Modelling euro area GDP growth with dynamic factor models. *

Andrei Tănase, Valeriu Nalban
National Bank of Romania

June, 2012

Keywords: GDP estimation and forecast, dynamic factor models.
JEL codes : E27, E58, O11

Abstract

Using structural models to estimate and forecast quarterly GDP conditional on a large number of economic indicators raises parsimony problems, such as the need to define many identification restrictions and also to deal with very sparse asymptotic distributions of the coefficients. Factor models have the role to fill this gap in macroeconomic modelling. The basic assumption is that the correlation between the observed variables can be sufficiently well explained by a few common orthogonal factors. The estimated factors can then be used to estimate any variable of interest. In the empirical section of this paper, using the Area Wide Model database, we obtain near-term-forecasts (NTF) of euro area GDP using the dynamic factor model. The specification of the model is such that we can decompose the deviations of the NTF's from the long run mean, which can be interpreted in a similar manner as the output gaps.

*Correponding author: Andrei Tănase, Senior Economist, National Bank of Romania, Lipscani street no. 25, 3rd district, 030031, Bucharest, e-mail: `andrei.tanase@bnro.ro`
The authors wish to acknowledge very useful suggestions from department colleagues and participants at past seminars.

1 Introduction

Current quarterly GDP is published with a delay of almost two months. In order to estimate GDP, analysts have to use information that is correlated with economic activity and that is released with higher frequency. Available information is made of real economy indicators, confidence indicators, financial and external environment data.

Under the structural VAR framework, using both core variables (that are usually included in macroeconomic models) and additional indicators raises parsimony issues related to the number of causality restrictions needed for shocks identification and the quality of significance tests associated to the estimated coefficients. The basic assumption behind factor models is that dynamics and correlation between variables can be sufficiently well explained by a few orthogonal unobserved common factors. These can then be used as regressors for GDP growth and therefore solve the parsimony issues highlighted before. This kind of modelling has initially been supported by Sargent and Sims [20], that constructed a coincident and a leading indicator for the US economy. They have been followed, among others, by Engle and Watson [8], Quah and Sargent [19] and Forni and Reichlin [11]. Recent developments of factor models have been proposed by Forni et al [12] [13] and Stock and Watson [23], [24] that characterize the factors as having a dynamic structure and yielding the class of dynamic factor models (DFM). The majority of these papers contain empirical applications based on GDP estimation and forecast. The resulting estimates are interpreted as coincident and leading indicators of the economy. Other dynamic factor models applications on macroeconomic data are found in papers of Barhoumi K. et al [6], Angelini et al. [2], Jakaitiene and Dees [16], Forni et al [12], Stock and Watson [23], Altissimo et al. [1], Banerjee and Marcellino [3], Banerjee et al. [4], Breitung and Eickmeier [5] and Schumacher and Breitung [22].

In the empirical application of this paper, we employ the factor estimation methodology of Forni, Hallin, Lippi and Reichlin (FHLR) in order to estimate and forecast euro zone GDP conditioning on a large number of economic and financial variables contained in the Area Wide Model database with quarterly series as designed in Fagan et al. [9] and updated to year 2009. Unlike Stock and Watson (SW) who estimate the common factors with the principal components extracted from the correlation matrix of the observed variables, FHLR estimate the factors from the spectral density matrix. The advantage of their method is apparent when the effect of the factors on the observed variables is heterogeneous over time. This should be the case of economic variables, given that we can observe leading indicators for GDP (for example, some labour market indicators or some confidence indicators).

The rest of the paper is organized as follows: the second section presents the factor model, starting from the static to the dynamic version and then the FHLR estimation method, the third section is dedicated to the empirical application and the last section contains the summary and conclusions of the paper.

2 Common factors models

The assumption behind common factor models is that variability in the observed data can be well explained by a few orthogonal factors. The usual way of estimating these models is by principal components analysis. This is based on the spectral decomposition of the covariance matrix. The common factors are estimated with the eigenvectors. Each orthogonal eigenvector has an associated eigenvalue that is proportional to the explained variance and the higher is the associated eigenvalue of the factor, the more it explains in terms of variance of the data. If the series are highly correlated (as in the case of economic variables) then the first few principal components cover the majority of the variance of the observed data.

2.1 The static factor model

Let $\{X_{n,T}\}$ be the set of observed variables, where n is the number of variables and T is the number of observations. For simplicity, we assume that the variables are demeaned and standardized. Moreover, let Σ_X be the covariance matrix of the data. As stated before, the basic assumption of the common factor model is that $q < n$ common unobserved factors explain the majority of the variance in the data. Under linearity assumption between the observed variables and the factors, the model is

$$X_{[n\times T]} = \beta F_{[q\times T]} + \epsilon \tag{1}$$

where $\{F_{q,T}\} = \{f_{j,t}, j = 1\dots q, t = 1\dots T\}$ is the vector of common factors and $\beta_{[n\times q]}$ are the factor loadings. The errors $\epsilon_{[n\times T]}$ are uncorrelated with the factors. The identification condition for the factors is their orthogonality.

Given that the factors are defined as having mean zero and unit variance, from (1) in comes that the variance of the data Σ_X can be written as $\Sigma_X = \beta\beta' + \Sigma_\epsilon$.

For $q = n$, the common factor model is equivalent to principal component analysis. This consists of writing the variance of observed variables $\{x_{i,t}, i = 1\dots n, t = 1\dots T\}$ as the variance of orthogonal principal components $\{r_{i,t}, i = 1\dots n, t = 1\dots T\}$ (see, for example, Tsay [26]).

Deriving the principal components is based on the following property of any positive definite symmetrical matrix Σ_X (see, for example, Peracchi [18]): given the eigenvalues $\lambda_1 \geq \lambda_2 \geq \ldots \lambda_n$, then there is a matrix P such that

$$\Sigma_X = P\, diag[\lambda_i]\, P' \qquad \text{unde} \qquad P'P = I_n \tag{2}$$

where $P = \{P_i = (p_{i,1}, \ldots p_{i,n})', i = 1 \ldots n\}$ is the set of eigenvectors of Σ_X. The principal components, equal to the first q eigenvectors that have the largest associated eigenvalues, explain the following proportion of variance

$$\frac{\sum_i^q Var(r_i)}{\sum_i^n Var(r_i)} = \frac{\sum_i^q \lambda_i}{\sum_i^n \lambda_i}$$

The equivalence between principal component analysis and factor models suggests to estimate the first q factors with the first q principal components. Then, from (2), it comes out that the estimated factor loadings are $\hat{\beta} = (\hat{\beta}_1, \ldots, \hat{\beta}_q) = (\sqrt{\hat{\lambda}_1}\hat{p}_1, \ldots, \sqrt{\hat{\lambda}_q}\hat{p}_q)$. The *communality* is $c_i = \hat{\beta}_{i,1}^2 + \ldots \hat{\beta}_{i,q}^2$, while the *uniqueness* is $(1 - c_i)$.

2.2 The dynamic factor model (DFM)

In order to formalize the dynamic factor model, we will follow the notation in Forni, Hallin, Lippi şi Reichlin (FHLR), references [12] and [13].

Let $\{X_{n,T}\} = \{x_{i,t}, i = 1 \ldots n, t = 1 \ldots T\}$ be the set of observed variables. Each variable x_i is modelled as the sum of its common component χ_i and a specific shock ξ_i. The common component is the linear projection of the variable of the common factors:

$$x_{i,t} = \chi_{i,t} + \xi_{i,t} \tag{3}$$

$$\chi_{i,t} = b_{i,1}(L) f_{1,t} + b_{i,2}(L) f_{2,t} + \ldots + b_{i,q}(L) f_{q,t} \tag{4}$$

where L is the *lag* operator and $\{B_{q,T} = b_{i,j}(L), i = 1 \ldots n, j = 1 \ldots q\}$ is the set of time varying factor loadings. These are assumed to satisfy standard scale conditions, more precisely the sum of their squares is finite. The dynamic factor model is analog to *state-space* models, with a measurement equation for the observed variables (3) and a transition equation for the common components (4). Therefore, the potentiality of dynamic factor models for estimation and forecasting application is high.

Both Forni, Hallin, Lipppi and Reichlin (FHLR, references [12] şi [13]) and Stock şi Watson (SW, [23] şi [24]) have proposed to estimate the common factors by principal component analysis. The difference between the two approaches lies in the source of the principal components: SW have a classical approach, using the covariance matrix, while FHLR use the spectral density matrix. Theoretically speaking, the two approaches are equivalent. In practice, FHLR approach is to be preferred when the influence of the factors on the observed variables is heterogeneous, that is each factor inflence different lags of the observed variables. Moreover, FHLR methodology include a nonparametric estimation of the spectral density (Bartlett filter) and this ensures a higher precision when the data is short.

Filtering techniques proposed by FHLR

FHLR propose two filtering techniques in order to estimate the common components of the variables. The first one is a two-sided filter and this carries on the usual end-of-sample problems of this kind of filters. The second one is a one-sided filter and this is obtained with a sequential procedure.

Let $\Sigma(\theta)_X$ denote the spectral density matrix of the data $\{X_{n,T}\}$, computed at frequencies $\{\theta_1, \ldots, \theta_{2M+1}\}$, equally distant in the interval $[-\pi, \pi)$. Moreover, let $\{\lambda_j, \mathbf{p}_{nj}^T(\theta_h), j = 1 \ldots q\}$ be the set of q eigenvalues and associated eigenvectors computed at frequencies θ_h. The common components estimated with the two-sided filter are

$$\hat{\chi}_{i,t} = \sum_{k=-M}^{M} K_{i,k} L^k X_{i,t}, \tag{5}$$

where

$$K_{i,k} = \frac{1}{2M+1} \sum_{h=0}^{2M} K_i(\theta_h) \exp(ik\theta_h)$$

and

$$K_i^T(\theta_h) = \tilde{p}_{1,i}^T(\theta_h)\mathbf{p}_1^T(\theta_h) + \ldots + \tilde{p}_{q,i}^T(\theta_h)\mathbf{p}_q^T(\theta_h)$$

integrate the eigenvectors of $\Sigma(\theta)_X$ with the inverse Fourier transformation. In the above equations, L^k is the lag operator of order k and $\tilde{p}$ is the complement of vector p. As for the frequecies to be used, these are set as $\theta_h = 2\pi h/(2M+1)$.

The spectral density is estimated nonparametrically with the Bartlett kernel considering periods $\{(t-M)\ldots t \ldots (t+M)\}$. Parameter M (not necessarily equal to the parameter of the spectral density frequencies) is chosen arbitrarily, one criterium being the frequency of the data[1].

[1] In the empirical application on quarterly data, parameter M was fixed as being equal to 4.

Consistency of the above estimator is not altered if the error terms $\{\xi_{i,t}, i = 1 \dots n, t = 1 \dots T\}$ are dependent. Moreover, if the number of common factors chosen by the analyst is higher than the true one, then convergence of the estimated factors to the true ones does not change significantly.

The one-sided filter is a modified version of the two-sided filter and it proves to have better end-of-sample properties. Let $X_{i,T+h}$ denote the common component of variable X_i for period $(t+h)$. The forecasted common component $\hat{X}_{i,T+h}$ is

$$\hat{\chi}_{i,T+h\,|\,T} = \sum_{i'=1}^{n} \left[\Gamma^{\chi}_{nh} Z_n' \left(Z_n \Gamma^{T}_{n0} Z_n' \right)^{-1} Z_n \right] x_{i'T} \tag{6}$$

Here $\Gamma^{\chi^T}_{nh}$ is the covariance matrix of the common components for the h^{th} lag, Z_n^T groups the eigenvectors of matrix $(\Gamma^{\chi^T}_{nh} + \Gamma^{\xi^T}_{nh}, \Gamma^{\xi^T}_{nh})$, where $\Gamma^{\xi^T}_{nh}$ is the covariance matrix of order h for the specific shocks. Setting $h = 0$ one gets the *in-sample* version of the filter, used for estimating contemporary values of the common components.

Therefore, the common components are linear combinations of the common factors and these are orthogonal linear combinations of the factors. The above equality allows to derive the contribution of each variable to the common components. This one-sided filter is used in the following empirical application in order to estimate and forecast GDP.

3 Empirical application

In this section we develop an empirical application in which wee use the one-sided filter FHLR detailed in equation (6) to estimate and forecast the quarterly growth of euro zone GDP. The obtained values are referred to as *coincident* and *leading* indicators for the economic activity. The specification of the model is such that we can decompose the deviations of the NTF's from the long run mean, which can be interpreted in a similar manner as the output gaps. The software that we use is Matlab, v2010. The primary version of the codes (including the routine for cross-section estimation of the common component) was downloaded from the authors home page (see reference [27]).

The data

The conditioning information consists of around 50 variables that make up the Euro Area Wide model database as designed in Fagan et al. [9] and updated to year 2009. Moreover, we add euro area sectoral confidence indicators as published by the European Commission. The database is composed of quarterly series that are transformations of original series with daily, monthly or quarterly frequency. The process of database construction is detailed on an explanatory note available on EABCN website.

Tables 1 and 2 present the variables used in the application. The first table show the output components and the second table comprises all other regressors used in the model. We have grouped the explanatory variables as GDP components and their deflators, price indices, the balance of payments and external environment, labour market indicators, financial variables and confidence indicators.

The correlation coefficients between GDP and the selected variables are also included in the tables: $\rho(0)$ denotes the contemporaneous correlation, while $\rho(-1)$ and $\rho(1)$ denote the correlation between GDP and the first *lag* and the first *lead* of the variable. Values of $\rho(-1)$ that are close to 1 (in absolute value) underlines the leading feature of the variable for GDP (for example, private consumption, world GDP and number of unemployed). Moreover, the analysis reveals some variables that are poorly correlated with the growth rate of GDP (such as the public sector spending, some of the price indexes or world GDP deflators). These series are kept in the application to ensure the group representativeness.

The raw data that we use from Euro Area Wide model database ranges between 1995 Q1 and 2009 Q4. This interval mixes periods of economic boom (1997-1999, 2003-2007), moderated growth and economic recession respectively (2008-beginning of 2009); moreover, the sample includes the last quarters of 2009, with mild signs of economic recovery.

The vast majority of the variables are transformed as quarterly growth rates. This ensures stationarity of the series. Exceptions are the interest rates, the unemployment and the confidence indicators. The final data spans over the 1995Q2 - 2009Q4 (59 observations).

The estimation exercise

The results of a principal component analysis on the full sample reveals that the first three components explain roughly 70% of the variability in the data. Extrapolating this result, the number of common dynamic factors q is set equal to 3. As for the number of lags for the factors, this is set as $M = 4$. Therefore, the model that we estimate is

$$y_t = b_1(L^4)f_{1,t} + b_2(L^4)f_{2,t} + b_3(L^4)f_{3,t} + \xi_{y,t}$$

The total number of frequencies is set equal to 4 and they are equally spaced in $[0, 2\pi]$ interval[2].

[2]Setting the same number for the frequencies and the lags has computational advantages.

Each estimation exercise consists of estimation and forecast of GDP growth by employing filter (6) with $h = 0$ and $h = 1$. The samples are of expanding window type. The initial sample has $T_0 = 20$ observations and, for the subsequent ones, we add one observation at a time, totalling 39 samples. The first sample (1995Q2-1999Q1) is used for obtaining the GDP estimate for 1999Q1 and the forecast for 1999Q2. The last sample (2004Q3-2009Q4) is used for obtaining the GDP estimate for 2009Q4 and the forecast for 2010Q1. For each exercise the variables are standardized. This aspect is important for the interpretation of results as deviations from long-run mean and computation of estimates decomposition.

An important aspect of the estimation exercise regards the transformations of the series such as to resemble a real time estimation exercise. In general, the data is not available for the entire current quarter (for example, only one or two monthly observations are available). In this case the analyst uses the average or the latest released figure as an estimator of the quarterly figure. In order to replicate this bias between the quarterly figure and the estimate, we impose an observation error that has a normal distribution with mean 0 and standard deviation equal to the empirical standard deviation of the series. Another practical aspect of the estimation exercise is that for GDP and GDP components, the end-of-sample value is replaced with a missing value such as to exclude any information regarding the current exercise.

The model is designed for balanced panels only. Therefore, any missing data must be either imputed or replaced by estimates. We do this with the same methodology. First we estimate the factor loadings of the one-sided filter (6) with $h = 0$ on the available data. Then, the factor loadings are used to estimate the missing values for the incomplete series (either backward or forward estimation). This method is similar to other filtering applications for ragged-edged data (see, for example Marcellino şi Schumacher [17]).

Uncertainty of the estimates is quantified by constructing bootstrap confidence intervals with a cyclical bootstrap method as in Davidson and MacKinnon [7] and Fitzenberger [10]. This method is especially designed for time dependent data. The bootstrap series are constructed by appending randomly chosen errors blocks of length ($S = 4$). The steps of each bootstrap replication are: (i) computation of the model errors; (ii) random extraction and merging of blocks of errors of size ($S = 4$), denoted with ε_t^b; (iii) construction of pseudo-series $y_t^{b\star} = \hat{Y}_t + \varepsilon_t^b$; (iv) estimation and forecast of the common components on the $y_t^{b\star}$ series. Forecast errors are multiplied by an asymmetric uncertainty parameter that allows to make an assumption on some positive or negative scenarios:

$$Y_t^{h\star} = \hat{Y}_t + \lambda_1 * \mathbb{1}\{\varepsilon_t^b > 0\}\varepsilon_t^b + \lambda_2 * \mathbb{1}\{\varepsilon_t^b < 0\}\varepsilon_t^b$$

where $\lambda_1 = 2$ and $\lambda_2 = 1.5$, allowing for a slightly larger probability on positive scenarios.

Table 1: GDP components

Ind	Code	Name	Transformation	$\rho(-1)$	$\rho(0)$	$\rho(1)$
1	YER	GDP (Real)	q-o-q			
2	PCR	Private Consumption	q-o-q	0.5	0.7	0.5
3	GCR	Gov. Consumption	q-o-q	-0.2	0.0	-0.2
4	ITR	Gross Investment	q-o-q	0.5	0.9	0.6
5	XTR	Exports of Goods and Services (Real)	q-o-q	0.6	0.9	0.6
6	MTR	Imports of Goods and Services (Real)	q-o-q	0.5	0.9	0.7
7	YIN	GDP, Income Side	q-o-q	0.6	0.9	0.7
8	WIN	Compensation to Employees	q-o-q	0.4	0.6	0.7
9	GON	Gross Operating Surplus	q-o-q	0.4	0.9	0.5
10	TIN	Indirect Taxes (net of subsidies)	q-o-q	0.5	0.5	0.5
11	YFN	GDP at Factor Costs (WIN+GON)	q-o-q	0.5	0.9	0.7

Table 2: The additional groups of variables and their correlations with GDP growth

Ind	Code	Name	Transformation	$\rho(-1)$	$\rho(0)$	$\rho(1)$
Deflators						
12	YED	GDP Deflator	q-o-q	-0.2	0.0	0.3
13	PCD	Consumption Deflator	q-o-q	0.3	0.5	0.6
14	GCD	Gov. Consumption Deflator	q-o-q	0.0	-0.1	0.0
15	ITD	Gross Investment Deflator	q-o-q	0.2	0.5	0.6
16	XTD	Exports of Goods and Services Deflator	q-o-q	0.2	0.6	0.6
17	MTD	Imports of Goods and Services Deflator	q-o-q	0.4	0.6	0.5
18	YFD	GDP at Factor Costs Deflator	q-o-q	-0.3	-0.1	0.1
Prices						
19	HICP	Overall HICP (Non-sesonally adjusted)	q-o-q	0.1	0.2	0.2
20	HEX	HICP excluding energy (Non-seasonally adjusted)	q-o-q	-0.1	0.0	0.0
21	HEG	HICP energy	q-o-q	0.3	0.5	0.4
22	HICPSA	Overall HICP (Seasonally adjusted)	q-o-q	0.2	0.4	0.4
23	HEXSA	HICP excluding energy (Seasonally adjusted)	q-o-q	-0.2	-0.1	0.1
24	HEGWEI	Weight of the HICP energy on overall HICP	q-o-q	0.0	0.1	0.1
25	COMPR	Commodity Prices	q-o-q	0.5	0.4	0.1
26	POILU	Oil prices (in USD)	q-o-q	0.4	0.4	0.1
27	PCOMU	Non oil commodity prices (in USD)	q-o-q	0.5	0.5	0.1
BOP and external environment						
28	CAN_YEN	Current Account Balance/GDP	q-o-q	-0.2	-0.1	0.0
29	NFN_YEN	Ratio, Net Factor Income from Abroad/GDP	q-o-q	0.5	0.4	0.3
30	YWD	World GDP Deflator	q-o-q	-0.1	0.1	0.4
31	YWDX	World Demand Deflator, Composite Indicator	q-o-q	0.1	0.3	0.4
32	YWR	World GDP	q-o-q	0.7	0.8	0.5
33	YWRX	World Demand, Composite Indicator	q-o-q	0.7	0.9	0.7
Labour market						
34	LFN	Labour Force (persons)	q-o-q	0.3	0.3	0.4
35	LNN	Total Employment (persons)	q-o-q	0.5	0.7	0.8
36	UNN	Number of Unemployed	q-o-q	-0.5	-0.8	-0.8
37	URX	Unemployment rate (% labour force)	level	0.3	0.2	0.0
38	LEN	Employees (persons)	q-o-q	0.5	0.7	0.8
39	LPROD	Labour Productivity (YER/LNN)	q-o-q	0.6	0.9	0.4
40	ULC	Unit Labour Costs(WIN/YER)	q-o-q	-0.5	-0.7	-0.2
41	WRN	Wage per head	q-o-q	0.1	0.2	0.3
Financial variables						
42	STN	Short-Term Interest Rate (Nominal)	level	-0.2	0.1	0.3
43	LTN	Long-Term Interest Rate	level	0.0	0.1	0.1
44	SAX	Household's savings ratio	q-o-q	-0.4	-0.5	-0.3
45	EEN	Effective exchange rate (EER12)	q-o-q	0.1	0.1	0.2
46	EXR	Euro per USD exhange rate	q-o-q	-0.1	-0.1	0.1
Confidence indicators						
47	EA_INDU	Industrial confidence indicator (40%)	level	0.3	0.6	0.8
48	EA_SERV	Services confidence indicator (30%)	level	0.5	0.7	0.7
49	EA_CONS	Consumer confidence indicator (20%)	level	0.4	0.6	0.7
50	EA_RETA	Retail trade confidence indicator (20%)	level	0.2	0.4	0.5
51	EA_BUIL	Construction confidence indicator (5%)	level	0.2	0.3	0.4
52	EA_ESI	The Economic sentiment indicator	level	0.5	0.7	0.8

Estimation results

Figure 1 shows the deviation of GDP NTFs from corresponding long run mean, produced at each quarter. The variables are grouped according to the list in Tables 1 and 2. In the first three years of the sample corresponding to moderate economic growth, the gaps are slightly negative. Between 2004 and 2008Q2 the gaps from the long run mean have been minor, associated with stable economic growth. For the next quarters, the variables have significant negative contributions, in particular the output components, external environment variables and financial market variables. During this downturn period, consumption and commercial activities have dropped and this has been reflected also in the confidence indicators. After ECB started to use expansionary monetary policy tools (reducing key rate, refinancing operations), the contribution of financial variables changed significantly and was assessed to be drastically reduced or even slightly positive starting with 2009Q3. Exports recovery and accommodative monetary policy in other economies (especially quantitative easing measures used by Fed and Bank of England) reduced the negative contribution of external environment variables also. These signs of sluggish recovery improved the confidence indicators, the overall estimation pointing to pre-crisis values for the deviation of GDP NTF form its long run mean in 2009Q4 and 2010Q1 respectively.

Figure 2 shows the estimated common component of the GDP and a forecast of GDP using a benchmark AR(2) model, alongside with the actual quarterly GDP growth. Until global economic crises triggered in 2008, quarterly GDP growth rates and associated NTFs were positive and with an increasing tendency, ranging between 0.1% and 0.8%. Accordingly, the absolute forecast errors were small, with a mean of 0.2 percentage points for 2000Q1-2007Q4 period. The GDP growth in 2008-2009 is characterized as being highly volatile: positive observations until mid 2008, followed by five negative quarters (with a minimum of -2.6% in 2009Q1) and a positive observation in 2009Q3. During this interval, the dynamic filter smoothen the extreme values and therefore the estimated common component has a minimum of just -1.8% in 2009Q1. Associated NTFs produce greater forecast errors, but DFM resulted in smaller mean absolute forecast error compared to the benchmark AR(2) model (0.7 and 1 percentage points respectively). Overall, DFM proves to have better forecasting properties as compared to the benchmark model, being especially useful during turbulent periods. This is due to the capacity to rapidly incorporate additional information on the economic downturn and adjust the estimates accordingly. In addition, we underline the higher stability of the factor model estimates as compared to the AR(2) model.

Moreover, in Figure 2, we plot a fan-chart for the estimated GDP growth for 2009Q4 and the forecast for 2010Q1. The fan-chart represents confidence intervals from 50% to 90% probability mass around the point estimate. These intervals are constructed with the bootstrap technique detailed in the previous section. For the last quarter of 2009 and the first quarter of 2010 the expected GDP growth was -0.1% (with a 90% confidence interval of -0.5:0.2) and 0.25% (with a 90% confidence interval of -0.7:0.7). The first point estimate shows a continuation of negative dynamics of the economy, albeit moderated as compared to the first half of 2009. For 2010Q1, point forecast is positive. The confidence intervals are not symmetric around the point estimates and they assign almost the same probability to both positive and negative values underlining therefore the high degree of uncertainty regarding the continuation of the recession.

4 Summary and conclusions

The dynamic factor model is useful for obtaining estimates conditional on many regressors. In the application of this paper, the model is employed for estimation and forecast of euro zone quarterly GDP conditional on a large number of indicators characterizing the real economy, the financial economy and the global environment . This way, the model can quickly adjust the estimates by capturing real activity swings, unlike a benchmark AR(2) model that is not able to benefit from information provided by such leading indicators. Until 2008Q3, the gaps from the long run mean are relatively low. During the recession, the gaps are significantly negative and they are explained mainly by the output components, external environment variables and financial market variables. In the last quarters of 2009, signs of sluggish recovery are pointed by the improvement of all classes of variables, in particular the financial variables, the confidence indicators and the external environment.

Figure 1: Deviation of NTF GDP from long run mean

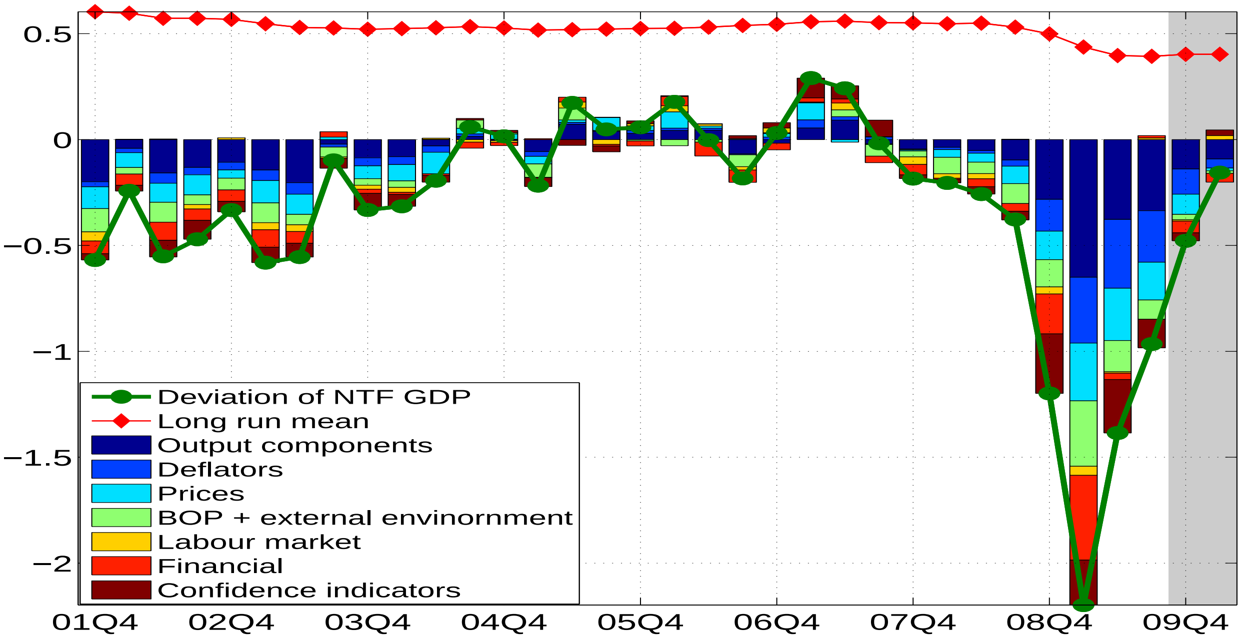
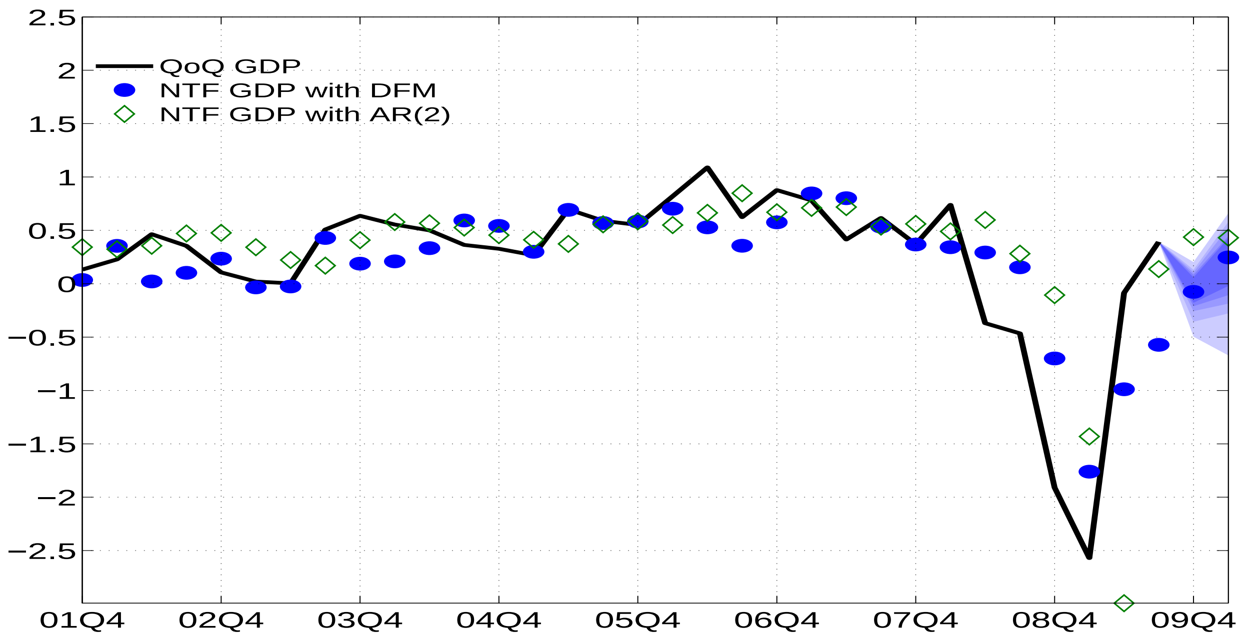


Figure 2: Near Term Forecast of GDP growth



References

[1] Altissimo F., Cristadoro R., Forni M., Lippi M., Veronese G. New eurocoin: tracking economic growth in real time. *CEPR, discussion paper* 2006; **5633.**

[2] Angelini E., Bańbura M., Gerhard Rünster. Estimating and Forecasting the Euro Area Monthly National Accounts from a Dynamic Factor Model. *ECB, WP* 2008; **953**.

[3] Banerjee A., Marcellino M. Are there any reliable leading indicators for US inflation and GDP growth? *Ineternational Journal of Forecasting* 2006; **22**: 137–151.

[4] Banerjee A., Marcellino M., Masten I. Leading Indicators for Euro Area Inflation and GDP growth? *EABCN* 2005.

[5] Breitung J., Eickmeier S. Dynamic factor models. *Deutsche Bundesbank, DP* 2005; **38**.

[6] Barhoumi K., Benk S., Cristadoro R., Reijer A.D., Jakaitiene A., Jelonek P., Rua A., Rnstler G., Ruth K., Nieuwenhuyze C. Short-term Forecasting of GDP using Large Monthly Datasets. A Pseudo-real Time Forecast Evaluation Exercise *ECB, OP* 2008; **84**.

[7] Davidson R., MacKinnon J.G. *Econometric Theory and Methods*. Oxford Univ. Press: NY, 2003.

[8] Engle R., Watson M.W. A One-Factor Multivariate Time Series Model of Metropolitan Wage Rates. *Journal of the Ameriacan Statistical Association* 1981; **76**: 774–781.

[9] Fagan G., Henry J., Mestre R. An area wide model (AWM) for the euro area *ECB, WP* 2001; **42**.

[10] Fitzenberger, B., 1998. The moving blocks bootstrap and robust inference for linear least squares and quantile regressions. *Journal of Econometrics* **82**, 235–287.

[11] Forni M., Reichlin L. Dynamic common factors in large cross-sections. *Empirical Econmics* 1996; **82**: 27–42.

[12] Forni M., Hallin M., Lippi M., Reichlin L. The Generalized Dynamic-Factor Model: Identification and Estimation. *The Review of Economics and Statistics* 2000; **82**: 540–554.

[13] Forni M., Hallin M., Lippi M., Reichlin L. The Generalized Dynamic-Facotor Model One Sided Estimation and Forecasting. *Working paper* 2003, final version published in *Journal of the Ameriacan Statistical Association* 2005; **100**: 830–840.

[14] Hamilton J.D. *Time series analysis*. Princeton University Press: New Jersey, 1994.

[15] Harvey A.C. Sepctral Analysis in Economics. *The Statistician* 1975; **24**: 1–36.

[16] Jakaitiene A., Dées S. Forecasting the world economy in the short term. *ECB, WP* 2009; **1059**.

[17] Marcellino M., Schumacher C. Factor-MIDAS for now- and forecasting with ragged-edged data: a model comparison for German GDP. *ECB, Discussion Paper* 2007; **34**

[18] Peracchi F. *Econometrics*. Wiley: Chichester (UK), 2001.

[19] Quah D., Sargent T.J. A Dynamic Index Model for Large Cross Sections. *Business Cycles, Indicators and Forecasting* 1983.

[20] Sargent T.J., Sims C.A. Business Cycle Modeling without Pretending to Have Too Much a Priori Economic Theory. *Federal Reserve Bank of Minneapolis, Working paper series* 1977.

[21] Schneider C., Arminger G. Factor analysis for extraction of structural components and prediction in time series. *Advances in Data Analysis*. Springer, 2007; 273 – 280.

[22] Schumacher C., Breitung J. Real-time forecasting of GDP based on a large factor model with monthly and quarterly data. *Deutsche Bundesbank, discussion paper* 2006; **33**

[23] Stock J.H., Watson M.W. Macroeconomic Forecasting Using Diffusion Indexes. *Journal of Business and Economic Statistics* 2002; **50**.

[24] Stock J.H., Watson M.W. Forecasting Using Principal Components from a Large Number of Predictors. *Journal of the Ameriacan Statistical Association* 2002; **97**: 1167 – 1179.

[25] Stock J.H., Watson M.W. Combination forecasts of output growth in a seven-country data set. *Journal of Forecasting*; **23**: 405 – 430.

Forecasting Using Principal Components from a Large Number of Predictors. *Journal of the Ameriacan Statistical Association* 2002; **97**: 1167 – 1179.

[26] Tsay R.S. *Analysis of financial time series*. Wiley: New Jersey, 2005.

[27] *http://www.economia.unimore.it/forni_mario/*