{0t}^h \tag{M0}$$

$$x_{t+h} = \alpha_1^h(L)x_t + \beta_{11}^h(L)z_{1t} + \gamma_1^h(L)Z_t + \epsilon_{1t}^h \tag{M1}$$

$$x_{t+h} = \alpha_2^h(L)x_t + \beta_{12}^h(L)z_{1t} + \epsilon_{2t}^h \tag{M2}$$

$$x_{t+h} = \alpha_3^h(L)x_t + \gamma_3^h(L)Z_t + \epsilon_{3t}^h \tag{M3}$$

$$x_{t+h} = \alpha_4^h(L)x_t + \sum_{k=1}^{5} \beta_{k4}^h(L)z_{kt} + \epsilon_{4t}^h \tag{M4}$$

$$x_{t+h} = \alpha_5^h(L)x_t + \sum_{k=2}^{5} \beta_{k5}^h(L)z_{kt} + \epsilon_{5t}^h \tag{M5}$$

$$x_{t+h} = \alpha_6^h(L)x_t + \beta_{16}z_{1t} + \sum_{k=3}^{5} \beta_{k6}^h(L)z_{kt} + \epsilon_{6t}^h \tag{M6}$$

$$x_{t+h} = \alpha_7^h(L)x_t + \sum_{k=3}^{5} \beta_{k7}^h(L)z_{kt} + \epsilon_{7t}^h \qquad (M7)$$

We considered the forecasting horizons $h = 1, 3, 6, 12$ and different time spans, starting from the sample 1985:1-1995:12 with $T = 132$ observations, and ending with $T = 186-h$ observations. We first estimated by OLS the pure autoregressive model $M0$ with maximum lags $l = 0, \ldots, 13$, thus obtaining the forecasts x_{0T+h}^{Tl}, $h = 1, 3, 6, 12$, $T = 132, \ldots, 186-h$. Then we computed the mean square errors $MSE_{0h}^{l} = \sum_{T=132}^{186-h}(x_{T+h} - x_{0T+h}^{Tl})^2/(54-h)$. Finally we retained the dynamic specification l^* minimizing the MSE. Models $M1$-$M7$ were estimated with l^* lags for the autoregressive part and with the maximum lag $g = 0, \ldots, 13$ for all the auxiliary regressors, thus obtaining the forecasts $x_{k,T+h}^{Tg}$, $h = 1, 3, 6, 12$, $T = 132, \ldots, 186 - h$, $k = 1, \ldots, 7$. Also in this case we computed $MSE_{kh}^{g} = \sum_{T=132}^{186-h}(x_{T+h} - x_{k,T+h}^{Tg})^2/(54 - h)$ and retained the dynamic specification g^* minimizing the MSE.

Table 5.2: Results for models $M0 - M3$

Forecasting horizon	model $M0$ simple AR	model $M1$ Z_t, z_{1t}	model $M2$ only z_{1t}	model $M3$ only Z_t
CPI index				
$h = 1$	0.445 (11)	0.410 (2)	0.416 (3)	0.449 (1)
$h = 3$	0.491 (9)	0.474 (3)	0.487 (3)	0.501 (0)
$h = 6$	0.593 (6)	0.573 (0)	0.574 (0)	0.606 (0)
$h = 12$	1.096 (1)	0.881 (4)	1.010 (5)	1.030 (1)
IP index				
$h = 1$	0.591 (11)	0.556 (2)	0.608 (1)	0.540 (1)
$h = 3$	0.920 (12)	0.917 (1)	0.916 (0)	0.947 (0)
$h = 6$	0.836 (8)	0.802 (4)	0.862 (0)	0.857 (0)
$h = 12$	0.910 (2)	0.900 (1)	0.897 (0)	0.887 (2)

Note that the forecasting models were re-estimated for each T (whereas the leading variables were individuated once and for all with the whole sample). The whole procedure (starting from the identification of the leading variables) was repeated twice, for the IP index and the CPI index.

Tables 5.2 and 5.3 show the results, i.e. the minimal $MSE_{0h}^{l^*}$ and $MSE_{kh}^{g^*}$, for the different forecasting horizons and the different models. The maximal lags l^* and g^* are in brackets.

Let us consider first models $M1$, $M2$ and $M3$, shown in Table 5.2. A first observation is that the IP index seems almost unpredictable for $h > 1$. Overall, the inclusion of the leading variables in $M1$ improves predictions with respect to the simple AR model. This is worth noticing, since in most cases univariate autoregressive forecasts for prices and production have proved hard to beat. Comparing columns 3 and 4 it is seen that financial variables perform well only for the CPI index.

Table 5.3: Results for models $M4 - M7$

forecasting horizon	model $M0$ simple AR	model $M4$ all blocks	model $M5$ all but z_{1t}	model $M6$ all but z_{2t}	model $M7$ all but z_{1t} and z_{2t}
CPI index					
$h = 1$	0.445 (11)	0.288 (2)	0.329 (2)	0.286 (2)	0.315 (1)
$h = 3$	0.491 (9)	0.446 (3)	0.455 (3)	0.447 (0)	0.454 (0)
$h = 6$	0.593 (6)	0.544 (1)	0.546 (0)	0.545 (1)	0.548 (1)
$h = 12$	1.096 (1)	0.886 (1)	0.889 (1)	0.890 (1)	0.891 (1)
IP index					
$h = 1$	0.591 (11)	0.592 (1)	0.575 (3)	0.586 (1)	0.566 (1)
$h = 3$	0.920 (12)	0.936 (1)	0.960 (3)	0.930 (1)	0.959 (0)
$h = 6$	0.836 (8)	0.876 (0)	0.848 (0)	0.872 (0)	0.841 (0)
$h = 12$	0.910 (2)	0.909 (0)	0.907 (0)	0.904 (0)	0.906 (0)

Shifting to Table 5.3, we see that enlarging the information space does not help predicting the IP index but causes an important improvement in forecasting prices, particularly for $h = 1$. Our interpretation is that we need at least four aggregates to get a good estimate of the common component, because the common component is driven by four common shocks. Note that in all of the models $M4$, $M5$ and $M6$ we do have four aggregates, but $M4$ and $M6$ perform better than $M5$, particularly for $h = 1$, indicating that the financial variables are important. The F-tests shown in table 5.4 are consistent with these results.

Table 5.4: F-tests for z_{1t} in model $M4$

forecasting horizon	1985:1-1996:6	1985:1-1997:6	1985:1-1998:6	1985:1-1999:6	1985:1-2000:6	5% critical value	10% critical value
CPI index							
$h = 1$	2.15*	2.54*	3.37**	3.66**	4.05**	2.70	2.13
$h = 3$	0.26	0.34	0.62	0.59	0.68	2.46	1.99
$h = 6$	0.30	1.22	1.05	0.95	1.31	3.09	2.35
$h = 12$	2.59*	2.27	2.42*	2.44*	2.80*	3.09	2.35
IP index							
$h = 1$	1.05	0.78	0.83	0.70	0.41	3.09	2.35
$h = 3$	3.81**	3.23**	2.66*	2.01	3.10**	3.09	2.35
$h = 6$	2.30	2.09	2.00	1.47	1.97	3.09	2.35
$h = 12$	0.30	1.32	1.06	1.31	1.30	3.94	2.75

5 Conclusions

We used a large data set, made up by 725 monthly macroeconomic time series concerning the main countries of the Euro area. We used this data to simulate out-of-sample predictions of the Euro area industrial production and consumer price indexes and to evaluate the role of financial variables in forecasting.

Our theoretical reference was the generalized dynamic factor model of Forni, Hallin, Lippi and Reichlin (2000). Our forecasting strategy was to identify the leading-lagging relations between the common components and try to summarize the relevant aggregate information by means of simple averages of leading variables, among which the average of the financial variables.

We found that the financial variables, as expected, are mostly leading with respect to the consumer price index and have an average time lead of about eight months. By contrast, the relation with the industrial production index is puzzling.

Aggregate leading information does not help predicting industrial production, but helps predicting inflation. Inflation forecasts improve substantially when more than one aggregate is included among the regressors. The financial predictor has a non-negligible predictive content both when used alone and jointly with the other aggregates.

References

[1] Altissimo, F., A. Bassanetti, R. Cristadoro, M. Forni, M. Lippi, L. Reichlin and G. Veronese (2001). A real-time coincident indicator for the Euro area. Banca d'Italia. Mimeo.

[2] Brillinger, D.R. (1981). *Time Series : Data Analysis and Theory*. Holden-Day, San Francisco.

[3] Chamberlain, G. and M. Rothschild (1983). Arbitrage, factor structure and mean-variance analysis in large asset markets. *Econometrica* 51, 1305-1324.

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

[5] Forni, M., M. Hallin, M. Lippi, and L. Reichlin (2001a). Coincident and leading indicators for the euro area *The Economic Journal* **101**, 62-75.

[6] Forni, M., M. Hallin, M. Lippi and L.Reichlin (2001b). The Generalized Dynamic Factor Model: one-sided estimation and forecasting. Mimeo.

[7] Forni, M. and M. Lippi (2001). The generalized dynamic factor model : representation theory. Econometric Theory, 17, 1113-1141.

[8] Gomez and Maravall (1999)

[9] Granger, C.W.J. and M. Hatanaka (1964). *Spectral Analysis of Economic Time Series*, Princeton: Princeton University Press.

[10] Sargent, T.J. and C.A. Sims (1977). Business cycle modelling without pretending to have too much *a priori* economic theory. In C.A. Sims, Ed., *New Methods in Business Research*, Federal Reserve Bank of Minneapolis, Minneapolis.

[11] Stock, J.H. and M.W. Watson (1999). Diffusion indexes. Mimeo.

[12] Stock, J.H. and M.W. Watson (2001). Forecasting output and inflation: the role of asset prices. Mimeo.

Appendix: Estimating the spectral-density matrix of the common components

In the first step of our procedure, we estimate the spectral-density matrix of the common components. Here we explain the procedure in detail. We start by estimating the spectral-density matrix of $\mathbf{x}_{nt}$. Let us denote the theoretical matrix by $\mathbf{\Sigma}(\theta)$ and its estimate by $\hat{\mathbf{\Sigma}}(\theta)$. The estimation is accomplished by using a Bartlett lag-window of size $M = 18$, i.e. by computing the sample auto-covariance matrices $\hat{\mathbf{\Gamma}}_k$, multiplying them by the weights $w_k = 1 - \frac{|k|}{M+1}$ and applying the discrete Fourier transform:

$$\hat{\mathbf{\Sigma}}_x(\theta) = \frac{1}{2\pi} \sum_{k=-M}^{M} w_k \cdot \hat{\mathbf{\Gamma}}_k \cdot e^{-i\theta k}.$$

The spectra were evaluated at 72 equally spaced frequencies in the interval $[-\pi, \pi]$, i.e. at the frequencies $\theta_h = \frac{2\pi h}{100}$, $h = -36, \dots, 36$.
Then we performed the dynamic principal component decomposition (see Brillinger, 1981). For each frequency of the grid, we computed the eigenvalues and eigenvectors of $\hat{\mathbf{\Sigma}}(\theta)$. By ordering the eigenvalues in descending order for each frequency and collecting values corresponding to different frequencies, the eigenvalue and eigenvector functions $\lambda_j(\theta)$ and $U_j(\theta)$, $j = 1, \dots, n$, are obtained. The function $\lambda_j(\theta)$ can be interpreted as the (sample) spectral density of the j-th principal component series and, in analogy with the standard static principal component analysis, the ratio

$$p_j = \int_{-\pi}^{\pi} \lambda_j(\theta) d\theta \Big/ \sum_{j=1}^{n} \int_{-\pi}^{\pi} \lambda_j(\theta) d\theta \tag{A.1}$$

represents the contribution of the j-th principal component series to the total variance in the system.

Letting $\mathbf{\Lambda}_q(\theta)$ be the diagonal matrix having on the diagonal $\lambda_1(\theta), \dots, \lambda_q(\theta)$ and $\boldsymbol{U}_q(\theta)$ be the $(n \times q)$ matrix $\begin{pmatrix} U_1(\theta) & \cdots & U_q(\theta) \end{pmatrix}$ our estimate of the spectral density matrix of the vector of the common components $\boldsymbol{\chi}_t = \begin{pmatrix} \chi_{1t} & \cdots & \chi_{nt} \end{pmatrix}'$ is given by

$$\hat{\mathbf{\Sigma}}_\chi(\theta) = \boldsymbol{U}(\theta)\mathbf{\Lambda}(\theta)\tilde{\mathbf{U}}(\theta) \tag{.2}$$

where the tilde denotes conjugation. Given the correct choice of q, consistency results for the entries of this matrix as both n and T go to infinity can easily be obtained from Forni, Hallin, Lippi and Reichlin (2000). Results on consistency rates can be found in Forni, Hallin, Lippi and Reichlin (2001a).

An estimate of the spectral density matrix of the vector of the idiosyncratic components $\boldsymbol{\xi}_t = \begin{pmatrix} \xi_{1t} & \cdots & \xi_{nt} \end{pmatrix}'$ can be obtained as the difference $\hat{\mathbf{\Sigma}}_\xi(\theta) = \hat{\mathbf{\Sigma}}(\theta) - \hat{\mathbf{\Sigma}}_\chi(\theta)$.

Starting from the estimated spectral-density matrix we can also obtain estimates of the covariance matrices of $\boldsymbol{\chi}_t$ at different leads and lags by using the inverse discrete Fourier transform, i.e.

$$\hat{\mathbf{\Gamma}}_{\chi k} = \frac{2\pi}{72} \sum_{h=-36}^{36} \hat{\mathbf{\Sigma}}_\chi(\theta_h) e^{i\theta_h k}.$$