The Euro-Sting revisited: the usefulness of financial indicators to obtain euro area GDP forecasts*

Maximo Camacho[a] and Agustin Garcia-Serrador[b]

a Universidad de Murcia

b BBVA Research

Abstract

This paper uses an extension of the Euro-Sting single-index dynamic factor model to construct short-term forecasts of quarterly GDP growth for the euro area by accounting for financial variables as leading indicators. From a simulated real-time exercise, the model is used to investigate the forecasting accuracy across the different phases of the business cycle. Our extension is also used to evaluate the relative forecasting ability of the two most reliable business cycle surveys for the euro area, the PMI and the ESI. We show that the latter produces more accurate GDP forecasts than the former. Finally, the proposed model is also characterized by its great ability to capture the European business cycle, as well as the probabilities of expansion and/or contraction periods.

Key words: real-time forecasting, eurozone GDP, business cycles

JEL classification: E32, C22, E27

---

* We thank R. Doménech, M. Jiménez and C. Ulloa for helpful comments. We also thank the editor and two anonymous referees for their valuable comments. M. Camacho would like to thank CICYT and Fundación Seneca for their support through grants ECO2010-19830 and 11998/PHCS/09. All the remaining errors are our own responsibility.

Corresponding author: Maximo Camacho, Universidad de Murcia, Facultad de Economía y Empresa, Departamento de Métodos Cuantitativos para la Economía, 30100, Murcia, Spain. E-mail: mcamacho@um.es

1. Introduction

The financial crisis erupted in the last 2007 called into question the ability of traditional forecasting methods which were unable to anticipate the turning point in time. This forecasting failure triggers some proposals that tried to improve upon the early detection of the unforeseen downturn. Focusing on euro area forecasting, the small scale dynamic factor model suggested by Camacho and Perez Quiros (2010) has been one of the most successful in computing accurate and timely assessments of short-term GDP developments. The good forecasting performance of their model comes from the fact that it mixes quarterly and shorter frequencies and that it deals with asynchronously published economic indicators which allow the model to compute real-time forecasts on the basis of timely updated data.

Although their model mixes the information contained in hard and soft data, it does not include financial indicators. However, according to Wheelock and Wohar (2009), financial variables (such as the slope of the yield curve) can be helpful as *leading* indicators in forecasting growth. In this context, the main contribution of this paper is the analysis of financial time series as leading indicators of output growth, in a dynamic factor model that accounts for asynchronous co-movements between the financial and the real activity indicators.

A further drawback of the analysis developed in Camacho and Perez Quiros (2010) is that their sample dated from the early nineties to 2007 which includes only one recession. In addition, their real-time forecasting analysis started in 2003 and ended in 2007 which become a period of relatively stable high growth. To overcome this potential shortcoming, we use enlarged historical time series of euro area GDP growth rate which are obtained from the AWM Database elaborated for the Euro Area Business Cycle Network (€ABCN).[1] According to its methodology, the historical data since the first quarter of 1970 are based on the aggregation of available country data. The main source is Eurostat, complemented by the OECD National Accounts, the OECD Main indicators, the BIS and the AMECO databases. The new data set used in this paper is now dated back to the early eighties which allow us to develop the following twofold exercise. First, since the common dynamic factor is now available since the eighties, we use recent techniques to support the view that the factor can become a reliable economic indicator of the euro area business cycle.

[1] The AWM Database and its methodology are available on the website of €ABCN (http://www.eacbn.org/area-wide-model)

Second, we develop a pseudo real-time exercise to evaluate the performance of the model to compute short-term forecasts of euro area GDP growth rates. The main feature of this forecasting analysis is that the data vintages are constructed by taking into account the lag of synchronicity in data publication that characterizes the real-time data flow. In addition, according to the standard literature on forecasting, the forecasts are carried out in a recursive way and with every new vintage, as the model is re-estimated and the forecasts for different horizons are computed. In the empirical analysis, we show that our model would have accurately forecasted GDP over the past 20 years. The model yields significant forecasting improvements over benchmark predictions computed from models that are only based on standard autoregressive specifications.

The dynamic factor model proposed in this paper can provide a reasonable benchmark to evaluate the value added in forecasting GDP of the two most followed business surveys for the euro area, namely the PMI (released by Markit Economics) and the ESI (Economic Sentiment Indicator, published by the European Commission). Although the two set of indicators are not highly misaligned in an historical comparison, sometimes they exhibit clear divergences in trends, especially around turning points. Our results suggest that the pseudo real-time forecasts of GDP computed from the model that uses European Commission indicators outperform those from the model that uses Markit indicators.

The structure of this paper is as follows. Section 2 outlines the model, shows how to mix frequencies, states the time series dynamic properties, and describes the state space representation. Section 3 contains data description and the main empirical results. Section 4 concludes and proposes several future lines of research.

2. The model

To use data with quarterly and monthly frequencies, we follow Mariano and Murasawa (2003). Let y_t^* and y_t be the quarterly and monthly growth rates of GDP, which are assumed to be observable each moth. These authors show that

$$y_t^* = \frac{1}{3}y_t + \frac{2}{3}y_{t-1} + y_{t-2} + \frac{2}{3}y_{t-3} + \frac{1}{3}y_{t-4} \quad , \tag{1}$$

which represents the quarterly growth rate as the weighted sum of five monthly growth rates.

The dynamic properties of the model are defined flowing the lines proposed by Camacho and Perez Quiros (2010), which is an extension of the dynamic factor model suggested by Stock and Watson (1991). Let us assume that the variables used in the model admit a dynamic factor representation. In this case, the variables can be written as the sum of two stochastic components: a common component, x_t, which represents the overall business cycle conditions, and an idiosyncratic component, which refers to the particular dynamics of the series. The underlying business cycle conditions are assumed to evolve with $AR(p_1)$ dynamics

$$x_t = \rho_1 x_{t-1} + ... + \rho_{p1} x_{t-p1} = e_t \,, \tag{2}$$

where $e_t \sim i.i.d.N\left(0, \sigma_e^2\right)$.

Apart from constructing an index of the business cycle conditions, we are interested in computing accurate short-term forecasts of GDP growth rates. To compute these forecasts, we start by assuming that the evolution of the 3-month growth rates depends linearly on x_t and on their idiosyncratic dynamics, u_t^y, which evolve as an $AR(p_2)$

$$y_t = \beta_y x_t + u_t^y \,, \tag{3}$$

$$u_t^y = d_1^y u_{t-1}^y + ... + d_{p2}^y u_{t-p2}^y + \varepsilon_t^y \,, \tag{4}$$

where $\varepsilon_t^y \sim i.i.d.N\left(0, \sigma_y^2\right)$. In addition, the idiosyncratic dynamics of the k monthly indicators can be expressed in terms of autoregressive processes of p_3 orders:

$$z_t^i = \beta_i x_t + u_t^i \,, \tag{5}$$

$$u_t^i = d_1^i u_{t-q}^i + ... + d_{p3}^i u_{t-p3}^i + \varepsilon_t^i \,, \tag{6}$$

where $\varepsilon_t^i \sim i.i.d.N\left(0, \sigma_i^2\right)$. Finally, we assume that all the shocks e_t, ε_t^y, and ε_t^i, are mutually uncorrelated in cross-section and time-series dimensions.

The model can be easily represented in state space form. Let us first assume that all the variables included in the model were observed at monthly frequencies for all periods. Since GDP is used in quarterly growth rates, y_t^*, according to expressions (3)-(4) it enters into the model as

$$y_t^* = \beta_y \left(\frac{1}{3} x_t + \frac{2}{3} x_{t-1} + x_{t-2} + \frac{2}{3} x_{t-3} + \frac{1}{3} x_{t-4} \right) + \left(\frac{1}{3} u_t^y + \frac{2}{3} u_{t-1}^y + u_{t-2}^y + \frac{2}{3} u_{t-3}^y + \frac{1}{3} u_{t-4}^y \right). \tag{7}$$

To keep away the noisy signals that characterize the monthly growth rates of hard indicators, they are used in annual growth rates. Soft indicators are used in levels since by construction their levels exhibit high correlation with the annual growth rate of their reference series. Calling Z_i^* the annual growth rates of hard or the level of soft variables, the dynamics of these variables relationship are captured by

$$Z_{it}^* = \beta_i \sum_{j=0}^{11} x_{t-j} + u_t^i , \qquad (8)$$

with $i = 1, 2, \ldots, k_1$.

Finally, following the suggestions of Wheelock and Wohar (2009), financial indicators are treated as leading indicators of the current business conditions. Accordingly, we establish the relationship between the level (in the case of term spread) or annual growth rate (in the case of total credit) of the financial indicator, Z_{ft}^*, and the h-period future values of the common factor, which represents the overall state of the economy, as follows:

$$Z_{ft}^* = \beta_f \sum_{j=0}^{11} x_{t+h-j} + u_t^f . \qquad (9)$$

As it is shown in the Appendix, this model can be easily stated in state space representation and estimated by using the Kalman filter. However, we assumed that the data do not contain missing data which were clearly an unrealistic assumption since our data exhibits ragged ends and mixing frequencies problems. Fortunately, Mariano and Murasawa (2003) show that the Kalman filter can be used to estimate model's parameters and to infer unobserved components and missing observations. These authors propose replacing the missing observations with random draws ϑ_t, whose distribution cannot depend on the parameter space that characterizes the Kalman filter.[2] Hence, although this procedure leaves the matrices used in the Kalman filter conformable, the rows containing missing observations will be skipped from the updating in the recursions and the missing data are replaced by estimates. In this way, forecasting is very simple since forecasts can be viewed as missing data located at the end of the model's indicators.

3. Empirical results

3.1. Preliminary analysis of data

[2] We assume that $\vartheta_t \sim N\left(0, \sigma_\vartheta^2\right)$ for convenience but replacements by constants would also be valid.

The data set managed in this paper spans the period from January 1980 to August 2010.[3] Regarding the indicators used in the paper, we only choose those that verify three properties. First, they must exhibit high business cycle comovements with the GDP growth rate. Second, for a given quarter they should refer to data of this quarter published before the figure of GDP becomes available in the respective quarter. Third, they must be relevant in the model from both theoretical and empirical standpoints.

Shortlisted indicators at an early stage are characterized by a strong link with the GDP cycle, trying to cover the main productive sectors (industry and services) and the main components of the GDP demand side (private consumption, investment and net exports). Many of these indicators were hard data, but they are subject to significant revisions and are released with significant delays, while both confidence surveys and financial variables are not usually revised and available on a timely basis. As a result of these criteria along with above required properties, selected data were finally determined. The indicators finally included in our model and their respective release lag time are listed in Table 1 and can be classified as hard, soft and financial indicators. The hard indicators are measures of economic activity such as real GDP, unemployment rate, industrial production and export. Typically, hard indicators are published with a reporting lag between 1 and 1.5 months. Soft indicators are based on opinion surveys concerning households (consumer and service confidence) and manufacturing (industry confidence) and are released on a timely basis. Finally, among the financial indicators, we include total credit to households which is typically released with a 1-month delay and the term spread (10-year bond rate minus 3m Euribor) which is available with no reporting lags.

All the variables are seasonally adjusted, including calendar adjustments and outlier detection and correction. GDP enters in the model as its quarterly growth rate; industrial production, exports and credit enter in annual growth rates; unemployment rate and term spread enter with no transformation. Before estimating the model, the variables are standardized to have a zero mean and a variance equal to one.[4]

3.2. In-sample analysis

Selecting the indicators that must be included in a dynamic factor model is still a developing area. Boivin and Ng (2006), with US data, and Caggiano, Kapetanios, and Labhard (2011),

[3] To facilitate the analysis, following Giannone, Reichlin and Small (2008) financial data enter into the model as monthly averages since the bulk of information compiled from the indicators is monthly.

with European data, have found that selecting a smaller subset of the available large data sets, and using the factors summarizing the information in that smaller subset of data in the forecasting equation, substantially improves forecast performance. A universe of potentially available time series is still an open question in empirical studies regarding factor models.

In this paper, the selection of the euro area indicators to be used in the dynamic factor model, from those previously considered, follows the recommendations suggested by Camacho and Perez Quiros (2010).[5] Following Stock and Watson (1991), we start with a model that includes measures of real economic activity such as GDP, unemployment, industrial production, and export. The estimated factor loadings, which measure the correlation between the economic indicators and the common factor, appear in the row labeled as M1 in Table 2. GDP has a positive factor loading. Therefore, industrial production and exports, which also exhibit positive factor loadings, are procyclical. In contrast, unemployment, which is clearly a countercyclical indicator, has a negative factor loading. In all cases, the factor loadings are statistically significant. Finally, the percentage of the variance of GDP explained by the model is about 64%.

The delay in the publication of many of these four indicators makes it difficult to assess the performance of economic activity in real time. To overcome this problem, soft indicators, which are available with very little delay, are further added to the estimation whenever the increase in the size of the data set raises the percentage of the variance of GDP explained by the common factor, but only when the variables to be added have statistically significant loading factors. Otherwise, the information provided by the potential indicators is assumed to be mainly idiosyncratic and it is not included in the model.

Following this principle, we extend the initial set of hard indicators from the Economic Sentiment Indicators (ESI) for the euro area as a whole which have the advantage of being available with almost no publication delay.[6] The surveys used to construct these indicators are conducted by the European Commission Directorate General for Economic and Financial Affairs (DG ECFIN) and are used to release confidence indicators linked to different sectors. As a first attempt, we enlarge model M1 with the confidence indicators related to industrial, services and consumer sectors since they cover the major sectors of the euro area. The empirical estimates, labeled as M2 in Table 2, reveal that the loading factors

[4] Therefore, final forecasts are computed by multiplying initial forecasts of the model by the sample standard deviation, and then adding the sample mean.
[5] All the dynamic factor models use $p1$=6 and $p2$=$p3$=2.
[6] Comparative assessments between ESI and PMI indicators appear in the forecasting analysis.

are positive as in the case of GDP and statistically significant. In addition, the percentage of the variance of GDP explained by the model becomes stable around 65%.

Although we tried with other soft indicators such as ESI construction, ESI retail sales, both for the euro area as a whole, and also with German indicators such as the IFO business climate index and ZEW indicator of economic sentiment. The factor loadings estimated for these indicators were always non significant and in many cases with counterintuitive sign. We also obtained the same conclusion when we tried to use national indicators such as the composite ESI for Germany, France Italy and Spain.[7]

The final enlargement of the model is conducted by including two financial indicators, total credit to households and term spread. In this context it is worth quoting the recent work by Wheelock and Wohar (2009), who find that the contemporaneous correlation between GDP growth and the slope of the yield curve is not statistically different from zero for the US, the UK and Germany, whereas the correlation with the slope lagged from one to six quarters are uniformly positive and statistically significant. According to these results, financial indicators are assumed to lead the business cycle dynamics in *h* periods. To select the number of leads, we follow Camacho and Domenech (2012) and compute the log likelihood associated with lead times that go from one quarter to one and a half years. Using this procedure, we find that the maximum of the likelihood function is achieved when the financial indicators are allowed to lead the common factor by nine months. The estimated loading factors of the model that includes financial indicators, which are displayed in the last row of Table 2 and labeled as M3, show that the highest credit and spread and the greater expected GDP growth. The final percentage of the variance of GDP explained by the model rises to about 67%.

Our model is based on the notion that co-movements among the macroeconomic variables have a common element, the common factor that moves in accordance with the euro area business cycle dynamics. To check whether the estimated factor coincides with the euro area business cycle, Figure 1 plots the estimated common factor and the Eurocoin published by the CEPR, which is probably the leading coincident indicator of the euro area business cycle.[8] The similarities between their business cycle dynamics are striking, suggesting that they track the same business cycle.

[7] Results are omitted to save space but they are available from the authors upon request.

[8] To facilitate comparisons, the Eurocoin has been transformed to exhibit the mean and the variance of the common factor.

To further examine the business cycle information that can be extracted from the common factor, let us assume that there is a regime switch in the index itself. For this purpose, we assume that the switching mechanism of the common factor at time t, x_t, is controlled by an unobservable state variable, s_t, that is allowed to follow a first-order Markov chain. Following Hamilton (1989), a simple switching model may be specified as:

$$x_t = c_{s_t} + \sum_{j=1}^{p} \alpha_j x_{t-j} + \varepsilon_t \quad , \tag{10}$$

where $\varepsilon_t \sim i.i.d.N(0,\sigma^2)$. The nonlinear behavior of the time series is governed by c_{s_t}, which is allowed to change within each of the two distinct regimes $s_t = 0$ and $s_t = 1$. The Markov-switching assumption implies that the transition probabilities are independent of the information set at t-1, χ_{t-1}, and of the business cycle states prior to t-1. Accordingly, the probabilities of staying in each state are

$$p(s_t = i / s_{t-1} = j, s_{t-2} = h, ..., \chi_{t-1}) = p(s_t = i / s_{t-1} = j) = p_{ij} \, . \tag{17}$$

Taking the maximum likelihood estimates of parameters, reported in Table 3, in the regime represented by $s_t = 0$, the intercept is positive and statistically significant while in the regime represented by $s_t = 1$, it is negative and statistically significant. Hence, we can associate the first regime with expansions and the second regime with recessions. According to the related literature, expansions are more persistent than downturns (estimated p_{00} and p_{11} of about 0.98 and 0.76, respectively). These estimates are in line with the well-known fact that expansions are longer than contractions, on average. Finally, Figure 2 displays the estimated smoothed probabilities of recessions and shaded areas that refer to the periods classified as recessions by the CEPR euro area Business Cycle Dating Committee (BCDC). The figure illustrates the great ability of the model to capture the European business cycle and validates the interpretation of state $s_t = 1$ as a recession and the probabilities plotted in this chart as probabilities of being in recession.[9]

A final check of the ability of the common factor to capture the euro area businesses cycle dynamics is the recession diagnostic evaluation procedure recently proposed by Berge and Jorda (2011). Under the fact that the common factor rises in expansions and fails in recessions, the classification procedure calculates a trigger level of the index beyond which a recession can be considered as highly probable. In our case, the threshold value that

[9] It is worth mentioning that the dating committee considered that in 2001 the euro area experienced a prolonged pause in the growth of economic activity rather than a recession.

maximizes the net benefits of true positives (when recessions are correctly identified) minus the costs of false positives (when recessions are called in expansionary periods) under the assumptions that expansions and recessions have symmetric importance is -8.82. Hence, values above this threshold would be classified as expansions, according to the criteria of the euro area BCDC, and values below it as recessions.[10] The area under the receiver operating characteristic curve, which measures business cycle classification skills of the common factor is 0.82 (standard deviation of 0.01) which is clearly above the coin-toss classifier value of 0.5 (Figure 2).

3.3. Simulated real-time analysis

Among many others, Stark and Croushore (2002) suggest that the analysis of in-sample forecasting performance of competitive models is questionable since the results can be deceptively lower when using real-time vintages. This happens because the in-sample analysis misses three aspects of real-time forecasting: (i) the recursive estimation of the model parameters as h-step-ahead forecasts are computed (ii) the real-time data flow, i.e., the fact that the indicators are released by the statistical agencies in charge of publishing new releases at different points in time; and (iii) the real-time data revisions that affects some economic indicators whose latest releases includes the new data and revisions of previous releases.

However, Croushore and Stark (2001) have pointed out that although developing real-time data sets is conceptually simple, producing real-time vintages is, as in our case, sometimes unfeasible since the historical records of some time series have been lost. In the context of dynamic factor models, an interesting alternative to the real-time forecasting analysis is the pseudo real-time forecasting exercise developed in this paper. This forecasting analysis takes into account the recursive estimation of the model and the real-time data flow (and hence the publication lags) without considering data revision, due to data availability constraints.

Accordingly, this proposal is based on trying to mimic as close as possible the real time analysis that would have been performed by a potential user of our dynamic factor models when forecasting at each period of time, on the basis of different vintages of data sets which are constructed from the latest available set of data. The experiment considers that the releases of each vintage contain missing data at the end of the sample reflecting the calendar of data releases following the reporting lags outlined in Table 1. Because the data is released

[10] Notice that calling recessions from a zero threshold is not optimal.

in blocks and the releases follow a relatively stable calendar, we can reproduce each 15 days the typical end of the sample unbalanced panel faced by the forecaster due to the lack of synchronization of the data releases. Hence, the experiment is labeled as "pseudo" because the vintages are not obtained in pure real time but from the latest available data set.

Since we wanted to forecast GDP growth for almost twenty years from 1990.1 to 2010.2, the first data vintage of this experiment refers to data up to 1989.01 as it would be known on June 15, 1989.[11] In each forecasting day, nine-month blocks of forecasts are computed from the model, which incorporate backcasts (forecasting last quarter's GDP before its official release), nowcasts (predicting current quarter GDP) and short-term forecasts (predicting next quarter's GDP). The first day on which the model produces forecasts of 1990.01 is June 15, 1989 and the vintages are then updated on the first day and on the fifteenth day of each month up to May 15, 2010, leading to 492 different vintages.

The nine-month blocks of forecasts implies that each GDP figure is forecasted 18 times. For example, the first forecast of the figure of GDP corresponding 2009.3 computed on February 15th, the date of publication of GDP corresponding to 2008.4 due to the nine-month blocks of forecasts developed in the analysis. Hence, the forecasts are computed each fifteen days until November 1st. To analyze the relation between the real-time forecasts and the information available at each forecasting period, Figure 3 plots the root mean-squared forecast errors (RMSE) for the 18-period forecasts of each GDP figure. The figure clearly shows how the flow of data releases during the forecasting period helps to produce more accurate forecasts. The first improvement in GDP forecasts accuracy occurs around three months after the first estimate. At that time a new GDP figure was already available and some soft and hard indicators lead to a more reliable forecast. The second significant forecast improvement befalls when confidence data are available for the quarter to be forecasted. The third improvement takes place when quantitative data for that quarter are also partially available as well as an additional figure for GDP is published. To a lesser extent, the forecast accuracy continues to improve as updated quantitative data are included.

This forecasting exercise allows us to asses the relative importance of forecasting from updated information sets. For this purpose, plots of actual data and pseudo real-time predictions can be found in Figure 4. This figure shows the simulated real-time backcasts (straight lines) of euro area GDP, which are updated each fifteen days, as well as the corresponding final quarterly data (dashed lines), which are equally distributed among the

[11] According to the nine-month blocks of forecasts computed from the model, the first day on which the model produces forecasts of 1990.01 is June 15, 1989.

respective days of the quarter to facilitate comparisons. The figure shows that the series of actual releases and pseudo real-time forecasts possess a high degree of conformity. Forecasts follow sequential patterns that track the business cycle marked by the evolution of GDP releases.[12]

Table 4 shows the mean-squared forecast errors (MSE), which are the average of the deviations of the predictions from the final releases of GDP available in the data set. Results for backcasts, nowcasts and forecasts appear in the second, third and fourth columns of the table, respectively. In addition to the factor model described in Section 2 (labeled as MICA), two benchmark models are included in the forecast evaluation. The former is an autoregressive model of order two (AR) which is estimated in real-time producing iterative forecasts, and the latter is a random walk (RW) model whose forecasts are equal to the average latest available real-time observations. Finally, the pseudo real-time forecasting exercise constitutes a natural framework to evaluate the value added in forecasting GDP of the two most watched business surveys for the euro area, namely the PMI (released by Markit Economics) and the ESI (Economic Sentiment Indicator, published by the European Commission). For this purpose, the forecasts have also been computed from a dynamic factor model that substitutes the ESI industry and service confidence indicators by their PMI counterparts and the ESI consumer confident indicator by retail sales.[13] This model is labelled in the table as MICA2.

The MSE leads to a ranking of the competing models according to their forecasting performance. However, it is advisable to test whether the forecasts made with the dynamic factor model are significantly superior to the others models' forecasts. To analyze whether empirical loss differences between two or more competing models are statistically significant, the last three rows of the table shows the pairwise test introduced by Diebold and Mariano (DM, 1995).

The immediate conclusion obtained when comparing the forecasts is that the gains in using the dynamic factor model in forecasting GDP with respect to ARIMA models depend on the forecast horizon. In the backcasting exercise, the differences between the MSE results using the factor model and the benchmark models are noticeable. The relative MSE of the dynamic factor model versus RW and AR are 0.374 to 0.466 and, according to the *p*-values of the DM test, the differences are statistically significant. Although the gains in nowcasting

[12] Although the graphs corresponding to nowcasts and forecasts are omitted to save space, they also track the GDP dynamics well but with some delays since they use poorer information sets to compute predictions.
[13] Notice that Markit does not publish consumer confident indicator.

GDP with the dynamic factor model are still considerable (relative MSE of about 0.80), the *p*-values of the DM test indicate now that the differences are not significant. However, the dynamic factor model does not seem to exhibit reductions in MSE when forecasting GDP with respect to the MSE achieved by ARIMA models. Notably, the reductions in MSE for the dynamic factor model that uses the ESI confident indicators compared to the model that uses the PMI confident indicators are large (relative MSE about 0.70) regarding the forecast horizon. The last row of Table 4 shows that the differences are statistically significant.

To analyze the stability of the forecasting performance over time, the table also includes the MSE within recessions and within expansions, which are computed from the periods that have been identified by CEPR as recessions and expansions. The figures of the table show that the forecasting accuracy of the models varies considerably over the business cycle. In recessions, although there is a marked deterioration for all models, the relative loss in forecasting accuracy from the benchmark ARIMA models is magnified with respect to that in the expansionary periods. The intuition is that the evolution of GDP in expansions is quite flat around its historical average. Therefore, the relative reductions obtained from the dynamic factor model diminish considerably in expansions.

To analyze further the potential instabilities as the consequence of booms or recessions in forecasting accuracy, we use the fluctuation tests proposed by Giacomini and Rossi (2010). For this purpose, Figure 5 plots the centered local loss differences of the Modified Diebold-Mariano (MDM) test suggested by Harvey, Leybourne, and Newbold (1997).[14] In each period, the figure plots the MDM which is computed over a rolling subsample of 40-month window. Negative (positive) values of such differences indicate that the dynamic factor model that uses ESI indicators produces better (worse) forecasts than the competitors. The figure also includes the two-sided critical values of equal forecasting accuracy and shaded areas that refer to the recessions periods suggested by euro area business cycle dating committee.

According to Figure 5, the relative forecast performance of the dynamic factor model that uses the ESI confident indicators with respect to ARIMA models and the dynamic factor model that uses the PMI indicators change over time. In particular, the MICA model delivers accurate forecast especially in slowdowns and periods around business cycle turning points. This result suggests that the forecast accuracy gains of the dynamic factor model can be down weighted if they are analyzed over a fixed evaluation period. Simple autoregressive models might be difficult to beat in rather tranquil times while the strength of the dynamic factor

[14] The modified DM test is more appropriate in the context of short sample forecasting scenarios.

models based on early available indicators is to contain early information on booms and recessions.

4. Conclusions

This paper proposes an extension of the Stock and Watson (1991) single-index dynamic factor model and evaluates it for forecasting exercises of the euro area quarterly GDP growth. The model has the advantage of combining information from real and financial indicators with different frequencies, short samples and publication lags. Using the Kalman filter, the model computes estimates of the unobserved common coincident component and of any missing values in the different series used to estimate the model.

Our results indicate three interesting features. First, we find that the common factor reflects the behavior of the euro area GDP growth during expansions and contractions properly. Second, we show that financial indicators such as the slope of the yield curve and the growth rate of real credit are useful for forecasting output growth especially when assuming that some financial variables lead the common factor. Finally, we provide a simulated real-time exercise that is designed to replicate the data availability situation that would be faced in a true real-time application of the model. We show that the model is a valid tool to be used for short-term analysis.

The analysis in this paper highlights some lines for future research. First, although the model presented in this paper provides timely estimates of the state of real activity, it does not provide measures of the economic activity at frequencies higher than monthly. This is still a developing area but several ongoing studies such as Aruoba, Diebold and Scotti (2009) are exploring this possibility. Second, although we examine the forecasting accuracy of the model by using a pseudo real-time exercise that accounts for recursive estimations and the typical delays observed in data publications, it uses final data vintages and, hence, ignores statistical revisions to earlier data releases. Although the actual data vintages that would have been used by real-time forecasters are hard to be obtained, we believe that allowing for such revisions is an interesting exercise for further assessing forecasting accuracy of our model in real-time.

Appendix

Without loss of generalization, we assume that our model contains only GDP, one non-financial indicator and one financial indicator, which are collected in the vector $Y_t = \left(y_t^*, Z_{it}^*, Z_{ft}^*\right)'$. For simplicity sake, we also assume that $p_1 = p_2 = p_3 = 1$, and that the lead for the financial indicator is $h = 1$. In this case, the observation equation, $Y_t = Z\alpha_t$, is

$$
\begin{pmatrix} y_t^* \\ Z_{it}^* \\ Z_{ft}^* \end{pmatrix} = \begin{pmatrix} 0 & \frac{1\beta_y}{3} & \frac{2\beta_y}{3} & \beta_y & \frac{2\beta_y}{3} & \frac{1\beta_y}{3} & 0 & \cdots & 0 & \frac{1}{3} & \frac{2}{3} & 1 & \frac{2}{3} & \frac{1}{3} & 0 & 0 \\ 0 & \beta_i & \beta_i & \cdots & & & & \cdots & \beta_i & 0 & \cdots & & \cdots & 0 & 1 & 0 \\ \beta_f & \beta_f & & & & & & \beta_f & 0 & 0 & \cdots & & & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} x_{t+1} \\ x_t \\ \vdots \\ x_{t-11} \\ u_t^y \\ \vdots \\ u_{t-4}^y \\ u_t^i \\ u_t^f \end{pmatrix}. \quad \text{(A1)}
$$

It is worth noting that the model assumes contemporaneous correlation between non-financial indicators and the state of the economy, whereas for financial variables, the correlation is imposed between current values of the indicators and future values of the common factor.

The transition equation, $\alpha_t = T\alpha_{t-1} + \eta_t$, is

$$
\begin{pmatrix} x_{t+1} \\ x_t \\ \vdots \\ x_{t-11} \\ u_t^y \\ \vdots \\ u_{t-5}^y \\ u_t^i \\ u_t^f \end{pmatrix} = \begin{pmatrix} \rho_1 & \cdots & 0 & 0 & 0 & & \cdots & & & 0 \\ 1 & & & 0 & & & \cdots & & & 0 \\ \vdots & \ddots & & \vdots & & & \cdots & & & \vdots \\ 0 & \cdots & 1 & 0 & & & \cdots & & & 0 \\ 0 & \cdots & & 0 & d_1^y & & 0 & 0 & & 0 \\ \vdots & \cdots & & \cdots & & \ddots & & & & \vdots \\ 0 & \cdots & & & & & 1 & 0 & & 0 \\ 0 & \cdots & & & & & 0 & 0 & d_1^i & 0 \\ 0 & \cdots & & & & & 0 & 0 & 0 & d_1^f \end{pmatrix} \begin{pmatrix} x_t \\ x_{t-1} \\ \vdots \\ x_{t-12} \\ u_{t-1}^y \\ \vdots \\ u_{t-6}^y \\ u_{t-1}^i \\ u_{t-1}^f \end{pmatrix} + \begin{pmatrix} e_{t+1} \\ e_t \\ \vdots \\ e_{t-11} \\ \varepsilon_{t-1}^y \\ \vdots \\ \varepsilon_{t-6}^y \\ \varepsilon_t^i \\ \varepsilon_t^f \end{pmatrix}, \quad \text{(A2)}
$$

where $\eta_t \sim iN(0,Q)$ and $Q = diag\left(\sigma_e^2, 0, \ldots, 0, \sigma_y^2, 0 \ldots 0, \sigma_i^2, \sigma_f^2\right)$.

References

Aruoba, B., Diebold, F., and Scotti, C. 2009. Real-time measurement of business conditions. *Journal of Business and Economic Statistics*, in press.

Berge, T., and Jorda, O. 2011. The classification of economic activity into expansions and recessions. *American Economic Journal*; *Macroeconomics* 3: 246-277.

Camacho, M., and Domenech, R. 2012. MICA-BBVA: A factor model of economic and financial indicators for short-term GDP forecasting. *SERIEs*, forthcoming.

Camacho, M., and Perez Quiros, G. 2010. Introducing the Euro-STING: Short Term INdicator of euro area Growth. *Journal of Applied Econometrics* 25: 663-694.

Diebold F., and Mariano, R. 1995. Comparing predictive accuracy. *Journal of Business and Economic Statistics* 13: 253-263.

Giacomini, R., and Rossi, B. 2010. Forecast comparisons in unstable environments. Journal of Applied Econometrics 25: 595-620.

Giannone, D., Reichlin, L., and Small, D. 2008. Nowcasting: The real-time informational content of macroeconomic data. *Journal of Monetary Economics* 55: 665-676.

Harvey, D., Leybourne, S., and Newbold, P. 1997. Testing equality of prediction mean squared errors. *International Journal of Forecasting* 13: 287-291.

Mariano, R., and Murasawa, Y. 2003. A new coincident index of business cycles based on monthly and quarterly series. *Journal of Applied Econometrics* 18: 427-443.

Stock, J., and Watson, M. 1991. A probability model of the coincident economic indicators. In Kajal Lahiri and Geoffrey Moore editors, *Leading economic indicators, new approaches and forecasting records*. Cambridge University Press, Cambridge.

Wheelock D., and Wohar, M. 2009. Can the term spread predict output growth and recessions? A survey of the literature. *Federal Reserve Bank of St. Louis Review* 91: 419-440.

Table 1: Final variables included in the model

	Series	Effective Sample	Source	Publication delay	Data transformation
1	Real GDP (GDP)	80.1-10.2	Eurostat	1.5 months	SA, QGR
2	Unemployment rate (UR)	04.01-10.06	Eurostat	2 months	SA, L
3	Industrial production (IP)	91.01-10.06	Eurostat	1.5 months	SA, AGR
4	Exports (Exp)	90.01-10.06	European Commission	1.5 months	SA, AGR
5	ESI Industry (ESII)	85.01-10.08	European Commission	0 months	SA, L
6	ESI Consumption (ESIC)	85.01-10.08	European Commission	0 months	SA, L
7	ESI Services (ESIS)	95.04-10.08	European Commission	0 months	SA, L
8	Credit to households (LHH)	04.01-10.07	European Central Bank	2 months	SA, AGR
9	Term spread (Beta)	94.01-10.08	BBVA Research	0 months	SA, L

Notes: SA means seasonally adjusted. QGR, AGR and L mean quarterly growth rates, annual growth rates and levels.

Table 2: Loading factors

Model	GDP	UR	IPI	Exp	ESII	ESIC	ESIS	LHH	Beta	% var
M1	0.20 (0.03)	-0.002 0.001	0.17 (0.01)	0.11 (0.01)	---	---	---	---	---	64.43%
M2	0.10 (0.04)	-0.005 (0.003)	0.04 (0.01)	0.03 (0.01)	0.03 (0.01)	0.02 (0.01)	0.02 (0.01)	---	---	64.91%
M3	0.10 (0.04)	-0.005 (0.003)	0.04 (0.01)	0.03 (0.01)	0.03 (0.01)	0.02 (0.01)	0.02 (0.01)	0.011 (0.005)	0.02 (0.01)	67.40%

Notes. Factor loadings (*t*-ratios are in parentheses) measure the correlation between the common factor and each of the indicators appearing in columns. See Table 1 for a description of the indicators.

Table 3. Markov-switching estimates

c_0	c_1	α_1	σ^2	p_{00}	p_{11}
0.34 (0.11)	-5.30 (0.61)	0.88 (0.02)	3.51 (0.28)	0.98 (0.01)	0.76 (0.10)

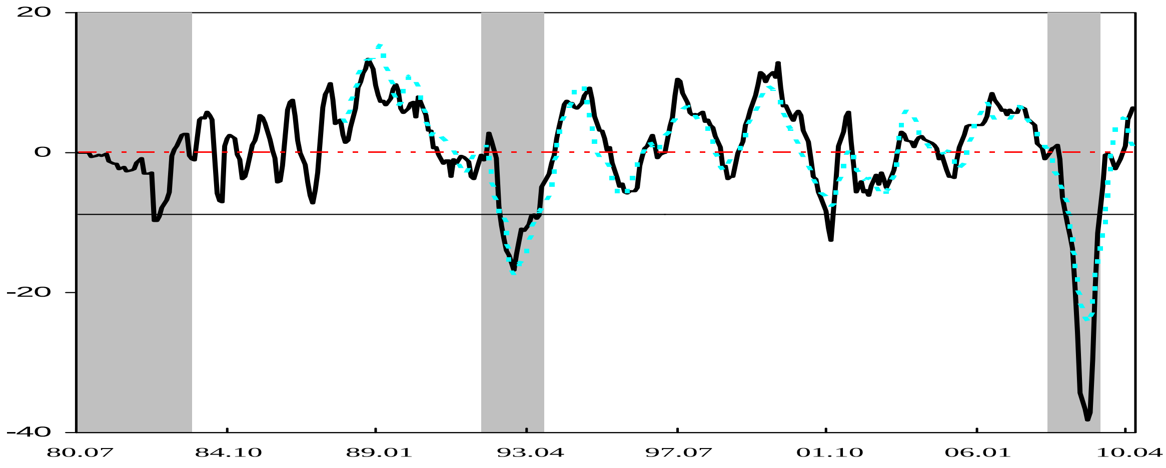
Notes. The estimated model is $x_t = c_{s_t} + \alpha_1 x_{t-1} + \varepsilon_t$, where x_t is the common factor and $\varepsilon_t \sim iidN(0,\sigma)$, and $p(s_t = i/s_{t-1} = j) = p_{ij}$.

Table 4: Predictive accuracy

	Backcasts	Nowcasts	Forecasts
Mean Squared Errors			
MICA	0.138 E: 0.080 R: 0.449	0.298 E: 0.108 R: 0.916	0.403 E: 0.223 R: 1.358
RW	0.370 E: 0.111 R: 1.746	0.361 E: 0.107 R: 1.707	0.377 E: 0.113 R: 1.779
MICA/RW	0.374	0.824	1.070
AR	0.297 E: 0.098 R:1.350	0.369 E: 0.127 R: 1.651	0.377 E: 0.114 R: 1.769
MICA/AR	0.466	0.806	1.071
MICA2	0.151 E: 0.128 R: 0.374	0.348 E: 0.178 R: 1.062	0.4647 E: 0.177 R: 1.962
MICA/MICA2	0.914	0.856	0.867
MS	0.212 E: 0.119 R: 0.708	0.321 E: 0.163 R: 1.165	0.403 E: 0.168 R: 1.012
MICA/MS	0.652	0.928	1.333
TAR	0.297 E: 0.085 R: 1.420	0.322 E: 0.189 R: 0.981	0.323 E: 0.157 R: 1.205
MICA/TAR	0.466	0.923	1.205
STING	0.254 E: 0.176 R: 0.667	0.298 E: 0.252 R: 0.750	0.385 E: 0.194 R: 1.400
MICA/STING	0.544	0.898	1.047
Equal predictive accuracy tests			
MICA vs RW	0.004	0.356	0.629
MICA vs AR	0.001	0.259	0.614
MICA vs MICA2	0.002	0.006	0.023
MICA vs MS	0.015	0.693	0.237
MICA vs TAR	0.005	0.637	0.128
MICA vs STING	0.022	0.516	0.682

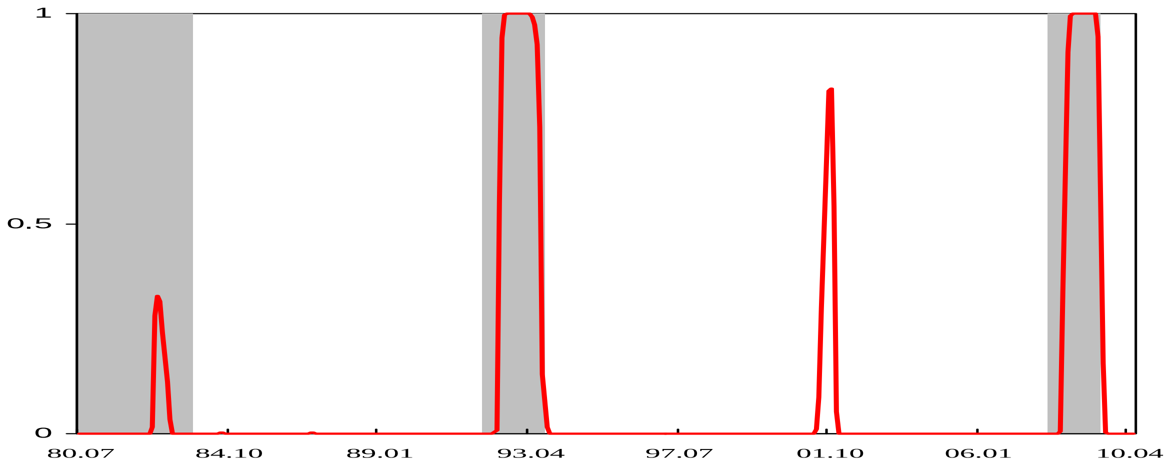
Notes. The forecasting sample is 1990.1-2010.1, which implies comparisons over 492 forecasts. Entries in rows one to thirteen are Mean Squared Errors (MSE) of dynamic factor model with ESI indicators (MICA), Random Walk (RW), autoregressive of order two (AR), dynamic factor model with PMI indicators (MICA2), Markov-switching of order two (MS), TAR of order two (TAR), and the Euro-STING model (STING), along with the relative MSEs over that of ESI. R and E refer to recessions and expansions periods according to CEPR. The last six rows show the *p*-values of the Diebold-Mariano (DM) test of equal forecast accuracy.

Figure 1. Common factor



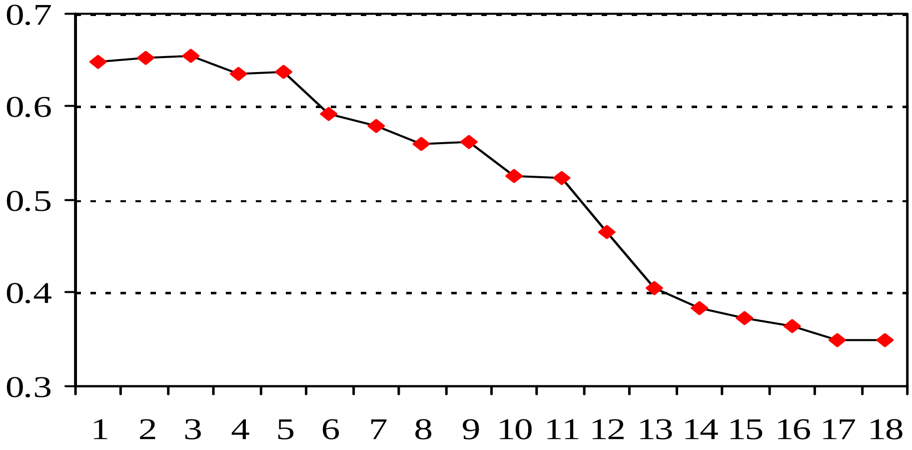
Notes. Straight line refers to the common dynamic factor (1980.07-2010.06). Dotted line refers to euro-coin (1988.01-2010.06). Straight horizontal line marks the optimal threshold of -8.82 so that values below it would be classified as recessions. The New-Eurocoin has been transformed to exhibit the mean and variance of the common factor. Shaded areas are the recessions suggested by euro area business cycle dating committee.

Figure 2. Smoothed recession probabilities



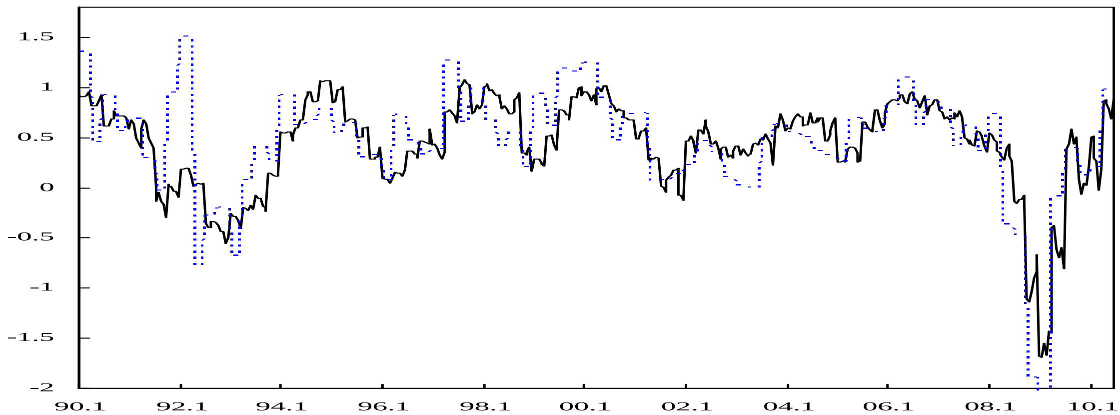
Notes. Shaded areas are the recessions suggested by euro area business cycle dating committee.

Figure 3. Evolution of forecasting accuracy



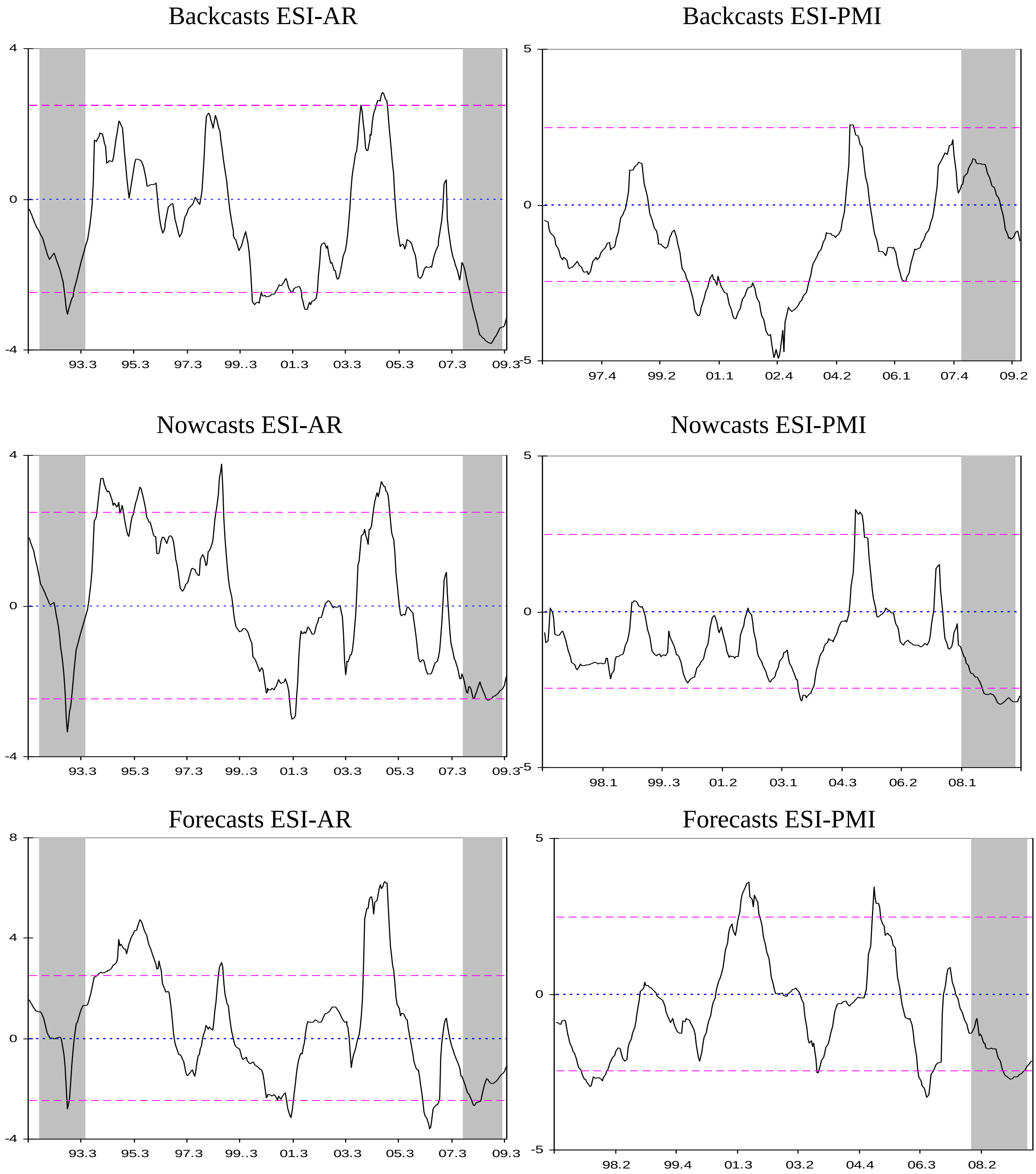
Notes. The figure plots the sample (root) average of the standard errors associated with the 18 times that each GDP figure is forecasted in real-time.

Figure 4. Pseudo real-time backcasts and actual realizations



Notes. The dotted line refers to actual realizations of GDP growth while the straight line refers to pseudo real-time backcasts computed from the dynamic factor model.

Figure 5. Rolling modified Diebold-Mariano test statistic



Notes. Shaded areas are the recessions suggested by euro area business cycle dating committee.