EUROPEAN CENTRAL BANK

EUROSYSTEM

WORKING PAPER SERIES

NO 894 / MAY 2008

THE ROLE OF COUNTRY-SPECIFIC TRADE AND SURVEY DATA IN FORECASTING EURO AREA MANUFACTURING PRODUCTION

PERSPECTIVE FROM LARGE PANEL FACTOR MODELS

by Matthieu Darracq Pariès
and Laurent Maurin

EUROPEAN CENTRAL BANK

EUROSYSTEM

WORKING PAPER SERIES

NO 894 / MAY 2008

THE ROLE OF COUNTRY-SPECIFIC TRADE AND SURVEY DATA IN FORECASTING EURO AREA MANUFACTURING PRODUCTION

PERSPECTIVE FROM LARGE PANEL FACTOR MODELS[1]

by Matthieu Darracq Pariès

and Laurent Maurin[2]

In 2008 all ECB publications feature a motif taken from the €10 banknote.

1 The authors would like to thank L. Reichlin and D. Giannone for having made available the codes used in their papers. Comments received from G. Korteweg as well as participants to the December 2006 DIW macroeconometric workshop in Berlin, and the July 2007 KOF Workshop in Zurich are acknowledged.

2 Both authors: European Central Bank, Kaiserstrasse 29, 60311 Frankfurt am Main, Germany;
e-mail: matthieu.darracq_paries@ecb.europa.eu; Laurent.Maurin@ecb.europa.eu
Corresponding author: Laurent Maurin

Address
Kaiserstrasse 29
60311 Frankfurt am Main, Germany

Postal address
Postfach 16 03 19
60066 Frankfurt am Main, Germany

Telephone
+49 69 1344 0

Website
http://www.ecb.europa.eu

Fax
+49 69 1344 6000

ISSN 1561-0810 (print)
ISSN 1725-2806 (online)

CONTENTS

Abstract

Several factor-based models are estimated to investigate the role of country-specific trade and survey data in forecasting euro area manufacturing production. Following Boivin and Ng (2006), the emphasis is put on the role of dataset selection on the empirical performance of factor models. First, spectral analysis is used to assess the information content for euro area manufacturing production of external trade and surveys data of the three largest economies as well as two medium-sized highly opened economies. Second, common factors are estimated on four datasets, following two methodologies, Stock and Watson (2002a, 2002b) and Forni et al. (2005). Third, a rolling out of sample forecast comparison exercise is carried out on nine models. Compared to univariate benchmarks, our results are supportive of factor-based models up to two quarters. They show that incorporating survey and external trade information improves the forecast of manufacturing production. They also confirm the findings of Marcellino, Stock and Watson (2003) that, using country information, it is possible to improve forecasts for the euro area. Interesting, the medium-sized highly opened economies provide valuable information to monitor area wide developments, beyond their weight in the aggregate. Conversely, the large countries do not add much to the monitoring of the aggregate, when considered separately.

Keywords: Factor models, Dataset, Forecasting.

JEL classification: E37, C3, C53.

Non-Technical Summary

Data availability for the euro area makes it possible to conduct conjunctural analysis through the monitoring and modeling of aggregate statistics constructed for the region as a whole, an approach called direct. Alternatively, the indirect or bottom-up approach is based on statistical sources and econometric tools specific to each country, relying on *ex post* aggregation to portray area wide economic fluctuations. While a branch of the literature has started to analyze the relative performance of the two approaches in a forecasting context, so far, no clear conclusion has been reached. This paper intends to bring a new insight to the problem, in a panel data context.

Our main objective is to investigate whether some country-specific dimensions can improve upon the analysis provided using area wide data only. As there is a good rationale for assuming that some country statistics should be given more weight in euro area conjunctural analysis than their mechanical contribution in the euro area aggregate, we do not impose aggregating relationships.

Our study contributes to the debate on the appropriate design of data panels for factor models. This question has regained interest recently, as it has been shown that a large dataset is not always better for estimating the factors. However, we adopt here a relatively modest approach and do not construct the optimal dataset using only statistical methods. Comparing the forecast performance obtained using various datasets, we investigate the extent to which the country dimension can improve upon forecasts based on area wide area statistics. More precisely, we restrain our description of euro area business cycle to manufacturing production and we focus on the information brought by external trade as well as detailed business surveys.

1 Introduction

Data availability for the euro area makes it possible to conduct conjunctural analysis through the monitoring and modeling of aggregate statistics constructed for the region as a whole, an approach called direct. Alternatively, the indirect or bottom-up approach is based on country-specific statistical sources and country-specific econometric tools. It relies on ex post aggregation to portray area wide economic fluctuations. Comparing an extensive set of models, univariate and vector autoregressions, single equation models and factor-based methods, Marcellino, Stock and Watson (2003) suggest that, concerning the euro area, forecasts constructed by aggregating country-specific models are more accurate than forecasts based on aggregated data only. However, so far, the literature analyzing the relative performance of the two approaches in a forecasting context has remained mostly inconclusive. In this paper, we intend to bring a new insight to the problem, in a panel data context and remaining within an econometric framework which mostly regards the euro area through aggregate statistics. Our main objective is to investigate whether some specific country dimensions can improve upon the analysis provided using aggregate data only. We restrain our description of euro area business cycle to manufacturing production which is defined as industrial production excluding construction, energy, mining and quarrying. Our main criteria to evaluate the role of country-specific data is limited to out-of-sample forecasting performance.

Factor models have emerged as an interesting alternative for short-term forecasting of real activity. Indeed, large-panel factor models provide the most appropriate framework to deal with the issue at stake, as shown by their extensive use in the recent literature on macroeconomic forecasting. However, the extent to which factor model methods require the use of a large dataset remains unclear. Boivin and Ng (2006) show that a large dataset is not always better for estimating the factors. When a block of series contains a strongly correlated idiosyncratic component, adding it the dataset reduces the efficiency of the factor estimates. Drawing heavily on the factor model literature, our study also contributes to the debate on the appropriate design of data panels for factor models, relying on forecast performance criteria.

Nonetheless, we do not rely on a statistical procedure to expand the dataset but, taking a more modest approach, we explore two sets of economic indicators for 5 euro area countries which, on the basis of judgment, could be expected to improve the conjunctural analysis for the euro area. The countries considered are the three largest euro area economies as well as two medium-sized highly opened economies, the Netherlands and Belgium. The indicators refer to external trade in goods and business survey data. A benchmark dataset including series followed to monitor euro area activity is extended with such information, and the forecasts obtained from the four resulting datasets are compared. There is indeed a good rationale for assuming that some country statistics should be given more weight in euro area conjunctural analysis than their mechanical contribution in the euro area aggregate. We consider first external trade statistics which are monthly indicators available at a fairly detailed level of decomposition, by country and by type of goods. It might be expected that exports of intermediate or capital goods from one country lead manufacturing production in another one. The relationships may depend on the country dimension, reflecting the relative specialization of each economy, and may also differ for intra-euro area trade and extra-euro area trade. Depending on the type of goods, they may be shifted in time: the production of intermediate goods may lead that of consumer goods, a delay which may be even longer for capital goods. These considerations also suggest that country and sector specific business surveys can present appealing properties to forecast euro area manufacturing production. Overall, the alternative datasets compared include external trade (intra and extra euro area trade) and business surveys information along the country and type of goods dimensions.

In order to assess the information content of each block of time series, a pseudo-real time forecast comparison exercise is performed. Banbura and Runstler (2007) show that a proper accounting of publication lags reinforces the role of some types of data in the forecasting exercise. In addition, the authors propose a methodology to derive forecast weights and forecast precision measures associated with data groupings in the panel. While interesting, we do not implement their method as our dataset is less affected by publication lags and does not mix frequencies. Moreover, while remaining in the field of the methods based on large-panel factors, we want to use different techniques to check for the robustness of the results obtained on alternative datasets. Indeed, available studies on the relative performance of the various methods carried out on US data, like Stock and Watson (2004), Boivin and Ng (2005) or D'Agostino and Giannone (2006) have not reached a clear consensus. Therefore, we consider both the static principal components method of Stock and Watson (2002a and 2002b) (SW), and the two-step approach based on dynamic principal components of Forni, Hallin, Lippi, and Reichlin (2005) (FHLR).

The paper consists of five sections. In the second section, we provide basic evidence based on frequency analysis to support our ex ante judgement that some country-specific data could have a stronger co-movement with euro area manufacturing production than suggested by their economic size. In the third section, both a generalized dynamic factor model and an approximate static factor model are estimated on four datasets. In the fourth section, the estimated factors are used to forecast euro area manufacturing production. A forecast comparison exercise is performed out-of-sample to rank the methods and the datasets. Finally, the fifth section concludes.

2 The information content of external trade and surveys

An analysis is carried out to support the setup of the datasets used in the forecast comparison exercise. The co-movements between the euro area manufacturing production, the external trade and the business surveys of five euro area countries are characterized. The countries consist of the three largest euro area economies, Germany, France, Italy, as well as two medium-sized highly opened economies, the Netherlands and Belgium. The co-movements are characterized using three indicators computed in the frequency domain.

2.1 The indicators used

The instantaneous correlation provides a limited description of the co-movements between the variables of a dataset as the possible time delay in the interaction between variables is not taken into account. Such information is provided by analysis in the frequency domain which allows a decomposition of a stochastic process into a sum of non-correlated waves of different periodicity.[1] Three indicators are used to concentrate information on the co-movement between the time series: dynamic correlation, time-lag and cohesion. These are based on the estimated cross-spectrums between time series.

ρ, the dynamic correlation between two time series corresponds to the ratio between the co-spectrum, r, the real part of the cross spectrum, and the product of the spectral densities, S. It is real and belongs to $[-1, 1]$.[2] It is the correlation coefficient between the real waves of

[1]For a description and use of frequency analysis, see M.B. Priestley (2001).

[2]It corresponds to the real part of coherency. Croux, Forni and Reichlin (2001) show that being obtained by averaging over opposite frequencies, the indicator preserves the information on the de-phase between the time series.

frequency ω:

$$\rho_{xy}(\omega) = \frac{r_{xy}(\omega)}{\sqrt{S_{xx}(\omega)S_{yy}(w)}} \tag{1}$$

The time-lag (TL) between two time series is also computed. Over the frequency band, this indicates the average delay between two time series[3]:

$$TL(\omega) = \frac{\Phi_{xy}(\omega)}{\omega} \qquad \Phi_{xy}(\omega) = \arctan\left(-\frac{q_{xy}(\omega)}{r_{xy}(\omega)}\right) \tag{2}$$

Finally, to concentrate the information on the dynamic correlation across the indicators by type of goods, the cross cohesion (coh) is estimated. Weighting the dynamic correlation between all possible pairs of series in two vectors, the cross-cohesion indicates the amount of co-movement between two sets of information, at each frequency.

$$coh_{xy}(\omega) = \frac{\sum_i \sum_j \alpha_i \alpha_j \rho_{x_i y_j}(\omega)}{\sum_i \sum_j \alpha_i \alpha_j} \tag{3}$$

The three statistics estimated depend on the cross-spectrum. In order to estimate it consistently, the frequency band is divided into 18 intervals and the correlation function is smoothed with a Bartlett window of size 15, close to the square root of the number of observations. Monthly series seasonally adjusted and covering the period from January 1989 to August 2007 are used (224 observations). In all the cases, the series are de-meaned and standardized. When necessary, they are made stationary by taking the monthly rate of growth.

2.2 Euro area manufacturing production and the country components

In the first step, we investigate the co-movements between euro area manufacturing production and the six main components of industrial production at the country level: production of energy, construction, production of intermediate goods, of capital, of durable consumer goods and of non durable consumer - the four later adding up to manufacturing production.

The official weights in euro area statistics are used to take out the country from the euro area aggregate and compute the manufacturing production for the rest of the euro area. Time-shift and dynamic correlation are aggregated over three intervals. From 0 to 3, the low frequency band (Lf) corresponds to cycles with a period of more than one year. It includes the business cycles which last between two and eight years according to the literature. From 4 to 12, the

[3] At frequency ω, $q_{xy}(\omega)$ is the imaginary part of the cross spectrum, named quadrature, and $\Phi_{xy}(\omega)$ is the phase angle shift between the two series.

Tab. 1: Correlation with euro area manufacturing production

		Int.	Cap.	Dur.	NDur.	En.	Const.
Germany	Cont.	63.5	44.2	49.4	28.7	-13.1	30.9
	Lf	82.7	74.7	55.4	35.8	-20.7	10.9
	Mf	64.8	52.8	45.7	24.1	-17.4	34.2
	Hf	59.1	32.8	50.0	31.7	-8.3	32.8
France	Cont.	62.3	49.5	36.8	29.0	-8.5	23.8
	Lf	84.6	57.2	64.0	40.6	-17.2	36.6
	Mf	53.4	44.0	16.8	20.0	-17.3	25.0
	Hf	65.6	52.2	47.2	35.6	-2.9	24.8
Italy	Cont.	55.6	48.4	52.2	49.6	0.4	23.8
	Lf	82.8	57.9	56.3	52.4	-6.5	36.6
	Mf	56.7	49.6	39.5	54.2	-17.3	25.0
	Hf	49.6	48.6	59.4	46.1	16.1	24.8
Holland	Cont.	30.1	22.4	29.4	10.8	-6.3	15.6
	Lf	61.1	39.2	56.5	31.2	-7.9	16.5
	Mf	18.4	9.1	20.0	1.9	-13.2	20.8
	Hf	34.6	27.3	32.2	14.7	-1.7	13.9
Belgium	Cont.	24.9	4.1	31.2	36.1	-4.9	47.8
	Lf	51.7	15.5	47.5	25.3	-7.2	36.2
	Mf	24.3	5.0	32.0	26.1	-20.5	45.6
	Hf	23.3	0.8	29.9	44.9	3.9	50.3

Note: (%). Cont. stands for contemporaneous, Int. for intermediate goods, Cap for capital goods, Dur. for consumer durable goods, NDur. for consumer non durable goods, En. for energy and Const. for construction.

Tab. 2: TIME-SHIFT WITH EURO AREA MANUFACTURING PRODUCTION

		Int.	**Cap.**	**Dur.**	**NDur.**	**En.**	**Const.**
Germany	Cont.	0.1	-0.3	0.1	0.6	-3.7	0.8
	Lf	0.6	-1.2	-0.6	2.3	-12.8	4.8
	Mf	0.0	0.0	0.2	0.5	-3.2	0.5
	Hf	0.0	-0.4	0.0	0.2	-1.1	0.0
France	Cont.	0.0	-0.2	1.7	0.1	-3.5	0.0
	Lf	0.4	-1.2	-0.4	-0.6	-11.0	-0.7
	Mf	-0.1	-0.2	1.7	0.2	-3.2	0.4
	Hf	0.0	0.0	2.4	0.2	-1.3	-0.3
Italy	Cont.	0.0	0.0	-0.1	0.1	3.2	0.0
	Lf	-0.2	0.4	0.8	0.3	8.1	-0.7
	Mf	-0.1	0.0	-0.3	0.1	2.8	0.4
	Hf	0.0	-0.1	0.0	0.0	2.2	-0.3
Holland	Cont.	0.1	0.1	0.2	0.2	-3.5	0.5
	Lf	1.3	-2.0	0.0	1.9	-8.7	2.4
	Mf	0.0	0.6	0.3	-0.3	-2.8	0.6
	Hf	-0.1	0.1	0.0	0.5	-2.8	-0.4
Belgium	Cont.	-0.1	0.8	-0.2	0.0	-3.7	0.5
	Lf	1.5	-3.9	-0.4	-1.3	-9.0	2.0
	Mf	-0.4	0.3	-0.1	0.2	-3.5	0.6
	Hf	-0.2	3.0	-0.3	0.1	-2.4	-0.1

Note: (%). Cont. stands for contemporaneous, Int. for intermediate goods, Cap for capital goods, Dur. for consumer durable goods, NDur. for consumer non durable goods, En. for energy and Const. for construction.

medium frequency band (Mf) includes all the phenomenons which have a period of three to twelve months (one quarter to one year). Finally, from 13 to 18, the high frequency band includes phenomenons which have a periodicity of less than one quarter and therefore do not appear in national accounts. The co-movements between the components of manufacturing production in the five euro area countries and the euro area manufacturing production are analyzed in Tables 1 and 2.

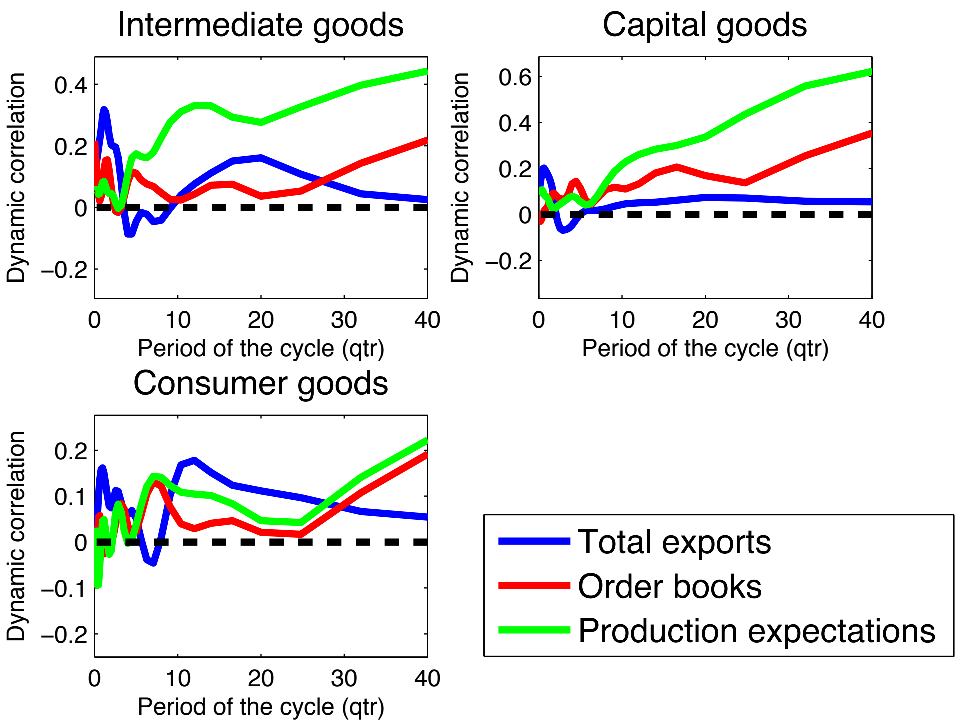
Table 1 shows that the instantaneous correlation provides only a part of the information about the correlation embodied in the sample, as the dynamic correlation over frequency bands can vary substantially.[4] Indeed, in almost of the cases, the correlations at low and medium frequencies is above the static correlation, while the correlation at high frequency is below. To some extent, there is a relationship between the size of the economy and the intensity of its spillovers to the rest of the euro area. At low frequency, Germany, France and Italy are more correlated with the rest of the euro area than the Netherlands. However, the correlation between the size of the economy and the intensity of the spillovers is limited. First, the ranking is not satisfied across the three largest euro area economies. Except for the production of capital goods, where Germany co-moves more intensively with the euro area, no clear distinction can be made as no country co-moves more strongly with the rest of the euro area. This is not explained by composition effects since the result holds at the detailed level, for intermediate goods, consumer goods, and to a lesser extent for energy. Second, at low and medium frequencies, the Belgium production of consumer goods (durable and non-durable), co-moves with the manufacturing production of the rest of the euro area more strongly than that of the Netherlands. For the countries considered, construction co-moves with the euro area manufacturing production , with a positive dynamic correlation lying in a range of 11% to 37% at business cycle frequencies and 16% to 48% contemporaneously. However, this component of manufacturing production has the weakest correlation with manufacturing production apart from energy. On the opposite, production of intermediate goods is the most correlated with manufacturing production, from 83% to 84% for the BIG3 economies, and 52% to 66% for the SMALL2 economies, much above the instantaneous correlation. For all the countries considered except Italy, at business cycle frequencies, this component leads the euro area manufacturing production (see Table 2). The estimated lead is above one month in the case of Belgium and the Netherlands.

[4]The static correlation cannot be easily recomputed from the indications contained in the table. While dynamic correlation aggregated over the whole frequency band is equal to the static correlation, dynamic correlation within a frequency band is not a simple average of the values taken within the band.

2.3 Country-specific trade and surveys data and euro area manufacturing production

In the second step, the co-movements between the euro are manufacturing production and the country specific trade and business survey data are analyzed to explain why the series considered are included in the dataset. For external trade, exports values from EUROSTAT external trade statistics are considered.[5] Half of the series refers to intra-euro area trade in goods and half refers to extra-euro area trade in goods. For surveys, data from the European Commission business surveys are used. These include surveys on export order books, on order books, and on production expectations. For both external trade and business survey data, the decomposition into the main economic categories is considered: capital goods, intermediate goods and consumer goods.

Fig. 1: Cross cohesion with euro area manufacturing production



The cross cohesion is computed to aggregate the country dimension and analyze the co-movements between euro area manufacturing production and external trade on the one hand

[5]Values are preferred to volumes as they are more reliable, available over a longer time span and more timely. Exports are preferred to imports as they are less sensitive to fluctuations in the oil bill which is likely to reflect a separate factor.

(intra and extra exports), and production expectations on the other hand. The results are shown in Chart 1 for each of the broad economic categories separately. The Chart confirms that the relationship depends on the frequency band considered and on the type of goods. At business cycle frequency,[6] the co-movements are more intense for intermediate goods, for which the averaged dynamic correlation of production expectations with the euro area manufacturing production lies above 30% at business cycle frequency. More generally, for all type of goods, extra exports and intra exports are correlated with manufacturing production. However, the strongest co-movements are observed with surveys and among all the indicators, production expectations having the strongest correlation with manufacturing production. Among surveys, order books assessment displays a much weaker correlation.

Overall, Tables 1, 2 and Chart 1 confirms that surveys convey information on movements in manufacturing activity over the business cycle. They show that exports are positively correlated with manufacturing production. Interestingly, the cycles in intermediate goods are the most correlated with the manufacturing cycles. These are leading the euro area manufacturing cycle in the case of Belgium and Netherlands. In terms of country specific dimensions, a more systematic analysis could be done. Indeed, the dataset choices have been mainly limited by data availability considerations.[7] In the next section, factor analysis is carried out to concentrate the amount of co-movement in the dataset and provide forecasts of euro area manufacturing production.

3 Estimating the Factor space

Most of the notation in this section are borrowed from D'Agostino and Giannone (2006). Consider a dataset that consists of n demeaned and standardized time series, $X_t = (x_{1t}, ..., x_{nt})'$, each of them representing a time series ($t = 1, ..., T$). X_t is assumed to follow an approximate factor structure and each serie is decomposed into a common component, χ, and an idiosyncratic component, ξ:

$$X_t = \chi_t + \xi_t \tag{4}$$

The idiosyncratic component is allowed to be weakly cross-correlated in the sense of Bai and Ng (2002) and weakly serially correlated while the common component is generated by q

[6]Consistently with the definition above, this corresponds to the part of the graph between 8 and 32 quarters on the x-axis.

[7]The five countries considered amounts to 80% of euro area industrial production excluding construction.

common dynamic factors:

$$X_t = \sum_{j=0}^{s} \Lambda_j L^j f_t + \xi_t \tag{5}$$

where f is a $q \times 1$ vector of common dynamic factors.

Equation 5 can also be written in a static form:

$$X_t = \underline{\Lambda} F_t + \xi_t \quad \text{with} \quad F_t = \left(f_t' ... f_{t-s}' \right) \tag{6}$$

where F is a $r \times 1$ a vector of static factors (with $r = q(s+1)$). The r static factors have the following dynamics:

$$F_t = \mathcal{A}(L) F_{t-1} + u_t$$

where $\mathcal{A}(L)$ is a matrix polynomial lag operator with coefficient matrices of size $r \times r$. We assume that the roots of $|I_r - \mathcal{A}(z) z|$ lie outside the unit circle.

Let us define the order-k covariance matrix of X_t by Γ_k. Given the independence of the common and idiosyncratic components, it can be decomposed into

$$\Gamma_k = \Gamma_k^{\chi} + \Gamma_k^{\xi} \tag{7}$$

where $\Gamma_k^{\chi} = \underline{\Lambda} \Gamma_k^F \underline{\Lambda}'$ is of rank r, Γ_k^F is the covariance matrix of F_t at lag k and Γ_k^{ξ} is the covariance matrix of ξ_t at lag k. A consistent estimator $\widehat{\Gamma}_k$ of the covariance matrix given the dataset $X_t = (x_{1t}, ..., x_{nt})'$ for $t = 1, ..., T$ is

$$\widehat{\Gamma}_k = \frac{1}{T-k+1} \sum_{t=k}^{T} X_t X_{t-k}' \tag{8}$$

3.1 Two methods

In the method proposed by Stock and Watson (2002a) (SW), the authors assume a static form representation of the data as in Equation (6) and use sample principal components to extract the factors $\widehat{F}_t$. The r principal components are given by

$$\widehat{F}_t = V_r' X_t \tag{9}$$

where V_r is the $n \times r$ matrix of the r eigenvectors associted with the r largest eigenvalues of the sample covariance matrix $\widehat{\Gamma}_0$.

The common component is then derived from

$$\widehat{\chi}_t = \widehat{\Gamma}_0 V_r \left(V_r' \widehat{\Gamma}_0 V_r \right)^{-1} V_r' X_t \tag{10}$$

Differently, in the methods proposed by Forni et al (2000 and 2005) (FHLR), the authors make use of the dynamic factor representation of the observables as in Equation (5). An estimation of the common and idiosyncratic components of the dynamic-factor representation can be provided by a principal component analysis of the spectral density matrix of the dataset.

Let us denote $\Sigma(\theta)$ the spectral density matrix of X_t at frequency $\theta \in [-\pi, \pi]$. We have

$$\Sigma(\theta) = \Sigma_\chi(\theta) + \Sigma_\xi(\theta) \tag{11}$$

where $\Sigma_\chi(\theta)$ is the rank q spectral density matrix of the common component and $\Sigma_\xi(\theta)$ is the spectral density matrix of the idiosyncratic component.

The spectral density matrix of X_t can be estimated by

$$\widehat{\Sigma}(\theta) = \frac{1}{2\pi}\sum_{k=-m}^{m} \alpha_k \widehat{\Gamma}_k e^{-i\theta k} \tag{12}$$

where α_k are weights satisfying the conditions: $\alpha_0 = 1$ and $0 \leq \alpha_k \leq 1, \forall k \leq m$.

A consistent estimate of the spectral density matrix of the common and idiosyncratic components are given by

$$\widehat{\Sigma}_\chi(\theta) = \mathcal{V}_q(\theta)\mathcal{D}_q(\theta)\mathcal{V}_q(\theta)' \tag{13}$$

$$\widehat{\Sigma}_\xi(\theta) = \widehat{\Sigma}(\theta) - \widehat{\Sigma}_\chi(\theta) \tag{14}$$

where $\mathcal{D}_q(\theta)$ is the diagonal matrix having the first q largest eigenvalues of $\widehat{\Sigma}(\theta)$ on the diagonal, and $\mathcal{V}_q(\theta)$ is the $n \times q$ matrix of the corresponding eigenvectors. Using the inverse-Fourrier transform, the covariance matrices are then computed as

$$\widehat{\Gamma}_k^\chi = \frac{2\pi}{2m+1}\sum_{j=-m}^{m} \widehat{\Sigma}_\chi(\theta_j) e^{i\theta_j k} \tag{15}$$

$$\widehat{\Gamma}_k^\xi = \frac{2\pi}{2m+1}\sum_{j=-m}^{m} \widehat{\Sigma}_\xi(\theta_j) e^{i\theta_j k} \tag{16}$$

with $\theta_j = \frac{2\pi}{2m+1}j$ and $j = -m, ..., m$.

At this stage, the factor space and the common component is consistently estimated through this spectral domain principal component analysis like in the Forni et al. (2000) methodology. However, the procedure to estimate the dynamic factors leads to a two-sided filter which is not well suited for forecasting purposes. In order to solve this issue, Forni et al (2005) use the estimates of the covariance matrices of the common and idiosyncratic components to formulate a generalized principal component problem:

$$\widehat{\Gamma}_0^{\chi} V_{rg} = \widehat{\Gamma}_0^{\xi} V_{rg} D_{rg} \qquad (17)$$
$$s.t.\ V'_{rg} \widehat{\Gamma}_0^{\xi} V_{rg} = I_r$$

where D_{rg} is the diagonal matrix having the first r largest generalized eigenvalues of the pair $\left(\widehat{\Gamma}_0^{\chi}, \widehat{\Gamma}_0^{\xi}\right)$. V_{rg} is the $n \times r$ matrix of the corresponding eigenvectors. The methodology puts a higher weight on the variables which have a higher degree of commonality. Moreover, the method shifts each variable in time on the basis of the cross-correlation at all leads and lags.

Then, the factor space and the common components are obtained as

$$\widehat{F}_t^g = V'_{rg} \widehat{X}_t \qquad (18)$$
$$\widehat{\chi}_t^g = \widehat{\Gamma}_0^{\chi} V_{rg} \left(V'_{rg} \widehat{\Gamma}_0 V_{rg}\right)^{-1} V'_{rg} X_t \qquad (19)$$

Indeed, in the case of Stock and Watson (2002b), the common component of X_t is computed using the contemporaneous covariance matrix, while in the case of Forni et al (2005), it is estimated using the contemporaneous covariance matrix of the common components, estimated in a first step.

3.2 Description of the various datasets

The methods described above are implemented on four different datasets which correspond to a set of aggregate euro area statistics augmented with country specific business surveys and external trade series.

The first dataset, denoted BENCHMARK, comprises standard variables used to forecast euro area activity and inflation.[8] It consists of 91 series distributed in three blocks. The first one contains 73 series concerning the euro area, the second block contains 13 series for the United States and the third one contains 5 series related to the world markets. The euro area block includes loans indicators and monetary aggregates, real activity data, production price series, survey data and financial data. Real activity data contain components of manufacturing production (namely capital goods, intermediate goods, non-durable and durable consumer goods), retail sales, labor market data, and external trade. Survey data include series from the European Commission, namely business survey, consumer survey, retail and construction

[8]The dataset is comparable to the one used by Giannone, Reichlin and Sala (2006).

surveys, and passenger car registration. Financial data include exchange rates, interest rates, and equity price indices. The US block is a sub sample of the euro area block. The world markets block includes oil prices, raw material prices and gold price.

The three other datasets expand the benchmark and merge it with country-specific data on the main industrial groupings for extra and intra euro area exports of goods, production expectations, assessment on order books and export order books.[9] The second dataset, denoted BIG3, adds data for the three largest euro area economies (Germany, France and Italy) and contains 136 series. The third one, denoted SMALL2, merges BENCHMARK and the two medium-sized highly opened economies considered in the sample, the Netherlands and Belgium. It contains 121 series Finally, the fourth dataset, denoted ALL, complements the euro area dataset with the same series for the five countries at the same time: Germany, France, Italy, Spain, the Netherlands and Belgium. This dataset contains 166 series.

Tab. 3: BAI AND NG (2002) CRITERIAS

	BENCHMARK			ALL			BIG3			SMALL2		
nstat	ICp1	ICp2	ICp3	ICp1	ICp2	ICp3	ICp1	ICp2	ICp3	ICp1	ICp2	ICp3
1	0.83	0.84	0.81	0.83	0.84	0.82	0.83	0.83	0.81	0.83	0.83	0.81
2	0.80	0.81	0.77	0.80	0.81	0.76	0.80	0.81	0.77	0.80	0.81	0.77
3	**0.79**	**0.80**	0.74	**0.79**	**0.81**	0.74	**0.78**	**0.80**	0.74	**0.79**	**0.80**	0.74
4	0.80	0.82	**0.74**	0.80	0.82	0.73	0.79	0.81	**0.73**	0.79	0.81	**0.73**
5	0.82	0.84	0.74	0.81	0.84	0.73	0.81	0.84	0.73	0.81	0.84	0.74
6	0.84	0.87	0.75	0.83	0.86	**0.72**	0.83	0.87	0.74	0.84	0.87	0.75
7	0.87	0.91	0.77	0.85	0.89	0.73	0.86	0.90	0.76	0.87	0.91	0.76
8	0.91	0.95	0.79	0.87	0.92	0.73	0.90	0.94	0.78	0.90	0.95	0.78
9	0.95	1.00	0.81	0.90	0.95	0.74	0.94	0.99	0.80	0.94	0.99	0.81
10	0.99	1.04	0.84	0.92	0.98	0.76	0.98	1.03	0.83	0.98	1.04	0.83

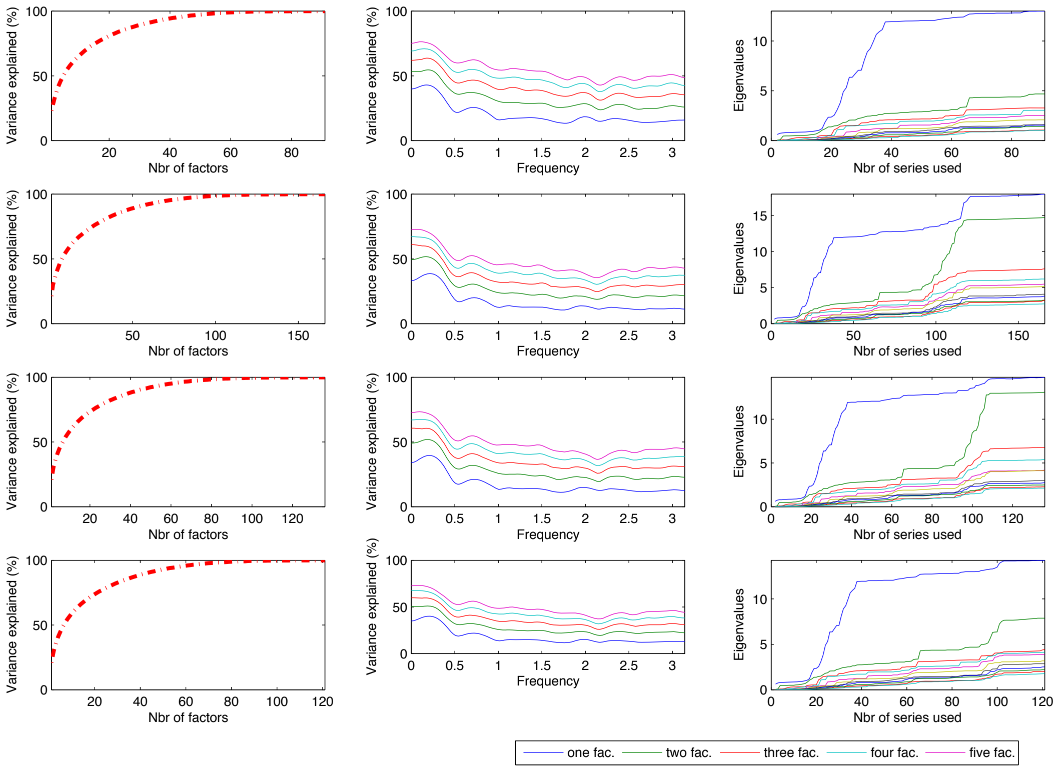
3.3 Extracting the common component of euro area manufacturing production

In the case of static factors like SW, the estimation of the common component requires only to fix r, the number of factors, while in the case of dynamic factors, it requires to fix the number of dynamic factors q, the frequency grid over which the spectrum is estimated and the window used to smooth the cross spectrum.

Bai and Ng (2002) propose to minimize an information criteria to determine the number of

[9] For each item, three series are considered: capital goods, intermediate goods and consumer goods.

Fig. 2: NUMBER OF FACTORS IN THE DATASETS



Note: Results obtained with the BENCHMARK (1st line), ALL (2nd line), BIG3 (3rd line) and SMALL2 (4th line).

Tab. 4: MANUFACTURING PRODUCTION: SHARE OF THE VARIANCE EXPLAINED (%)

	Benchmark	BIG3	SMALL2
Static factors			
1	43.1	35.5	41
2	45.4	39.1	41.2
3	45.8	39.1	41.3
4	62.7	40.3	41.3
5	74.1	66.9	75.4
6	75.6	68.8	75.5
7	76.3	73.7	75.8
8	76.7	74.5	75.9
9	77.1	74.5	75.9
10	77.1	74.7	76.1
Dynamic one-sided			
1	47.1	35.4	41.5
2	54.6	38.8	41.6
3	54.5	46.4	46.7
4	68.1	58.3	63.8
5	68	60.2	63.7
6	67.9	60.2	63.8
7	67.9	60.2	64.5
8	68.1	62.2	64.3
9	68.1	61.8	64.6
10	68	61.8	64.7
Dynamic two-sided			
1	49.1	42.9	46.5
2	59.9	55.4	56.8
3	69.2	67	67.7
4	71.7	70.3	71.8
5	74.8	73.3	75.1

static factors:

$$r = Arg \min_r \left(\ln(\widehat{V}_r) + r.g(N,T) \right) \qquad where \qquad \widehat{V}_r = \frac{1}{NT} \sum_{t,i} \widehat{\xi}^2_{t,i}$$

Where $\widehat{V}_r$ is the average residual variance obtained when r factors are estimated and g is a penalty function. Bai and Ng (2002) propose three penalty functions which define three information criteria. These criteria have different properties in finite sample but are asymptotically equivalent.

$$\begin{aligned} g_1(N,T) &= \left(\frac{N+T}{NT}\right) \ln \left(\frac{NT}{N+T}\right) \\ g_2(N,T) &= \left(\frac{N+T}{NT}\right) \ln \left(\min(\sqrt{T}, \sqrt{N})\right) \\ g_3(N,T) &= \frac{\ln\left(\min(T,N)\right)}{\min(T,N)} \end{aligned}$$

These criteria are computed over the four datasets for 1 to 10 factors (see Table 3). The results show that the number of factor to retain varies in a narrow range across datasets, from three to four in most of the cases. While the two first criteria always indicate three factors, the third criteria always indicate more factors, four in all but one cases. Chart 2 also presents the share of variance explained by the first eigenvalues, contemporaneously but also in the frequency domain (respectively in the left column and in the center column). Again, independently of the dataset, the chart shows the strong degree of co-movement among the series. 20 factors explain around two-third of the total variance of BENCHMARK, SMALL2 and BIG3 while 10 factors explain around half. In the case of ALL, the share of the variance explained by the 20 first factors, one-seventh of the numbers of series included in the dataset, remains above one-half.

More heuristic criteria can be used to determinate the number of common factors q, the dynamic rank of the variance-covariance matrix, and r, the static rank (see Forni et al (2000)). First, there should be a substantial gap between the variance explained by the q first principal components and the variance explained by the other. If data are generated by a dynamic factor model with q factors, then when incrementing the dataset, the average over frequency of the first q empirical eigenvalues should diverge, whereas the average of the other ones should remain relatively stable. This criteria is used to analyze the number of factors driving the four datasets considered. Starting with a small number of time series relatively to the number included in the final dataset, the dataset is incremented by adding the series one by one up to the final panel. At each step, the spectral density matrix is estimated, the dynamic eigenvalues are computed and averaged over the grid of frequencies. The results are displayed in the right hand side of Chart 2. The horizontal axis indicates the size of the sample N, which ranges from 2 to 91 for the BENCHMARK dataset and 2 to 166 for the largest dataset (ALL). The vertical axis plots the average over frequencies of the 10 first estimated eigenvalues. It appears that the first two eigenvalues exhibit a relatively constant positive slope, while the remaining ones appear rather flat. In the frequency domain, the charts show the percentage of the variance explained by 1 to 5 dynamic factors. We see that the gap between the share of the variance explained by the first three eigenvalues and the first four eigenvalues is larger than the gap between the first four and the first five.

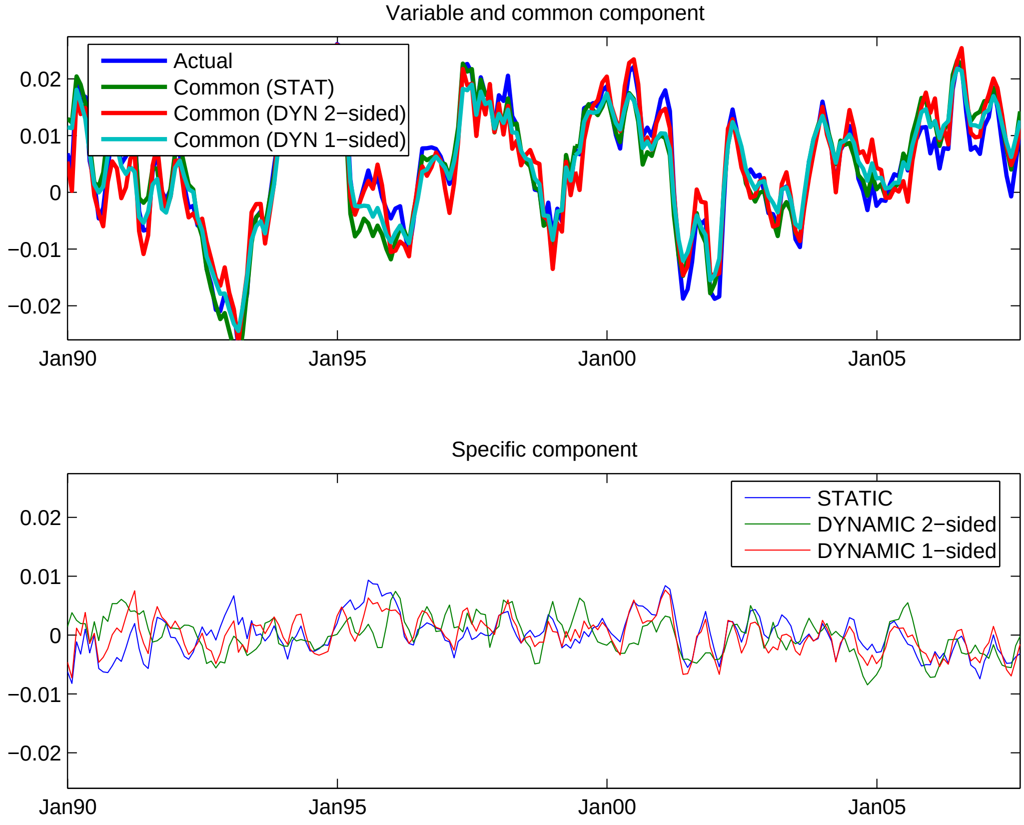
Finally, the gap between the number of static and dynamic factors also suggests that the time dimension is important (since $r = q(s + 1)$, $q = 2$ and $r = 4$ implies $s = 2$).

Table 4 also reports the share of the variance explained depending on the number of factors and the method used, but focusing on euro area manufacturing production only. Notice that the number of factors remains surrounded by some uncertainty as Chart 2 and Table 3 show that we could retain 3 to 4 static factors. Forni et al. (2000) and Boivin and Ng (2006) show that, while underestimating the number of factors results in large efficiency losses, overestimating it has little implication. Therefore, we retain 2 dynamic factors and 4 static factors. These numbers lie in the range suggested in the literature on real activity.[10]

Based on these parameters, Chart 3 presents the common and specific components estimated with the three methods, the one based on Stock and Watson (2002b) methodology, called STATIC, the one based on Forni et al. (2005), called DYN 1-sided, and the one based on Forni et al. (2002) methodology, called DYN 2-sided. The differences obtained with the two first methods, based on the same number of factors, reflect the differences in the estimated factors that solve a generalized eigenvalue problem in which the series are weighted according to their common to idiosyncratic ratio in Forni et al. (2005).

[10]For instance, d'Agostino and Giannone (2006) retain 10 factors to explain a dataset of around 150 series for the US economy. Giannone, Reichlin and Sala (2005) retain 2 dynamic factors to forecast the federal fund rates using data from the Beige book. Boivin and Ng (2005) also retain 2 dynamic factors to concentrate the information contained in a panel of 147 monthly data.

Fig. 3: ESTIMATED COMPONENTS OF EURO AREA MANUFACTURING PRODUCTION



Note: ALL, BENCHMARK dataset augmented by information on the 5 countries considered. Common component estimated with 2 dynamic factors and 4 static factors. Three-month moving averages reported.

4 Forecast comparison exercise

4.1 Specification of the diffusion forecasts

In this section, we use the various factor models to specify the best h-step ahead linear prediction of euro area manufacturing production, x_{IP} which corresponds to the line IP of the vector of standardized data X_t:

$$x_{IP,t+h|t} = \mathbb{E}^{linear}\left[x_{IP,t+h} \mid \Omega_t\right]$$

with $\Omega_t = \text{span}\left\{X_{t-i}, i \in [0, +\infty)\right\}$.

Given the orthogonality assumed between common and idiosyncratic components, the forecast equation can be written as

$$\begin{aligned} x_{IP,t+h|t} \quad &\approx \quad \mathbb{E}^{linear}\left[\chi_{IP,t+h} \mid \text{span}\left\{F_{t-i}, i \in [0, +\infty)\right\}\right] \\ &\quad\quad +\mathbb{E}^{linear}\left[\xi_{IP,t+h} \mid \text{span}\left\{x_{IP,t-i}, i \in [0, +\infty)\right\}\right] \end{aligned}$$

Thereafter, we follow closely Boivin and Ng (2005) in order to derive a battery of forecast diffusion processes. More detailed explanations on the underlying assumptions on the probabilistic structure and its approximation can be found in this paper. First, we consider unrestricted diffusion forecasts of the form

$$x_{IP,t+h|t} = \widehat{\gamma}_{IP}\left(L\right)F_t + \widehat{\varphi}_{IP}\left(L\right)x_{IP,t} \tag{20}$$

where $\widehat{\gamma}_{IP}\left(L\right)$ and $\widehat{\varphi}_{IP}\left(L\right)$ are obtained by regressing $x_{IP,t+h|t}$ on the pf lags of F_t and pa lags of $x_{IP,t}$. This equation can be estimated using the two factor spaces $\widehat{F}_t$ from the SW approach and $\widehat{F}_t^g$ from the FHLR approach. The corresponding diffusion forecasts will be referred thereafter as *Unrestricted Static* and *Unrestricted Dynamic* , as in Boivin and Ng (2005). These equations are also estimated under the constrain of no autoregressive terms.

Second, we make use of the common component structure to formulate the index forecasts. To begin with, we estimate a restricted form in which only the common component is introduced and the idiosyncratic component is assumed to be $i.i.d.$

$$x_{IP,t+h|t} = \widehat{\underline{\gamma}}_{IP}\left(L\right)\chi_t \tag{21}$$

This equation is applied with SW and FHLR common components and are called *Common Static* and *Common Dynamic* respectively. Those diffusion forecasts can be performed through a sequence of one-step ahead forecasts. However, the results are not reported here.

Instead, we only investigate the nonparametric projection of the FHLR common component as follows:

$$x_{IP,t+h|t} = \widehat{\chi}^{g}_{IP,t+h|t} \tag{22}$$

$$\widehat{\chi}^{g}_{IP,t+h|t} = \widehat{\Gamma}^{\chi}_{IP,h} V_{rg} \left(V'_{rg} \widehat{\Gamma}_0 V_{rg} \right)^{-1} V'_{rg} X_t \tag{23}$$

where $\widehat{\Gamma}^{\chi}_{IP,h}$ is the IPth row of $\widehat{\Gamma}^{\chi}_{h}$ and the idiosyncratic component is assumed $i.i.d.$ This diffusion forecast is referred as *Dynamic Non-Parametric*. The analogous non-parametric projection of the common component based of the SW factor space is not performed.

Finally, since $\widehat{\chi}_{IP,t+h|t} = \underline{\widehat{\Lambda_i F}}_{t+h|t} \neq \underline{\widehat{\Lambda}_i} \widehat{F_{t+h|t}}$ when parameters are not observed, we have another possibility to forecast the comon component. The following index forecast is considered:

$$x_{IP,t+h|t} = \underline{\widehat{\Lambda_i}}' \widehat{\mathcal{A}}(L)\, F_{t+h-1|t} + \widehat{\phi}(L) \xi_{IP,t+h-1|t} \tag{24}$$

The factor space and loadings are either SW or FHLR and the corresponding diffusion forecasts are called *Sequential Static* and *Sequential Dynamic*.

The forecast properties of factor-based forecasts have been compared in a number of studies (Stock and Watson (2002a), Forni et al. (2005), Boivin and Ng (2006), Kapetanios and Marcellino (2003), d'Agostino and Giannone (2006)). Overall, it is widely recognized that factor-based models do improve upon univariate forecasts (ARIMA models) and multivariate ones (VAR). However, it seems particularly difficult to largely beat an AR forecast out-of-sample and the improvement resulting from factor based forecasts is limited, at least in small sample.

The literature is less conclusive on the method which performs best to extract the factors and forecast the idiosyncratic component. Stock and Watson (2002a) include lags of the dependent variable as additional predictors for the idiosyncratic component while Forni et al. (2005) exploit the orthogonality of the components and forecast them separately. Boivin and Ng (2006) conclude that using the factor structure to forecast the idiosyncratic component does not clearly improve the forecast but that optimizing the value of the parameters remains important in the dynamic model: the step used to decompose the frequency, the length of the window as well as the method used to smooth the cross-spectrum. Marcellino, Stock and Watson (2006) show that sequential forecasts (the forecast h-step ahead is computed by rolling on one step-ahead forecasts) outperform direct forecasts.

More recently, the sensitivity of the forecast to the dataset used to extract the factors has also been investigated. Boivin and Ng (2006) show that big is not always better: Adding more time series, for which the specific components are strongly correlated with others, reduce the efficiency of the factor estimates and therefore can be detrimental to the estimation and even more to the forecast. The question of forecasting the aggregate vs forecasting the components has also been renewed with the factor-based approach. Marcellino Stock and Watson (2003) show that the aggregation of country based forecasts provide a better forecast than a direct euro area forecast. However, their conclusion remains based on several independent forecasts.

4.2 Pseudo-out-of-sample forecast comparison

In this sub-section, we analyze the forecast performance of the different datasets across the various factor-based diffusion forecasts presented previously. Monthly seasonally adjusted data are considered over the period January 1989 to August 2007 (224 observations). When observations are missing, the Stock and Watson EM algorithm is used to estimate them. First, missing observations are substituted to the unconditional means. This new dataset is then used to estimate the factors and loadings in the first step. During the $(n+1)$th step, the factors and loadings extracted in the previous iteration are used to generate a new estimation.[11] The exercise is performed starting with a dataset comprising 60 observations, a minimum requirement to estimate the factors and the coefficient in the factor-based diffusion equations.

The prediction accuracy is evaluated at a given horizon using the root mean squared forecast errors ($RMSFE$) metric. For comparison purposes, two univariate models are added to the nine factor-based diffusion forecasts: a random walk with drift on levels, and an autoregressive model on first differences. The forecast comparison exercise is conducted on a rolling out of sample basis. Given a dataset, the estimation of the factors and all the diffusion forecasts are estimated with the information available until month t. Those equations are then used to forecast euro area manufacturing production from $t+1$ to $t+12$ months ahead. The monthly forecasts are transformed into quarterly frequency, 1 to 4 quarters ahead. Afterwards, the sample is incremented by one observation (one month), and all the steps are redone: factor estimations, estimation of the equations, forecasts, and quarterly aggregation. For each estimation, the number of parameters estimated remains constant. The autoregressive terms and the factors in (20) are estimated with two lags. The diffusion forecasts based on the common component (see equation (21)) introduce current and lagged terms. The sequential forecast of

[11] For a detailed description of the method, see Stock and Watson (2002a). For an application to mixed frequency panel, see Schumacher and Breitung (2006). For a more general discussion and a comparison of various methods, see Angelini, Henry and Marcellino (2006).

equation (24) assume that the idiosyncratic component follows an autoregressive process of order 1 and that the factors are described by a VAR(2). In all the diffusion forecasts, the number of static factors are 4 in the SW and FHLR methods, and we assume 2 dynamic factors in the first step of the FHLR approach.

Given the relatively small number of observation used for the first estimations, the equations that require the estimation of fewer parameters may be advantaged, and the exercise may provide more support to the random walk with drift and the autoregressive models. However, the rolling Root Mean Square Forecast Errors (RMSFE) should provide information on this effect since it is expected to diminish over time.

Tables 5 to 8 present the RMSFE of the factor-based forecasts relative to the random walk with drift for the 4 datasets. The forecast comparison exercise provides two sets of results. First of all, it allows to examine the relative merits of the various diffusion forecasts in predicting euro area manufacturing production, given the various datasets used. Second, across the range of diffusion forecasts developed here, we explore the role of the dataset on the forecast performance. This provides some elements to answer the question at the core of this paper, namely whether country-specific trade and survey data improve on the forecast performance derived from aggregate euro area dataset.

Regarding the first set of results, note that, in terms of forecast error, the univariate autoregressive model performs slightly better than a random walk with drift at all horizons, by around 5 to 10%. Turning to factor based forecasts, they generally do improve on the random walk with drift forecast and on the autoregressive forecast in the short run, up to three quarters. However, at three and four quarters, some factor-based forecasts present only a marginal improvements on the univariate forecasts. The worse outcomes tend to be recorded with the *Sequential* and *Common* diffusion forecasts. There is indeed no significant evidence in favor of the sequential forecasts, as found in Marcellino et al. (2006). Note however that the diffusion forecasts we selected do not allow a direct comparison of sequential versus direct forecast since the specification (20) is not the direct forecast equivalent to (24).

Regarding the forecasts based on the common component, the relative weak performance compared with *Unrestricted* forecasts in particular should be weighted against the higher number of coefficient to estimate and it is not sure that information criteria would support the unrestricted factor structure. This supports the possibility of nuisances introduced by the higher number of parameters to be estimated when factors are used instead of the common component.

In terms of RMSFEs, the improvements of *Unrestricted* forecasts compared with univariate models are substantial, especially at a one-quarter horizon. The gains can reach 25 to 30% on average across the different methods and datasets. At longer horizons, the RMSFEs are generally around 20% lower than with random walk with drift forecast. The presence of autoregressive terms in the specification of the *Unrestricted* diffusion forecasts is not strongly improving the forecast performance. It seems to help forecasting at short horizons and with the SW static factors.

Finally, the *Dynamic non-parametric* forecasts performs relatively well at one-quarter and two-quarter horizons. This diffusion forecast delivers the lowest RMSFEs for 3 out of our 4 datasets. At longer horizons however, the performance deteriorates and becomes significantly worse than the *Unrestricted* forecasts.

Regarding the relative performance of models based on static versus dynamic factors, we do not find strong evidence in favor of the static factors, contrastingly with Boivin and Ng (2005). Actually, for the datasets BENCHMARK, BIG3 and SMALL2, the diffusion forecasts including dynamic factors tend to perform slightly better.

We turn now to the analysis of the sensitivity of forecast performance to the different datasets. In order to reach conclusions which are robust to alternative diffusion forecast specifications, we compute, for each dataset and each horizon, the best linear forecast combination of the nine diffusion forecasts, the one which minimizes the weighted sum of the RMSFEs. Table 9 reports the results of this exercise: for each horizon and each dataset, the table presents the smallest RMSFE obtained by linear weighting of the diffusion forecasts.

We note first that augmenting the benchmark dataset with survey and trade flows information on the three largest euro area members (dataset BIG3) does not result in an improved forecast performance. At all horizons, the RMSFEs are actually slightly higher than with the BENCHMARK dataset. On the contrary, the dataset SMALL2 which merges information for the euro area with survey and trade data for the Netherlands and Belgium, delivers significantly better forecasts than the BENCHMARK dataset at one-quarter and two-quarter horizons notably. Those results can be explained by considering that information on the three large countries is already largely contained in the euro area aggregate dataset, while information on smaller countries can be under represented in aggregate statistics, considering they smaller mechanical weight. Finally, augmented with data for the 5 countries considered, the dataset

ALL performs better than the dataset BIG3 and similarly to the dataset SMALL2 for horizons at and beyond two quarters.

To illustrate further the role of the information set on forecast performance, Figures 4 and 5 plot the rolling RMSFEs (averaged over a two-quarter window) at one-quarter and two-quarter horizons, for the best forecast combination based on each of the four datasets. A visual inspection of the series shows that, during periods of sharp cyclical developments (around 1993, 1997, 1999-2001 or 2003), the BIG3 forecast is most of the time worse than the BENCHMARK forecast. However, SMALL2 or ALL forecasts can improve substantially on the BENCHMARK one. This is particularly true for the period 1999-2002.

5 Concluding remarks

The forecast comparison exercise at the core of this paper contributes in several dimensions to the applied econometrics literature using large-panel factor models. First, focusing on the euro area manufacturing production, we showed that the performance of factor-based diffusion forecasts should be examined across the range of available methods since none proves to be dominant. Second, we illustrated the sensitivity of the forecast properties to the information set, providing examples where more information was deteriorating the prediction abilities of our diffusion forecasts, examples that tend to confirm that more data is not always better in terms of forecasting performance. This suggests that a careful attention should be dedicated to the construction of the dataset especially when the economist faces a situation in which the number of observation is small. Finally, we provided evidence that selected country dimensions along trade and survey statistics could significantly improve on the forecast performance derived from aggregate euro area dataset. This appears especially interesting for the two medium-sized highly opened economies considered.

However, we did not provide any systematic method to construct an appropriate dataset. In the exercise, information has been added to the dataset on the basis of economic judgment without constraints and our results confirmed our prior intuition. As far as euro area analysis is concerned, the best way to use information on country and on the type of goods remains to be determined. One could envisage a more systematic selection procedure for the dataset, based on statistical criteria.

Tab. 5: RMSFE AT 1 TO 4 QTR, DATASET BENCHMARK

h	1	2	3	4
Random walk with drift	2.2	2.2	2.2	2.2
Autoregressive	90.0	94.0	95.4	96.7
Unrestricted static no autoreg	74.4	83.0	72.6	77.2
Unrestricted dynamic no autoreg (DD)	70.2	79.3	77.1	75.1
Unrestricted static autoreg	73.7	81.7	72.7	78.7
Unrestricted dynamic autoreg (DDAR)	71.4	79.4	76.7	76.6
Sequential static	87.2	93.7	93.4	94.7
Sequential dynamic (SD)	76.9	88.3	88.5	93.1
Dynamic non-parametric (DNP)	68.6	76.7	83.2	81.2
Common static	88.4	92.0	94.9	95.6
Common dynamic (COMD)	86.2	91.1	93.3	95.0

Tab. 6: RMSFE AT 1 TO 4 QTR, DATASET ALL

h	1	2	3	4
Unrestricted static no autoreg	71.3	82.6	81.0	87.0
Unrestricted dynamic no autoreg (DD)	72.6	82.5	80.0	81.9
Unrestricted static autoreg	70.0	79.5	79.1	86.7
Unrestricted dynamic autoreg (DDAR)	72.8	82.1	77.3	82.4
Sequential static	85.6	95.3	93.4	94.6
Sequential dynamic (SD)	82.8	91.7	89.1	95.5
Dynamic non-parametric (DNP)	73.2	82.4	86.4	88.0
Common static	80.1	87.6	90.0	91.9
Common dynamic (COMD)	83.3	90.3	92.3	93.1

Tab. 7: RMSFE at 1 to 4 qtr, dataset BIG3

h	1	2	3	4
Unrestricted static no autoreg	74.97	83.77	73.23	78.34
Unrestricted dynamic no autoreg (DD)	70.65	80.68	76.65	74.15
Unrestricted static autoreg	74.01	82.31	73.24	79.68
Unrestricted dynamic autoreg (DDAR)	71.96	81.11	76.67	76.15
Sequential static	87.69	93.91	93.52	95.22
Sequential dynamic (SD)	87.33	94.13	93.16	95.10
Dynamic non-parametric (DNP)	68.77	77.49	83.09	81.63
Common static	88.35	92.35	94.77	95.56
Common dynamic (COMD)	86.63	91.60	93.59	95.10

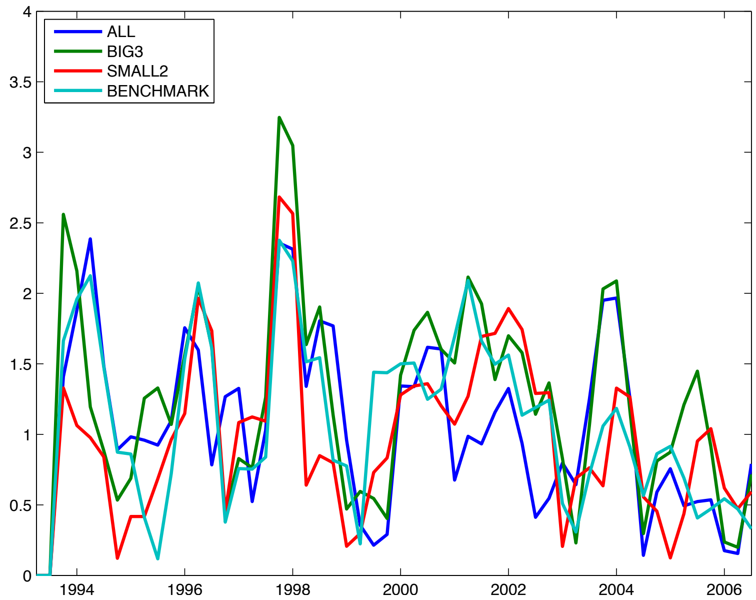
Tab. 8: RMSFE at 1 to 4 qtr, dataset SMALL2

h	1	2	3	4
Unrestricted static no autoreg	74.5	83.4	73.3	78.9
Unrestricted dynamic no autoreg (DD)	69.2	78.2	76.6	73.6
Unrestricted static autoreg	73.1	81.7	73.1	80.0
Unrestricted dynamic autoreg (DDAR)	70.8	78.5	76.3	75.5
Sequential static	87.7	93.8	93.5	95.1
Sequential dynamic (SD)	75.0	89.1	90.1	92.4
Dynamic non-parametric (DNP)	67.6	76.7	82.7	81.0
Common static	88.7	92.3	94.9	95.7
Common dynamic (COMD)	86.7	91.5	93.5	95.1

Tab. 9: Best dataset at 1 to 4 qtr (based on RMSFE)

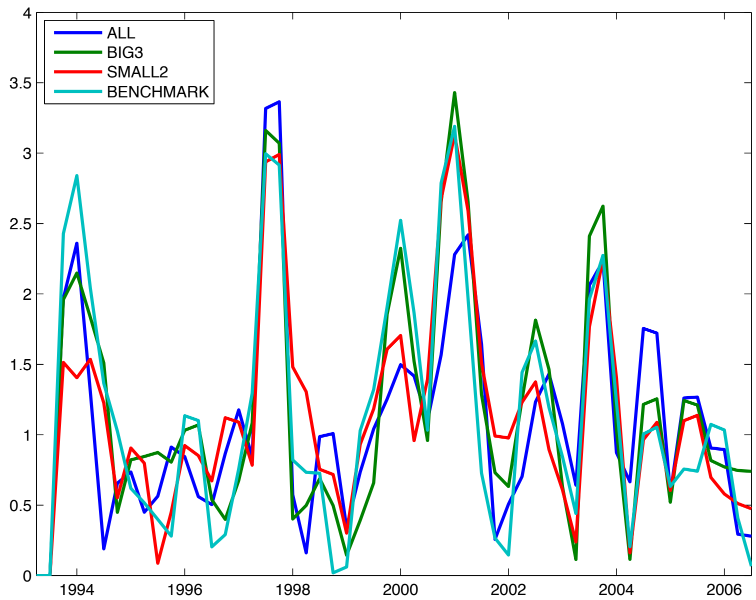
h	1	2	3	4
BENCHMARK	55.0	65.5	62.3	61.6
ALL	57.0	62.0	65.4	55.5
BIG3	63.7	66.8	65.7	64.5
SMALL2	49.3	61.4	65.0	56.9

Fig. 4: OUT OF SAMPLE RMSFE, 1 QTR AHEAD



Note: pooled forecast, moving average over the last 2 quarters.

Fig. 5: OUT OF SAMPLE RMSFE 2 QTR AHEAD



Note: pooled forecast, moving average over the last 2 quarters.

References

[1] D'Agostino A. and Giannone D., (2006), "Comparing alternative predictors based on large-panel factor models", ECB Working Paper, 680.

[2] Angelini E., Henry J. and Marcellino M., (2006), "Interpolation and backdating with a large information set", Journal of Economics Dynamics and Control, 30, pp. 2693-2724.

[3] Bai J. and Ng S. (2002), "Determining the number of factors in approximate factor models", Econometrica, 70, 191-221.

[4] Banbura M. and Runstler G. (2007), "A Look into the Factor Model black box: publication lags and the role of hard and soft data in forecasting GDP", ECB Working Paper, 751.

[5] Boivin J. and Ng S. (2006), "Are more data always better for factor analysis?", Journal of Econometrics, 132, pp. 169-194.

[6] Boivin J. and Ng S. (2005), "Understanding and comparing factor-based forecasts", International Journal of Central Banking, December 2005.

[7] Breitung J. and Schumacher C. (2006), "Real-time forecasting of GDP based on a large factor model with monthly and quarterly data", Deutsche Bundesbank Working Paper.

[8] Cheung C. and Demers F., (2007), Evaluating Forecasts from Factor Models for Canadian GDP Growth and Core Inflation, Banque du Canada, Working Paper 2007/8.

[9] Croux C., Forni M. and Reichlin L. (2001), "A measure for co movement for economic variables: theory and empirics", Review of Economics and Statistics, 83(2), 231-241

[10] Iacobucci A. (2006), "Spectral analysis of economic time series", in New Tools of Economic Dynamics, J. Leskow, M. Puchet, L. F. Punzo (eds.)

[11] Forni, M., Hallin, M., Lippi, M. and Reichlin, L. (2005), "The Generalized Dynamic Factor Model: One-sided estimation and forecasting", Journal of the American Statistical Association, 100(471), pp 830-841.

[12] Forni, M., Hallin, M., Lippi, M. and Reichlin, L. (2000), "The Generalized Dynamic Factor Model: Identification and estimation", Review of Economics and Statistics, 82(4), pp 540-54

[13] Giannone D., Reichlin L. and Sala L. (2005), "Monetary Policy in Real Time". In Mark Gertler and Kenneth Rogoff, editors, NBER Macroeconomics Annual 2004, pp. 161-200. MIT Press.

[14] Giannone D., Reichlin L. and Small D.H., (2006), "Nowcasting GDP and inflation. The real-time informational content of macroeconomic data releases", ECB WP 633.

[15] Kapetanios G. and Marcellino M. (2003), "A Comparison of Estimation Methods for Dynamic Factor Models of Large Dimensions", Working Papers no. 489, Queen Mary, University of London, Department of Economics.

[16] Marcellino M., Stock J.H. and Watson M. (2006), 'A comparison of direct and iterated multistep AR methods for forecasting macroeconomic time series' Journal of econometrics, 135, pp. 499-526.

[17] Marcellino M., Stock J.H. and Watson, M. W. (2003). 'Macroeconomic forecasting in the Euro area: country specific versus Euro wide information', European Economic Review, Vol. 47, pp. 1– 18.

[18] Priestley M.B. (2001), Spectral analysis and time series, Academic Press, 11th edition.

[19] Stock J.H. and Watson, M. (2002a), "Macroeconomic Forecasting Using Diffusion Indexes", Journal of Business & Economic Statistics 20, 147-162.

[20] Stock J. H. and Watson M., (2002b), "Forecasting using principal components from a large number of predictors", Journal of the American Statistical Association, Vol 97, n 460.

European Central Bank Working Paper Series

For a complete list of Working Papers published by the ECB, please visit the ECB's website (http://www.ecb.europa.eu).

862 "Stock market volatility and learning" by K. Adam, A. Marcet and J. P. Nicolini, February 2008.

863 "Population ageing and public pension reforms in a small open economy" by C. Nickel, P. Rother and A. Theophilopoulou, February 2008.

864 "Macroeconomic rates of return of public and private investment: crowding-in and crowdingout effects" by A. Afonso and M. St. Aubyn, February 2008.

865 "Explaining the Great Moderation: it is not the shocks" by D. Giannone, M. Lenza and L. Reichlin, February 2008.

866 "VAR analysis and the Great Moderation" by L. Benati and P. Surico, February 2008.

867 "Do monetary indicators lead euro area inflation?" by B. Hofmann, February 2008.

868 "Purdah: on the rationale for central bank silence around policy meetings" by M. Ehrmann and M. Fratzscher, February 2008.

869 "The reserve fulfilment path of euro area commercial banks: empirical testing using panel data" by N. Cassola, February 2008.

870 "Risk management in action: robust monetary policy rules under structured uncertainty" by P. Levine, P. McAdam, J. Pearlman and R. Pierse, February 2008.

871 "The impact of capital flows on domestic investment in transition economies" by E. Mileva, February 2008.

872 "Why do Europeans work part-time? A cross-country panel analysis" by H. Buddelmeyer, G. Mourre and M. Ward, February 2008.

873 "The Feldstein-Horioka fact" by D. Giannone and M. Lenza, February 2008.

874 "How arbitrage-free is the Nelson-Siegel model?" by L. Coroneo, K. Nyholm and R. Vidova-Koleva, February 2008.

875 "Global macro-financial shocks and expected default frequencies in the euro area" by O. Castrén, S. Dées and F. Zaher, February 2008.

876 "Are sectoral stock prices useful for predicting euro area GDP?" by M. Andersson and A. D'Agostino, February 2008.

877 "What are the effects of fiscal policy shocks? A VAR-based comparative analysis" by D. Caldara and C. Kamps, March 2008.

878 "Nominal and real interest rates during an optimal disinflation in New Keynesian models" by M. Hagedorn, March 2008.

879 "Government risk premiums in the bond market: EMU and Canada" by L. Schuknecht, J. von Hagen and G. Wolswijk, March 2008.

880 "On policy interactions among nations: when do cooperation and commitment matter?" by H. Kempf and L. von Thadden, March 2008.

881 “Imperfect predictability and mutual fund dynamics: how managers use predictors in changing systematic risk” by G. Amisano and R. Savona, March 2008.

882 “Forecasting world trade: direct versus “bottom-up” approaches” by M. Burgert and S. Dées, March 2008.

883 “Assessing the benefits of international portfolio diversification in bonds and stocks” by R. A. De Santis and L. Sarno, March 2008.

884 “A quantitative perspective on optimal monetary policy cooperation between the US and the euro area” by S. Adjemian, M. Darracq Pariès and F. Smets, March 2008.

885 “Impact of bank competition on the interest rate pass-through in the euro area” by M. van Leuvensteijn, C. Kok Sørensen, J. A. Bikker and A. A. R. J. M. van Rixtel, March 2008.

886 “International evidence on sticky consumption growth” by C. D. Carroll, J. Slacalek and M. Sommer, March 2008.

887 “Labor supply after transition: evidence from the Czech Republic” by A. Bičáková, J. Slacalek and M. Slavík, March 2008.

888 “House prices, money, credit and the macroeconomy” by C. Goodhart and B. Hofmann, April 2008.

889 “Credit and the natural rate of interest” by F. De Fiore and O. Tristani, April 2008.

890 “Globalisation, domestic inflation and global output gaps: evidence from the euro area” by A. Calza, April 2008.

891 “House prices and the stance of monetary policy” by M. Jarociński and F. Smets, April 2008.

892 “Identification of New Keynesian Phillips Curves from a global perspective” by S. Dées, M. H. Pesaran, L. V. Smith and R. P. Smith, April 2008.

893 “Sticky wages: evidence from quarterly microeconomic data” by T. Heckel, H. Le Bihan and M. Montornès, May 2008.

894 “The role of country-specific trade and survey data in forecasting euro area manufacturing production: perspective from large panel factor models” by M. Darracq Pariès and L. Maurin, May 2008.

ISSN 1561-0810