Introducing the Euro-STING: Short-Term INdicator of Euro Area Growth [*]

Maximo Camacho

Universidad de Murcia

mcamacho@um.es

Gabriel Perez-Quiros

Banco de España and CEPR

gabriel.perez@bde.es

Abstract

We set out a model to compute short-term forecasts of the euro area GDP growth in real-time. To allow for forecast evaluation, we construct a real-time data set that changes for each vintage date and includes the exact information that was available at the time of each forecast. With this data set, we show that our simple factor model algorithm, which uses a clear, easy-to-replicate methodology, is able to forecast the euro area GDP growth as well as professional forecasters who can combine the best forecasting tools with the possibility of incorporating their own judgement. In this context, we provide examples showing how data revisions and data availability affect point forecasts and forecast uncertainty.

Keywords: Business Cycles, Output Growth, Time Series.

JEL Classification: E32, C22, E27

1 Introduction

Early assessments of ongoing developments in economic activity are of crucial interest to economic agents for successful decision-making. In the euro area, the lack of timely information associated with the publication of macroeconomic variables, the presence of missing values in the historical time series and the short length of the euro area aggregates make the day-to-day monitoring of economic activity especially problematic. In particular, although the generally accepted reference series to describe short-term economic developments is the euro area GDP growth rate, the "final" estimates of GDP growth (called *second* release) is published with a lag of about 14 weeks after the end of the respective quarter. Given this publication delay, forecasters from various relevant institutions try to anticipate the evolution of GDP growth by publishing their highly influential forecasts. Among these, the most significant are the European Commission's macroeconomic forecasts, the euro area GDP growth projection of DG ECFIN, the IFO-INSEE-INSAE economic forecast and the projections of the OECD Economic Outlook. Usually, the forecasts of these institutions are based on indicators that are available on a more timely basis and exhibit similar economic fluctuations to the reference series. The natural indicators are the two early announcements of the second release, which are called *flash* and *first* releases, respectively. Other candidates are monthly indicators, which are based on either economic activity data (hard indicators) or surveys (soft indicators) since they exhibit a much shorter publishing lag than the second releases.

However, even though these institutions use state-of-the-art forecasting methods, they have the possibility of partly basing their forecasts on judgements. Consequently, their forecasts cannot be easily replicated and their forecast failures are difficult to interpret. We seek to avoid this problem by using a simple algorithm which, while doing the job of forecasting euro area GDP growth at least as well as the professional forecasters do in real time, has the advantages of forecasting from a specific model which can be evaluated in terms of transparency, replicability and decomposition of the informational content of all the macroeconomic data released.

For this purpose, we consider a short-term forecasting model which modifies somewhat

2

the Stock and Watson (1991) *strict* dynamic factor model to allow for the particular data problems of real-time forecasting. Following the approximate Kalman filter suggested by Mariano and Murasawa (2003), the model is able to handle indicators which are available at different frequencies. As in Giannone, Reichlin and Small (2008), the gaps that characterize the ragged edges behind the asynchronous data publication are also filled in by using the Kalman filter. Finally, as in Evans (2005), the Eurostat data revision procedure when publishing GDP growth rates is explicitly modeled.

The paper is closely related to Banbura and Runstler (2007), Angelini et al. (2008) and Barhoumi et al. (2008), who use the *approximate* dynamic factor model proposed by Giannone et al. (2008) to compute euro area GDP forecasts which are continuously updated as well. As in these proposals, we diverge from the euro area univariate bridge equations employed by Runstler and Sedillot (2003) and Diron (2006) and from those which try to measure high-frequency objects (as real-time activity) on a daily or hourly basis, such as Aruoba, Diebold and Scotti (2009). However, we propose several contributions which clearly differentiate our approach from their proposal.

First, in contrast to the euro area applications of large-scale factor models but in line with Frale et al. (2008), our model is a small-scale factor model.[1] We are aware that users of approximate factor models have extensively stressed that strict factor models rely on the tight assumption that the idiosyncratic noises are cross-sectionally orthogonal. However, large-scale factor models are not exempt from theoretical assumptions that may not hold in empirical applications. Large-scale models estimate factors that are consistent when the number of variables and observations tend to infinity, under the assumptions that the idiosyncratic components are weakly correlated (in time series and cross-sections) and that the variability of the common component is not too small. However, empirical warnings may appear since the number of time series usually employed in euro area applications is finite and small (for example, Angelini et al, 2008, use 85 time series out of the several thousands of time series available), contrary to what theory requires. In addition, as documented by Boivin and Ng (2006), empirical warnings may also appear because

---

[1] According to the classification of Cubbada and Hecq (2008), technically our model should be considered as a medium-scale model. However, during the text, we will refer to our model as a small-scale model.

economic series are drawn from a small number of categories. If this is the case, adding more variables to the core of representatives of different categories may add mainly noise and cross-correlation of the idiosyncratic shocks, generating idiosyncratic correlations that might be larger than those warranted by theory. Additionally, Bai and Ng (2008) show that, when the variables are carefully selected, having zero loads for some of the variables of the large-scale models (i.e. reducing the scale of the model) may improve the forecasting performance of the models, even when the weak cross-correlation holds. However, in the euro area applications of factor models, the empirical reliability of the theoretical assumptions is not usually analyzed. Finally, Inklaar, Jacobs and Romp (2004) have recently suggested that a reliable indicator for the euro area can be constructed from a limited number of series that are selected using economic logic.

One of the advantages of using small-scale models is that, given their small dimension, it is relatively easier to check the empirical implications of the violation of the theoretical assumptions. For this purpose, we propose a method which is a modified version of Engle's (2007) McGyver method and consists of estimating the covariance matrix of the idiosyncratic shocks by blocks. Using this method, we find that, although some cross-covariance is present in the estimated idiosyncratic components, it leads to negligible impacts in the empirical estimation of factors.

Second, in contrast to the previously cited factor models which perform out-of-sample euro area forecast evaluations against basic benchmarks, we compare our forecasts not only against those benchmarks, but also against the forecasts made by the most influential professional forecasters. In addition, we include in this comparison the key ingredient of forecasting analysis in true real time. Typically, real-time forecasting specialists (see for example Stark and Croushore, 2002) stress that the evaluation of forecast errors from latest-available data is questionable, suggesting that comparisons between the forecasts generated from new models, competitors and benchmark forecasts should be based on real-time data rather than out-of-sample simulations. For this purpose, we construct a real-time data set which include the data vintages available at the time of each forecasting day of the past four years. This task was relatively easy thanks to the limited number of predictors used in the model compared with models with larger sets of indicators. It is worth noting

4

that, following the spirit of computing model-based forecasts by using exactly the same amount of information that was available the day on which the competitors published their forecast, we are precluded from using the Euro Area Business Cycle Network real-time data set. The reason is that these data are collected by monthly vintages (not updated daily), resulting in a significant delay of about half a year. Using our database, we find that our model outperforms all of the most recognized euro area GDP growth forecasts in terms of mean squared errors for different forecasting horizons. In addition, although we acknowledge some degree of uncertainty since Diebold-Mariano-based tests cannot reject in many cases the null of equal forecasts due to the short sample, we obtain that the null that the forecasts from our model encompass the forecasts of the competitors cannot be rejected in any case. Finally, using this real-time data set, we provide examples that show how data revisions and data availability affect point forecasts and forecast uncertainty.

Third, our model also contributes in explicitly modelling the data revision procedure followed by the Eurostat GDP data releases. For this purpose, we follow the US application of Evans (2005) and propose that preliminary announcements in the US are noisy signals of revised data. With respect to the euro area, the idea that the GDP publishing procedure can be modeled as noise instead of news has also been used in the theoretical proposal of Coenen, Levin, Wieland (2005). In addition, we show empirical evidence favouring the noisy version of the euro area GDP revision procedure. In particular, we find that revisions are much more correlated with preliminary announcements than with final estimates, that the noise-to-signal ratio for revisions is sizeable, that the forecast efficiency test fails to reject the noise hypothesis and that the revisions from flash to second are less volatile than the revisions from first to second.

Fourth, we also innovate in suggesting a statistical method for data selection which is based on the idea that the object of interest is ultimately the forecast of euro area growth rates. Bearing in mind the previous discussion about our model's specificities, we start the data selection by enlarging the model of Stock and Watson (1991) to include GDP growth and its early releases (first and flash). In addition, given that the model should provide early assessments of economic developments, we also include the most promptly available set of soft indicators. With this reasonable set of indicators, our method to

further enlarge the model is given by the criteria that we include an additional indicator only when it increases the percentage of variance of GDP growth explained by the common factor. The intuition for this method involves screening out those additional indicators that capture idiosyncratic dynamics that do not lead to a better fit for GDP growth through the common component.

In the spirit of forecasting economic time series using targeted predictors as in Bai and Ng (2008), we use our method to examine the need to use many predictors in forecasting euro area GDP growth. For this purpose, we consider our small-scale model which is estimated with data sets of different dimensions. First, we enlarge the model by including as additional predictors each of the 85 variables proposed by Angelini et al. (2008), which were not already included in our model, one by one. Second, we enlarge the model by including the first five factors computed for those 85 variables by using the methods of Stock and Watson (2002) and Giannone et al. (2008). In none of these cases does the variance of GDP growth explained by the factor increase, which leaves us with the evidence that our model is already capturing all the relevant information to predict GDP.

Fifth, although the assumption of forecasting with models of idiosyncratic white noises can be relaxed in large-scale models, the assumption is still present in the euro area empirical applications. The asymptotic theory ensures that this assumption has no major implications for obtaining consistent factors. However, in practice this implies that the forecasts of the indicators are simply linear combinations of the estimated factors since the particular dynamics of the idiosyncratic components are not explicitly modeled. In our parsimonious but dynamically complete specification, the dynamics of the idiosyncratic components have been carefully addressed. This allows us to adequately measure the effect of unexpected news in each indicator on future economic growth.

To sum up, in this paper we use a model which computes accurate short-term forecasts of euro area GDP growth in real time. The forecasts rely on the literature on small-scale models of coincident indicators, which accounts for the specificities of real-time forecasting and the full specification of the idiosyncratic component of each indicator. The name of the model is then based on these features: a model that combines the most commonly used Short-Term INdicators of Growth (STING).

The paper is organized as follows. Section 2 outlines the proposed methodology and analyzes how to deal with mixing monthly and quarterly frequencies of flow data, how to use early estimates of GDP growth and how to estimate the model. Section 3 evaluates the empirical reliability of our method. Section 4 concludes and proposes several further avenues of research.

2 The model

In this section, we develop a state space representation of a model to compute short-term forecasts of euro area GDP growth in real time from a data set that may include mixing frequencies, missing data and data revisions.

2.1 Mixing frequencies

This paper deals with the problem of mixing monthly and quarterly frequencies of flow data by treating quarterly series as monthly series with missing observations. Let G_t be a quarterly series which is observable every third period and whose logs are integrated of order one. In this paper, series with these characteristics are the time series of GDP (second), its announcements (first and flash) and employment. These series are the quarterly aggregates of monthly series, X_t, which are assumed to be observable in this section. Accordingly, we can construct quarterly time series from monthly series by adding the monthly values of the corresponding quarter

$$G_t = 3 \left(\frac{X_t + X_{t-1} + X_{t-2}}{3} \right), \tag{1}$$

which means that the quarterly levels are three times the arithmetic mean. However, handling this definition would imply using non-linear state space models, which is rather troublesome. Mariano and Murasawa (2003) avoid this problem by approximating the arithmetic mean with the geometric mean. It is worth noting that if monthly changes are small, the approximation error is almost negligible.[2] In practice, monthly changes

---

[2]For example, even if we assume a high constant growth of 1% each month (annual growth rate of more than 12%), the difference between the arithmetic and the geometric means is less than 0.4 percentage

in production and employment are small (less than a percentage point) so the geometric approximation is appropriate.

In this context, Proietti and Moauro (2006) and Aruoba, Diebold and Scotti (2009) propose dynamic factor models that permit exact filtering, which avoids the approximation proposed by Mariano and Murasawa (2003). However, their proposals are not free from problems. The former authors develop an exact filter in a non-linear framework which also involves approximations. The latter authors propose a filter that is developed in a dynamic factor model, where the trends of all the indicators used in the filter are assumed to be polynomial trends.

Hence, we assume that the flow data at any quarter is three times the geometric mean of the monthly issues within the given quarter:

$$G_t = 3 \left(X_t X_{t-1} X_{t-2} \right)^{1/3},\qquad(2)$$

which yields

$$\ln G_t = \ln 3 + \frac{1}{3} \left(\ln X_t + \ln X_{t-1} + \ln X_{t-2} \right).\qquad(3)$$

Taking the three-period differences for all t and after some algebra, we can express the quarter-on-quarter growth rates (g_t) of the quarterly series as weighted averages of the monthly-on-monthly past growth rates (x_t) of the monthly series

$$g_t = \frac{1}{3} x_t + \frac{2}{3} x_{t-1} + x_{t-2} + \frac{2}{3} x_{t-3} + \frac{1}{3} x_{t-4}.\qquad(4)$$

2.2 Flash, first and second GDP growth rates

Eurostat officially revises two times the GDP figures that correspond to a given quarter. The first estimate of GDP growth rate in the euro area, y_t^f, is released about 45 days after the end of the respective quarter and this is the so-called flash estimate. Although it is very useful to have an early estimate of GDP, the disadvantage of this flash estimate is that it is based on incomplete information. Using more comprehensive information, the revision of this figure is published about 20 days after the flash and this is the so-called first estimate, y_t^{1st}. In addition, as new information is available, the second estimate of

―――――――――――――――――――――――――――――――――

points.

8

GDP growth rate, y_t^{2nd}, incorporates an additional revision about 40 days after the first and this is the so-called second estimate. Under this revision process, let us call e_1 the revision between the flash and the first, and e_2 the revision between the first and the second.

In this paper, we follow Evans (2005) to propose that data revisions are modeled as noise

$$y_t^f \quad = \quad y_t^{2nd} + e_{1t} + e_{2t}, \tag{5}$$

$$y_t^{1st} \quad = \quad y_t^{2nd} + e_{2t}, \tag{6}$$

where e_{1t} and e_{2t} are independent mean zero revision shocks with variances and $\sigma_{e_1}^2$ and $\sigma_{e_2}^2$, respectively.[3]

The assumption that preliminary announcements are noisy signals of revised series is not new in the literature. Coenen, Levin, Wieland (2005) propose a model of data revisions where Euro area preliminary advances are modeled as final data plus an uncorrelated error. Aruoba (2008) and Swanson and van Dijk (2006) have recently found evidence indicating that preliminary data cannot be considered as rational forecasts of revised data. To show that this is a reasonable assumption in our model, we develop the standard analysis (see Aruoba, 2008) of data revisions to the relationships between first and second.[4] First, we obtain that the mean of the revisions is statistically significant (p-value of 0.007), suggesting that the initial announcements of the statistical agency are biased. Second, we show that revisions are much more correlated with preliminary announcements (correlation of 0.14) than with second (correlation of -0.04) estimates. Third, we find that the noise-to-signal ratio for revisions is 0.18, which is a sizeable value compared with the average of the number presented by Aruoba (2008) of 0.39. Finally, when preliminary data are optimal forecasts of revised data, we expect the variance of latter revisions (from first to second) to be small compared with the variance of prior revisions (from flash to second).

---

[3]For simplicity, we assume that e_{1t} and e_{2t} are uncorrelated. Adding correlation between errors is straightforward, but we think that the available sample of flashes and firsts is still too short to formulate elaborated models.

[4]The results of the analysis of the revisions from flash to second are qualitatively the same. However, we do not include them in the text because computations have been obtained from just 19 observations.

Hence, we expect the news contained in updated data to help to predict final data, which implies that prior revisions should be more volatile than latter revisions. This is not the case for the euro area GDP revision process since the variance of revisions from flash to second is half of the variance of the revisions from first to second.

It is worth noting that we do not explicitly consider all the vintages for all the revisions for all the variables included in our model but only for GDP growth since these are the most relevant in our real-time forecasting exercise. If we did that, we would be talking about more than one hundred variables in the specification. Therefore, we would not have a small-scale model and we could not follow the parsimonious principle of this paper. To account for the revisions of all the variables, the proper approach can be found in Altavilla and Ciccarelli (2007), but that would be beyond the scope of our paper.

2.3 State space representation

To consider the notion of co-movements among the GDP series and the economic indicators, the time series are modeled as the sum of two orthogonal components. The first component is the common factor, f_t, and reflects the notion that the series dynamics are driven in part by common shocks. The second component captures the idiosyncratic behavior of each series.

For clarity, let us start by assuming that all variables are always observed at a monthly frequency. Monthly growth rates of quarterly series and monthly growth rates of hard indicators are assumed to exhibit a direct relationship to the common factor which measures the common component of the monthly growth rates of these series. However, we have to treat the relationship between the common factor and the soft indicators differently. The reason for this is that, according to the European Commission (2006), the guiding principle for the selection of questions in the survey is the aim of achieving as high as possible coincident correlation of the confidence indicator with year-on-year growth of the reference series. Hence, taking first differences could involve over differentiation. Accordingly, we have to relate the level of soft indicators to the year-on-year common growth rate, which can be written as the sum of current values of the common factor and its last eleven lagged values.

Let us collect the r_h hard indicators in the vector Z_t^h and the r_s soft indicators in the vector Z_t^s. Let l_t be the quarterly employment growth rate, and let u_{1t}, u_{2t}, U_t^h, and U_t^s be the scalars and r_h-dimensional and r_s-dimensional vectors which determine the idiosyncratic dynamics of GDP, unemployment and the economic indicators, respectively. The *measurement equation* can be defined as

$$
\begin{pmatrix} y_t^{2nd} \\ Z_t^h \\ Z_t^s \\ l_t \\ y_t^{1st} \\ y_t^f \end{pmatrix} = \begin{pmatrix} \beta_1 \left(\frac{1}{3} f_t + \frac{2}{3} f_{t-1} + f_{t-2} + \frac{2}{3} f_{t-3} + \frac{1}{3} f_{t-4} \right) \\ \beta_2 f_t \\ \beta_3 \sum_{j=0}^{11} f_{t-j} \\ \beta_4 \left(\frac{1}{3} f_t + \frac{2}{3} f_{t-1} + f_{t-2} + \frac{2}{3} f_{t-3} + \frac{1}{3} f_{t-4} \right) \\ \beta_1 \left(\frac{1}{3} f_t + \frac{2}{3} f_{t-1} + f_{t-2} + \frac{2}{3} f_{t-3} + \frac{1}{3} f_{t-4} \right) \\ \beta_1 \left(\frac{1}{3} f_t + \frac{2}{3} f_{t-1} + f_{t-2} + \frac{2}{3} f_{t-3} + \frac{1}{3} f_{t-4} \right) \end{pmatrix} +
$$

$$
\begin{pmatrix} \frac{1}{3} u_{1t} + \frac{2}{3} u_{1t-1} + u_{1t-2} + \frac{2}{3} u_{1t-3} + \frac{1}{3} u_{1t-4} \\ U_t^h \\ U_t^s \\ \frac{1}{3} u_{2t} + \frac{2}{3} u_{2t-1} + u_{2t-2} + \frac{2}{3} u_{2t-3} + \frac{1}{3} u_{2t-4} \\ \frac{1}{3} u_{1t} + \frac{2}{3} u_{1t-1} + u_{1t-2} + \frac{2}{3} u_{1t-3} + \frac{1}{3} u_{1t-4} \\ \frac{1}{3} u_{1t} + \frac{2}{3} u_{1t-1} + u_{1t-2} + \frac{2}{3} u_{1t-3} + \frac{1}{3} u_{1t-4} \end{pmatrix} + \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \\ e_{2t} \\ e_{1t} + e_{2t} \end{pmatrix}, \quad (7)
$$

where $U_t^h = (v_{1t}, ..., v_{r_h t})'$, $U_t^s = (v_{r_h+1 t}, ..., v_{rt})'$, and $r = r_h + r_s$. The factor loadings, $\beta = \begin{pmatrix} \beta_1 & \beta_2' & \beta_3' & \beta_4 \end{pmatrix}'$, measure the sensitivity of each series to movements in the latent factor and have dimensions that make them conformable with each equation.

The dynamics of the model are achieved by assuming that

$$
f_t = a_1 f_{t-1} + ... + a_{m_1} f_{t-m_1} + \epsilon_t^f, \tag{8}
$$

$$
u_{1t} = b_1 u_{1t-1} + ... + b_{m_2} u_{1t-m_2} + \epsilon_t^{u_1}, \tag{9}
$$

$$
v_{jt} = c_{j1} v_{jt-1} + ... + c_{jm_3} v_{jt-m_3} + \epsilon_t^{v_j}, \tag{10}
$$

$$
u_{2t} = d_1 u_{2t-1} + ... + d_{m_4} u_{2t-m_4} + \epsilon_t^{u_2}, \tag{11}
$$

where $\epsilon_t^f \sim i.i.d.N\left(0, \sigma_f^2\right)$, $\epsilon_t^{u_1} \sim i.i.d.N\left(0, \sigma_{u_1}^2\right)$, $\epsilon_t^{v_j} \sim i.i.d.N\left(0, \sigma_{v_j}^2\right)$, with $j = 1, ..., r$, and $\epsilon_t^{u_2} \sim i.i.d.N\left(0, \sigma_{u_2}^2\right)$. All the covariances are assumed to be zero. The identifying

assumption implies that the variance of the common factor, σ_f^2, is normalized to a value of one.

More compactly, we use the expression for the measurement equation

$$Y_t = Hh_t + w_t, \tag{12}$$

with $w_t \sim i.i.d.N\left(0, R\right)$. In addition, the *transition equation* can be stated as

$$h_t = Fh_{t-1} + \xi_t, \tag{13}$$

with $\xi_t \sim i.i.d.N\left(0, Q\right)$. An extensive description of what these equations look like for the empirical model has been set out in Appendix A.

To handle missing observations, following Mariano and Murasawa (2003) we replace the missing observations with random draws θ_t from $N(0, \sigma_\theta^2)$ which are independent of the model parameters.[5] The substitutions allow the matrices to be conformable but they have no impact on the model estimation since the Kalman filter uses the data generating process of the normal distribution for them. In that sense, the missing observations simply add a constant to the likelihood function of the Kalman filter process. Let Y_{it} be the i-th element of the vector Y_t and R_{ii} be its variance. Let H_i be the i-th row of the matrix H which has α columns and let $0_{1\alpha}$ be a row vector of α zeroes. In this case, the measurement equation can be replaced by the following expressions

$$Y_{it}^* = \begin{cases} Y_{it} & \text{if } Y_{it} \text{ observable} \\ \theta_t & \text{otherwise} \end{cases}, \tag{14}$$

$$H_{it}^* = \begin{cases} H_i & \text{if } Y_{it} \text{ observable} \\ 0_{1\alpha} & \text{otherwise} \end{cases}, \tag{15}$$

$$w_{it}^* = \begin{cases} 0 & \text{if } Y_{it} \text{ observable} \\ \theta_t & \text{otherwise} \end{cases}, \tag{16}$$

$$R_{iit}^* = \begin{cases} 0 & \text{if } Y_{it} \text{ observable} \\ \sigma_\theta^2 & \text{otherwise} \end{cases}. \tag{17}$$

This trick leads to a time-varying state space model with no missing observations so the Kalman filter can be directly applied to Y_t^*, H_t^*, w_t^*, and R_t^*. Let $h_{t|\tau}$ be the estimate of

---

[5]Filling in missing observations with means, medians or zeroes would be valid.

h_t based on information up to period τ and let $P_{t|\tau}$ be its covariance matrix. With this notation, the *prediction equations* are

$$h_{t|t-1} = Fh_{t-1|t-1}, \tag{18}$$

$$P_{t|t-1} = FP_{t-1|t-1}F' + Q. \tag{19}$$

The prediction errors are $\eta_{t|t-1} = Y_t^* - H_t^* h_{t|t-1}$ with covariance matrix $\zeta_{t|t-1} = H_t^* P_{t|t-1} H_t^{*\prime} + R_t^*$. Hence, the log likelihood can be computed in each iteration as

$$l_t = -\frac{1}{2}\ln\left(2\pi\left|\zeta_{t|t-1}\right|\right) - \frac{1}{2}\eta'_{t|t-1}\left(\zeta_{t|t-1}\right)^{-1}\eta_{t|t-1}. \tag{20}$$

The *updating equations* are

$$h_{t|t} = h_{t|t-1} + K_t^*\eta_{t|t-1}, \tag{21}$$

$$P_{t|t} = P_{t|t-1} - K_t^* H_t^* P_{t|t-1}, \tag{22}$$

where the Kalman gain, K_t^*, is defined as $K_t^* = P_{t|t-1}H_t^{*\prime}\left(\zeta_{t|t-1}\right)^{-1}$. The initial values of $h_{0|0}$ and $P_{0|0}$ used to start the filter are a vector of zeroes and the identity matrix, respectively. Note that when at any date τ all the elements of the vector Y_τ are not observed, the updating equation is $h_{\tau|\tau} = h_{\tau|\tau-1}$ and time τ does not change the estimated dynamics of the model. This feature can be used to easily compute forecasts by adding missing data for all the variables in the model at the end of the sample.

As documented by Banbura and Rustler (2007), the Kalman filter allows computation of the contribution of each series to GDP forecasts. Substituting the prediction errors $\eta_{t|t-1}$ and (18) into the updating equation (21), one obtains

$$h_{t|t} = (I - K_t^* H_t^*)Fh_{t-1|t-1} + K_t^* Y_t^*. \tag{23}$$

Now, when the Kalman filter is close to its steady state, this expression becomes

$$h_{t|t} = M_t^*(L)Y_t^*, \tag{24}$$

with the elements of the matrix of lag polynomial $M_t^*(L) = (I - (I - K_t^* H_t^*)FL)^{-1}K_t^*$ measuring the effects of unit changes in the lags of individual observations on the inference

of the state vector $h_{t|t}$. Letting M_{jt}^* be each of these matrices, the inference on the state vector can be decomposed into a weighted sum of observations

$$h_{t|t} = \sum_{j=0}^{\infty} M_{jt}^* Y_{t-j}^*. \tag{25}$$

In this respect, $M_t^*(1) = (I - (I - K_t^* H_t^*) F)^{-1} K_t^*$ is a matrix that contains the cumulative impacts of the individual observations in the inference of the state vector.

Combining this relationship with the first row of equation (7) which shows that GDP can be decomposed into the sum of its unobservable components, one can compute the cumulative impact of each indicator on the forecast of GDP growth. For the empirical illustration stated in Appendix A, this measure can be easily obtained as follows

$$\psi_t = \beta_1 \left(\frac{1}{3} m_{1t}^* + \frac{2}{3} m_{2t}^* + m_{3t}^* + \frac{2}{3} m_{4t}^* + \frac{1}{3} m_{5t}^* \right)' + \left(\frac{1}{3} m_{13t}^* + \frac{2}{3} m_{14t}^* + m_{15t}^* + \frac{2}{3} m_{16t}^* + \frac{1}{3} m_{17t}^* \right)',$$
$$\tag{26}$$

where m_{it}^* is the i-th row of $M_t^*(1)$, and ψ_t is a vector which contains the cumulative forecast weight of each indicator.

3   Empirical results

3.1   Data description and indicators selection

The variables entering the proposed model are listed in Table 1 and plotted in Figure 1. Given their interest for real time forecasts, the particular date on which these series are published and the samples that they cover are also shown in the figure. Note that on the day on which the paper was written, February, 11th 2008, GDP growth and its announcements were available for 2007.3, but none of these figures were available for 2007.4.[6]

The list of indicators included in the dynamic factor model can be classified into three groups. The first group contains quarterly indicators. Apart from the second release of GDP and its early estimates (flash and first), this group includes the quarterly series of employment. The second group of indicators is formed by monthly hard indicators based

---

[6] To understand notation, for example 2007.1 or 07.1 refer to first quarter of year 2007 while 2007.01 or 07.01 refer to first month of year 2007.

on economic activity data. Specifically, these are the euro area Industrial Production Index (IPI, excluding construction), the Industrial New Orders index (INO, total manufacturing working on orders), the Euro area total retail sales volume, and the extra-Euro area exports. Table 1 shows that these indicators exhibit large publication delays that range from 35 to 52 days. The last group of time series consists of soft indicators, which are based on survey data. The soft indicators included are the Euro-zone Economic Sentiment Indicator (ESI), the German business climate index (IFO), the Belgian overall business indicator (BNB), and the Euro area Purchasing Managers confidence Indexes (PMI) in the services and manufacturing sectors. The main characteristic of soft indicators is that they are promptly available and can be observed in Table 1 since these indicators are available on a timely basis within the reference month.

The set of indicators included in the model have been selected by successive enlargements of the basic model based on Stock and Watson (1991). Accordingly, we start the model with the euro area counterparts of their core set of variables: IPI, Retail Sales and Employment.[7] We extend the model to include the GDP second and its preliminary announcements, the flash and first estimates. In addition, we include the soft indicators which are most promptly available in the Euro area: BNB, IFO, PMI services, PMI Manufacturing and ESI.[8]

Since macroeconomic data are very collinear, it is reasonable to conjecture that including (probably more noisy) additional variables may not improve forecasting accuracy and that it might be worth focusing on some key variables by following some statistical selection procedure. For this purpose, having defined the set of core variables, we propose a method to decide whether new indicators should be added to this core. The method, which is based on the assumption that the primary focus of the model is to provide forecasts of GDP growth, consists of adding a variable only when it increases the percentage of variance of GDP growth explained by the common factor. Accordingly, the method involves screening out those additional indicators that capture idiosyncratic dynamics and

---

[7]We do not include Personal Disposable Income because we do not have this series for the euro area.

[8]To select the soft indicators, we use those mentioned in the monthly calendar of JP Morgan to bring to the attention of their clients as being the relevant set of variables for analyzing euro-area developments. Within the soft indicators, we include two national series as they are the first to be released in the area.

that do not lead to better fit for GDP growth through the common component. To start with, we obtain that the percentage of variance explained by the factor which is computed from the core set of variables is 74.15%. Adding Exports and INO increase this percentage to 74.85% and to 78.10%, respectively. However, on the same criteria we rule out the inclusion of different financial indicators such as term premium and risk premium and other commonly used real variables such as vehicle registrations. It is worth noting that our final set of indicators is reasonable. Our indicators coincide (except for construction production) with those of aggregate economic activity in the set of key indicators available daily on a single webpage by Eurostat.[9]

To examine the extent to which we could omit valuable information by using a model which uses a reduced number of indicators, we consider the following exercise. We construct a database following precisely the description suggested in Angellini et al (2008, page 22). With the same data transformation that they propose, we extract five factors from their eighty-five series by using two methodologies, Stock and Watson (2002) and Giannone et al. (2008). These factors are able to explain about 50% of the time series variability of the indicators, so we consider that these factor extraction methods work well in this context. Finally, we use these two sets of five factors as additional explanatory variables in our Euro-STING model. We expect that if these factors contain additional information about the Euro area GDP growth that our small-scale model has omitted (recall that, apart from flash and first estimates, our model is "only" based on 11 variables), the variance of GDP explained by the extended model should be greater than in the original Euro-STING proposal. However, the variance of GDP explained by the extended models is lower than in the original specification. In particular, the percentage explained by the factors goes down from 78.1% to 73.0%, in the case of those estimated with Giannone et al. (2008) and to 74.4% in that of those estimated with Stock and Watson (2002). To consider all the possibilities, we also extend our model by including each of the eighty-five variables one by one (except for the variables which are already included in our model) and we obtain that the variance of GDP explained by the factor decreases with each ad-

---

[9] Obviously, Eurostat do not include all the soft indicators that we include in our model because some of them are produced by different institutions.

16

ditional variable included in the model. In particular, the variance of GDP explained by the factor in each of these enlarged models ranges from 71.0% to 77.5%. Hence, it seems that our model is capturing all the relevant information when computing the factor which is the common driving force of euro-area growth developments. Or, in other words, the additional variables typically used in euro area applications are capturing other types of co-movements across variables that do not necessarily lead to better fit for GDP, which is the key variable to be forecast in our paper. This result confirms that our initial selection of variables is appropriate.

In the context of comparing small versus large-scale factor models, this finding deserves some comments. Users of large-scale factor models have extensively stressed that strict factor models rely on the tight assumption that the idiosyncratic noise is cross-sectionally orthogonal. However, large-scale factor models are not exempt from theoretical assumptions that may not hold in empirical applications. It has been proved that large-scale models lead to factors which are consistent under certain assumptions. The basic assumptions behind these theoretical results are that the number of variables and observations must tend to infinity, that the idiosyncratic components should be only weakly correlated (in time series and cross section), and that the variability of the common component cannot be too small. However, there are reasonable grounds for raising some doubts about the empirical reliability of such theoretical assumptions in the euro area context. First, although the number of time series which are available in real time at a high degree of dis-aggregation and reduced cost is enormous, the number of time series effectively employed in euro area applications does not tend to infinity. For example, Angelini et al. (2008) and Banbura and Runstler (2007) use 85 and 76 time series, respectively. Second, cross correlation can be larger in empirical applications than is warranted by asymptotic theory. In addition, Boivin and Ng (2006) document that warnings may arise because of the composition of the empirically compiled data, since when data are drawn from a small number of broad categories the risk of correlated idiosyncratic errors increases dramatically. In this context, Bai and Ng (2008) also show that forecasting with targeted predictors, which in practice implies imposing zero loads for some of the variables of the large-scale models (i.e. reducing the scale of the model), may improve the forecasting performance even when

17

the weak cross-correlation holds.

On the basis of this evidence, we propose a checking method to examine the empirical reliability of the theoretical assumptions about the covariance of idiosyncratic components of factor models. The reduced dimensions of our model allow us to consider a modified version of Engle's (2007) MacGyver method to estimate the covariance matrix of the idiosyncratic components. For this purpose, we make the reasonable assumption that soft indicators are cross-correlated with each other, that hard indicators are also cross-correlated, but that hard and soft indicators are not cross-correlated with each other. We now estimate the cross-correlation matrix of blocks (the ones which refer to soft and hard indicators) separately and include the estimated cross-correlation parameters in the model. The correlations estimated for soft indicators range from a maximum of 0.25 to a minimum of 0.06. The hard indicators have much smaller cross correlations since their maximum is only about 0.02.

Apart from the size of the correlations, we obtain the interesting result that the impact of having estimated an strict factor model in terms of forecasting GDP growth and in terms of estimating the factors is negligible. The correlation between GDP growth rates obtained from a factor model which estimates the cross correlation of the idiosyncratic components and those obtained from our factor model (which assumes a diagonal cross-correlation matrix) is 0.98. In addition, the correlation between the factors obtained from these two alternative specifications is 0.97. Therefore, although strict factor models such as ours assume a diagonal cross-correlation matrix among idiosyncratic components that can potentially distort factor estimation and forecasting, we find that our model is empirically robust to this assumption at least in the application presented in this paper.

Depending on the nature of the data, the indicators used in this model are transformed in different ways.[10] The three GDP releases and employment are used in the form of quarterly growth rates. Hard indicators are transformed by taking monthly growth rates. However, soft indicators are included in levels. Although considering soft indicators in levels can complicate the estimation of the model, Section 2.3 described economic reasons for using them in this form. The model specification suggests that the potential problem

---

[10]These transformations imply that although some series are integrated they are not cointegrated.

of the inertia in these series is accounted for by considering that the series are related with a twelve-lag moving average of the monthly growth rates, which refers to the annual growth rates. Finally, all series have been normalized to have zero mean and unit variance.

Following the method outlined in Section 2, missing data are conveniently replaced by random numbers which have been generated from N(0,1). Figure 1 provides a clear overview of the importance of missing data in the euro area forecasting exercises. First, many series start too late. Retail sales, industrial new orders, exports, employment, BNB and PMI start in the second half of the 1990s, and flash and first GDP are only available for the last four and nine years, respectively. Second, hard indicators exhibit a publication delay of one or two months which leads to missing data at the end of the sample. Finally, quarterly series do not contain monthly figures and, subject to the standard publication delays, only the third month of each quarter is available.

3.2   In-sample analysis

The in-sample analysis was carried out using the latest data set available on February, 11th 2008. To illustrate the lack of synchronization in the data release dates, the last rows of this vintage are reported in Table 2. In this table we can observe the particularities of real-time forecasting. Data for quarterly series only appear in the third month of each quarter and, although the vintage refers to 2008, the figures for 2007.4 were still not available. Soft indicators contain data until January 2008 while hard indicators exhibit their typical publication delays of one or two months. On the next forecasting dates, but not in this vintage, preliminary estimates of GDP growth (flash and first) were already available for 2007.4.

To understand how the model predicts, recall that our interest is in short-term forecasting. For this purpose, the model has been developed to forecast a rolling window of nine months that moves forward in accordance with to the publication date of the second estimates. The day we wrote this paper, the latest available second release of GDP was for 2007.3, which was released on January, 9th 2008. Hence, from this date until April, 9th 2008 (the release date for the second GDP estimate for 2007.4), the forecast of GDP that our model produces covers the period from October 2007 to June 2008 (2007.4, 2008.1

19

and 2008.2). From April, 9th 2008, the forecasts will cover the nine months from January 2008 to September 2008 (the first three quarters of 2008). Since second estimates are among the observed variables in the measurement equation, these nine-month forecasts can be obtained directly from the Kalman filter iterations by imposing nine months of missing observations after the latest figure available for second releases. Accordingly, the in-sample data vintage reported in Table 2 shows missing observations for GDP growth from October 2007 to June 2008.

The model adopted in this paper is based on the notion that comovements 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 2 plots the factor (left-hand scale) and the Eurocoin (right-hand scale) published by the CEPR, which is probably the leading coincident indicator of the euro area business cycle. The similarities between their business cycle dynamics are striking suggesting that they track the same business cycle.

The maximum likelihood estimates of the factor loadings, which reflect the degree to which variation in each observed variable can be explained by the latent series, are reported in Table 3.[11] In all cases the estimates are positive and statistically significant, indicating that these series are procyclical, i.e. positively correlated with the common factor. Although all the series contain incremental information about the euro area business cycle pattern, there are some differences in the absolute sizes of the corresponding factor loadings. Our estimates show that real activity data exhibit the highest loading factors. In particular, the highest impact of the common component is on industrial production (0.21), closely followed by industrial new orders (0.19) and second GDP (0.12). However, loading factors of soft indicators tend to be lower than those of real activity data and all of them are below 0.07. As we will examine later on, this result should not necessarily be interpreted as evidence against survey data. These in-sample estimates may reflect the fact that ignoring the timeliness advantages of soft indicators would diminish their role in factor models when hard indicators are available.

Second GDP forecasts can be examined in Figure 3 and Table 4. Figure 3 plots

---

[11] Other maximum likelihood estimates are available from the authors upon request.

the monthly estimates of GDP quarterly growth rates along with their actual values. In accordance with the methodology employed in this paper, the Kalman filter anchors monthly estimates to actual whenever GDP is observed. Hence, for those months where GDP is known, the actual value and the estimates coincide. Table 4 (Panel A) shows how our model anticipates the next issues of the Eurostat data release process. The key releases are the second GDP growth rates for quarters 2007.4, 2008.1 and 2008.2, which we call lagged, current and future forecasts, respectively. In addition, this table presents the predicted values for the next three quarters of flash and first estimates and their standard deviations.

In addition, recall that one of the distinguishing advantages of our model is that it proposes a complete dynamic specification for all the indicators. This allows us to compute accurate forecasts not only for GDP but also for the whole set of indicators used to estimate the dynamic factor model. These forecasts are crucial for understanding how the next values of these indicators affect the GDP forecast.[12] Table 4 (Panel B) shows the forecasts for the next unavailable month of each indicator.

Let us now examine the cumulative forecast weights of each indicator. The question to be analyzed here is the relative importance of each indicator to forecast GDP growth. One example could be helpful for an intuitive understanding of this exercise. Suppose that, for a given month, we only have one indicator available, let's say, the BNB. Obviously, the BNB will have 100% of the weight because it is the only information set "responsible" for the changes in the forecast of GDP in that given month. When new information becomes available for that month, the relative weight of that series will decline. Table 5 shows the evolution over the latest months of these forecast weights (normalized to add up to one) in GDP growth. Firstly, we concentrate on quarterly series. In accordance with the anchoring characteristic of our proposal, the rows labelled as 2007.06 and 2007.09 show that, when the second GDP estimate is published, the cumulative forecast weights of all the indicator series in GDP forecasts are zero.[13] In addition, cumulative weights for

---

[12]Imposing white noise idiosyncratic dynamics would produce very naive forecasts since it would restrict them to being just weighted averages of the set of factors, with weights equal to the factor loadings.

[13]The intuition behind this result is that once GDP is available, its figure is a *sufficient* statistic to forecast GDP.

21

quarterly series are zero for the first two months of each quarter since the observations are missing and they do not add any information to the Kalman filter. It is important to notice that, using this specific data set, the flash and first estimates and employment always have zero weights. These series only have weights during the periods in which they are available but the corresponding GDP second estimate is not.[14]

Finally, the evolution of the weights of monthly indicators depends heavily on the nature of the indicators and the date on which they are computed. Cumulative weights to forecast second values up to 2007.11 are high and concentrated among hard indicators, basically IPI and INO. However, the reported weights for soft indicators are typically smaller for those forecasts. In line with the results of Banbura and Runstler (2007), this should not be interpreted as a failure of soft indicators to incorporate useful forecast information about GDP growth. It simply means that they contain limited information in addition to the real activity data when the latter have already been published. Once their more timely publication takes place, business surveys gain importance in short-term forecasts. Accordingly, the table reports significant improvements in the cumulative weights of soft indicators in the forecast of GDP growth for 2007.12. In this month, IPI, INO and Exports are not available and the two highest weights are for soft indicators (weights of PMIs are 0.30 and 0.21 for manufactures and services, respectively) followed by Sales (weights of 0.19). In 2008.01, hard indicators are no longer available so only survey data exhibit positive weights to forecast GDP.[15]

Figure 4 illustrates how the model can be used to evaluate the reaction of GDP forecasts to different next issues of the indicators. The ESI indicator was last updated on January, 31th 2008, the expected value of ESI for that date was 104.2, and the estimate of the GDP growth rate for 2008.1 associated with this expected value was 0.47. The day before the ESI update, we use the Kalman filter with the last vintage of ESI but where the observation of January 2008 was filled in with simulated values from -15 to 220 The figure plots the GDP forecasts associated with these simulated values of ESI and shows the GDP forecast changes due to potential ESI departures from its expected value. The actual realization

---

[14]In these cases flash and first cumulative weights are about 0.8.

[15]We will come back to the role of soft indicators in real-time forecasting in the next section.

of ESI was 101.7 which implied a decrease in the predicted GDP growth of about 0.06.

It is worth noting the logistic shape of GDP responses in Figure 4. The intuition behind these responses is simple. Recall that the state vector updates according to two sources of variation, the prediction error and the Kalman gain which decreases as the variance of the state vector increases. As the generated values of ESI diverges from the expected value, the forecast error increases but, for extreme departures of the indicator simulations from its expected value, the Kalman gain becomes negligible and the state vector remains almost unchanged so the forecasts of GDP growth become flat.

3.3 Real-time analysis

As Croushore and Stark (2001) have pointed out, developing a real-time data set is conceptually simple. However, producing real-time data requires a great amount of effort in practice since one has to handle old data sources that sometimes exist only in printed form. In addition, the data set should always follow the principle of putting whatever data were available at the time into the corresponding cell in order that, on each day of the forecast, only the time series information available on that day is used. However, it is worth constructing such databases since evaluating forecast errors by using latest-available data is questionable since measures of forecast error, such as root-mean-squared error and mean absolute error, can be deceptively lower when using latest-available data rather than real-time data (Stark and Croushore, 2002). In accordance with this principle, we have constructed a data set that gives forecasters a picture of the data that were available on any given day in the period 2004-2008.

It is worth noting that, following the principle of computing model-based forecasts by using exactly the same amount of information that was available on the day the competitor published its forecast, we are precluded from using the euro area Business Cycle Network real-time data set. The reason is that the existing data are collected into monthly vintages (not updated daily), which end up with a significant delay of about half a year. By contrast, for each day on which a particular series of our data set was updated, we collect the whole set of time series available at that moment in "vintages" that we call vint-mm/dd/yy. These vintages are kept fixed until the day that a new series was updated. Hence, we

have compiled different vintages which contain only the information available on the days of publication of new data, so we can mimic the forecasting procedure that a forecaster would have performed at any time during the last few years. The first vintage for which we could collect data for all indicators was vint-01/02/04 and therefore we start the real-time analysis with the forecast of GDP growth in 2003.4, for which estimates were still not available at that time. We end up with 424 different vintages for the period 01/02/04 to 02/05/08.

Using the first vintage of our data set, called vint-01/02/04, we estimate the model and compute the nine-month GDP forecasts that include lagged (2003.3), current (2004.1) and future (2004.2) forecasts, as described in the in-sample exercise. In order to keep the exercise feasible, we keep the estimated parameters fixed for the next 28 vintages until the second release for 2003.4 is published on 04/16/04. The model is then re-estimated and the procedure is recursively repeated until we reach the last vintage of our data set, vint-02/05/08, which was also used to perform the in-sample analysis.

To illustrate how the real-time forecasting exercise behaves in each forecasting period that includes the nine-month forecasts, we plot in Figure 5 the real-time forecasts that were made each day in two different forecasting periods. The top chart includes all the latest forecasts of GDP growth for 2007.4, i.e. forecasts made from 07/12/07 (publication day of 2007.1 GDP) to 02/11/08 (today). To evaluate uncertainty, the figure also displays the associated one standard deviation error bands. With this chart we can examine the model's reactions to the financial turbulence that took place during the summer of 2007. PMI services and manufactures were the first series to incorporate information about the business climate in our model. In September 2007, these series fell by about -3.82 and -1.13 points, with the former figure representing the largest decline in the history of PMI services. In addition, BNB, IFO and ESI also exhibited strong declines of about -1.8, -1.6, -3.1, respectively. As can be observed in the figure, the declines in GDP growth forecasts came soon after. As soon as these data were introduced in the model, the GDP forecast fell by 0.3 percentage points. After one month of low forecasts, the recovery of most of the survey data and the relatively better news that came from real activity data (especially from IPI which grew about 0.5% in August) let GDP forecasts rise, partially offsetting

24

the summer's falls.

The bottom chart of Figure 5 plots the forecasts of the second forecasting exercise. The figure shows the real-time forecasts of GDP for 2006.3 over the period from 04/12/06 to 01/10/07. This period is particularly interesting since, being a recent period, it reflects the revisions undergone by second-release GDP growth. The revision of this quarter is apparent in the difference between the data actually published as they were available in real time (bottom horizontal line) and the data as they appear in the current revision (top horizontal line). Although both flash and first release were about 0.52, which roughly coincides with the figure issued on 01/11/07, we were forecasting a second-release growth of almost 0.58 for 2006.3. However, Eurostat revised upwards this figure to 0.58 in the GDP time series published on 01/09/08. This example is illustrative of the importance of truly real-time exercises using current-vintage data sets instead of end-of-sample vintage data sets to assess real-time forecasting performance.[16] In addition, the exercise illustrates that, as suggested by Orphanides and van Norden (2002), real-time forecasts must be compared with the last vintage as *final* data in the measurement of output.

The relationship between new, updated incoming information and the forecast error is examined in Figure 6. This figure plots the sample average of the standard errors associated with each GDP forecast for the 275 days that each forecasting exercise lasts. Although the standard errors may vary somewhat from quarter to quarter, on average the uncertainty about the GDP forecast continuously decreases during the forecasting period. The forecast uncertainty falls about one-third during the first 200 days as information from the indicators becomes available to compute the forecasts. The variance then falls significantly following the flash releases. However, the falls in uncertainty provided by the first releases are of much less importance. This pattern indicates that the first releases provide little new information about GDP growth beyond that already contained in the flash estimates.

One additional exercise of interest is to examine the forecasting accuracy of our model with respect to the preliminary announcements of GDP growth. For this purpose, Table 6 lists the mean squared errors (MSEs) which compare the preliminary announcements with

---

[16]The out-of-sample data sets are based on final data that are cut in some date and sequentially enlarged.

the revised second-release GDP growth rates (vintage 01/09/08), along with the MSE from Euro-STING forecasts. These were calculated for different days of the forecasting process: the days before and after flash and first releases. According to this table, the Euro-STING forecasting accuracy on the days before flash releases (MSE of 0.027) is similar to but slightly worse than that of the preliminary announcements (MSE of 0.024 and 0.025).[17] However, the MSE of the Euro-STING forecast on the day on which the flash is released is 0.022, which reduces the MSE of the flash estimates themselves. A similar situation is seen with the first releases. But in this case, incorporating the information of first releases in the model leads to dramatic reductions in the MSE, which falls to 0.014. Accordingly, preliminary announced GDP cannot be considered as the most accurate forecast of the revised GDP figures. Using the upcoming information from all the indicators is important to improve upon the forecasting accuracy in real time.

Before ending the real-time forecasting section, let us take up again the issue of the importance of timely information contained in soft indicators. Figure 7 plots the relative cumulative forecast weights of all the observations corresponding to the first quarter of 2007 in the forecast of GDP for that quarter.[18] As can be seen, on the publication date of BNB for January (01/24/07), it was the only indicator available in that quarter from which to infer GDP growth for that period. Accordingly, BNB receives 100% of the relative forecast weight. As new information from other indicators becomes available, the relative forecast weights decrease until BNB is published in February (02/24/07), when there is a new peak. The intuition for this peak is that there are two values of the BNB that affect the inference of GDP for that quarter but only one issue at most for the other monthly indicators. Following the same reasoning, BNB weights decrease until the new peak corresponding to March 2007. After this peak, there is a long decline in BNB weights as hard indicators become available. The last dramatic decline relates to the publication

---

[17]It is worth pointing out that the main benefit of flash releases comes from just one quarter, 2005.Q4. Taking out this quarter, there is no additional information in the flash release which is not already contained in the Euro-STING model.

[18]We use 2007.1 because we wish to analyze the changes in weights in a single series and from that time onwards, PMIs were released before BNB. Weights for PMI are more difficult to interpret since they refer to manufactures and services.

of the flash estimate for this quarter on 05/15/07. Finally, weights collapse to 0 when second GDP growth for 2007.1 is published on 7/12/07. This real-time exercise reinforces the previous results that survey data contain valuable information for forecasting GDP growth apart from that contained in real activity data once their more timely publication is taken into account.

3.4 Forecasting accuracy

Let us come back to the main purpose of the paper: to show that an automated small-scale factor model algorithm is able to forecast the euro area GDP growth at least as well as professional forecasters. The former uses a clear, easy-to-replicate methodology, while the latter not only employ the best forecasting tools, but also are allowed to incorporate their own judgement.

To connect this analysis with the euro area large-scale dynamic factor applications for forecasting GDP growth such as those of Angelini et al. (2008) and Barhoumi et al. (2008), it is worth stressing that they only compare their out-of-sample forecasts with a set of benchmark models that might not be the most commonly used in practice. However, we make the effort of comparing in real time our forecasts with the most influential forecasts in the European forecasting arena. In addition, the relatively short sample for real-time comparison (16 quarters) is standard in the real-time forecasting literature. For example, Giannone et al. (2006) acknowledge that they use pure real-time vintages from July 2003 to March 2005, which implies a real-time analysis of 8 quarters, and Evans (2005) conducts real-time analysis from April, 11th 1993 to June, 30th 1999, which includes 26 quarters.

Visual inspection of Figure 8 shows that the real-time forecasting accuracy of our model is good. This figure plots the forecasts for the most immediate quarter of GDP growth of our nine-month forecasting exercise, which were carried out every day of the real-time forecasting period. The figure also displays the latest available GDP growth figures (vintage vint-02/05/07), which include the data revisions. In general, final values of GDP growth lie within the two-standard error band represented by the shaded area.

To assess the relative forecasting accuracy of the real-time forecasts, Table 7 shows the mean squared errors of our forecast and those of well-known forecasts of euro-area GDP

growth rate. These include the Eurocoin forecasts, the IFO-INSEE-INSAE economic forecasts, the European Commission macroeconomic forecasts, the projections of the OECD Economic Outlook, and the euro-area GDP growth projection of DG ECFIN.[19]

The Euro-STING forecasts are updated daily and on each of these days the model computes GDP growth forecasts for the next nine months. Competing forecasts are published with lower frequency and some of them are computed at different horizons. For this reason, the first three columns of this table compare forecasts made with different leads and lags with respect to the target GDP figure. Care has been taken to compare forecasts which have the same forecasting horizon and are available on the same day on which the competitor publishes its release. Note that this works to the detriment of Euro-STING, which could obtain better forecasts since it is updated daily and could use more up-to-date information than its competitors which do not change their forecasts for a whole month or even a quarter.

In terms of mean squared forecast error, our simple automated model beats almost all of its competitors at all the forecasting horizons. As shown in Table 7, the Euro-STING model outperforms the Eurocoin forecasts, the IFO-INSEE-INSAE economic forecasts and the European Commission macroeconomic forecasts for all the forecasting horizons. Relative to the DG ECFIN projections, the Euro-STING shows better forecasting performance as the forecasting horizon increases. The Euro-STING forecasts are also comparable with those of the best of its competitors, i.e. those of the OECD Economic Outlook.

In spite of the good forecasting performance, and as expected due to the size of the sample, the pairwise comparisons of forecasts with standard statistical Diebold and Mariano (1995) tests usually cannot reject the null hypothesis of no difference in forecasting accuracy. In particular, the p-values of the modified Diebold-Mariano tests go from 0.037 to 0.999. However, these results can be interpreted positively since we can conclude that the Euro-STING does the job of forecasting euro area GDP growth as well as the professional forecasters. In addition, the sign unambiguously favours our model, which would be a very rare outcome if our forecasts were indistinguishable from the most influential euro area GDP growth forecasts in the forecasting arena. Finally, it should be stressed

---

[19]See Appendix B for a description of these forecasts.

that the Euro-STING forecasts have a significant advantage since they are updated in a timely manner. Forecasts are updated daily as new information becomes available, which permits day-to-day monitoring in the euro area.

To provide more empirical evidence in favor of the Euro-STING forecasts, Table 7 also presents the p-values of the forecast encompassing test based on testing the significativity of α_1 in the OLS regression

$$y_t^{2nd} - \widehat{f}_{t,Euro-STING} = \alpha_0 + \alpha_1 \widehat{f}_{t,profesional} + \epsilon_t, \tag{27}$$

where $\widehat{f}_{t,Euro-STING}$ and $\widehat{f}_{t,professional}$ are the real-time forecasts of Euro-STING and of professional forecasters, respectively. Although we acknowledge the potential limitations of using our short sample, these p-values indicate that the forecasts of the Euro-STING model encompass the forecasts from the professional forecasters.

Finally, to evaluate further the results of our modeling effort, we include in the list of competitors a benchmark of non-forecastability and a standard AR(2). The relative MSEs of the Euro-STING with respect to these naïve benchmarks are 0.46 and 0.47 and the p-value of the null hypothesis of no difference in forecasting accuracy is 0.00 in both cases. Hence, the Euro-STING forecasts are more accurate than the average growth over the past sample and than standard autoregressive forecasts.

4 Conclusion

The monitoring of euro area economic developments in real time is a plentiful source of debate. How to deal with the lack of information hindering timely publication of macroeconomic variables, how to fill in missing values in time series, how to use euro area aggregates with short time spans, and how many variables should be included in the forecasting model are still open questions. Our paper contributes to the literature in this field by proposing a model which handles all these problems but is sufficiently tractable to allow economic analyses in real time. In particular, the model is based on the strict factor model used by Stock and Watson (1991), which is updated to deal with: a) ragged edges, as in Giannone, Reichlin and Small (2008) in order to allow for variable reporting lags; b) mixed frequencies, as in Mariano and Murasawa (2003) in order to bridge monthly with

quarterly GDP; c) data revisions for GDP growth, as in Evans (2005), which are modeled by assuming that the flash and preliminary estimates are equal to the true GDP plus uncorrelated noise.

With this simple medium-scale factor model, we show that our algorithm using a clear, easy-to-replicate methodology is able to forecast euro area GDP growth at least as well as professional forecasters. This is of prime importance since the latter not only employ the best forecasting tools, but also are allowed to incorporate their own judgement.

In addition, we propose several empirical contributions. First, we construct a new co-incident indicator of the euro area economy that is responsive to euro area business cycle dynamics. Second, we give some examples to illustrate our assertion that the analysis of forecasting accuracy in real time should rely on current-vintage data sets and not on end-of-sample vintage data sets, which may lead to unrealistic results. Third, we show that monthly indicators and flash announcements contain valuable information to reduce forecast uncertainty. Fourth, we find that once the timely publication of survey indicators takes place in short-term forecasts, business surveys gain importance with respect to economic activity data. Fifth, we propose a method to examine the empirical reliability of the assumption of uncorrelated idiosyncratic components. Finally, we propose a method of data selection in factor analyses.

We consider that the construction of a real-time database is a very useful contribution. The database contains 424 different vintages which include the information that was available to construct real-time forecasts each day during the last five years. We use this database to evaluate the forecasting accuracy of our model in a horse-race analysis against the main forecasts of the euro area GDP growth rate. The database may be very useful in empirical work on forecasting GDP growth in the euro area from factor models, since these proposals are usually compared with forecasts from simple models by using out-of-sample (latest available) data. This may give rise to misleading results.

In addition, the model used in this paper provides a solid foundation for (at least) two natural extensions. One of them has to do with the pre-seasonal adjustment of the series by Eurostat. The usefulness of extending the baseline model to handle non-seasonally adjusted series along the lines suggested by Harvey and Shephard (1993) is twofold. First, it

would allow researchers to examine how different procedures for handling seasonality may affect forecasting performance in real time. Second, it would provide a unified model for forecasting macroeconomic series in those countries that produce non-seasonally adjusted aggregates.

The other extension of the model has to do with anticipating changes in business cycle regimes. Dynamic factor models are probably the most appropriate framework in which to combine the two key features of the business cycle: the idea of comovements among macroeconomic aggregates and the dichotomy between expansions and recessions. Following the economic arguments put forward by Diebold and Rudebusch (1996), the extension seeks to unify the linear dynamic factor model proposed in this paper and the non-linear Markov-switching methodology.

Regarding the debate over forecasting euro area GDP growth with large-scale factor models (Barhoumi et al., 2008, Angelini et al., 2008) versus small-to-medium-scale factor models such as ours and Frale et al. (2008), we agree with the view of Aruoba, Diebold and Scotti (2009) that the matter is ultimately empirical. Although much professional attention has recently turned to "big data" approaches, Boivin and Ng (2006) and Bai and Ng (2008) have recently suggested that empirical forecast accuracy does not necessarily increase with the number of series included in the factor model. In addition, Banbura and Runstler (2007) found that most of the predictive content of their large-scale model is contained in a small set of variables. Hence, there is room for small-scale approaches and further empirical analyses should compare the performance of forecasts from these two approaches. In the context of forecasting euro area GDP growth, we find evidence in favor of using a small-scale model. We show that forecasting with a selected set of indicators does the job of forecasting relatively easier and entails no loss of valuable information or of accuracy of factor estimation.

To sum up, we think that the model presented in this paper, which describes and evaluates what we call Euro-STING forecasts, is a practical forecasting tool. It has a good forecasting record, is automatically updated when new information becomes available, provides a way of measuring the effects of new developments in GDP growth rate indicators and allows for extensions that could address in a common framework several problems,

such as seasonality and non-linearities, which historically have been analyzed separately from the forecasting exercise.

Appendix A

To illustrate what the matrices stated in the measurement and transition equations look like, let $0_{i,j}$ be a matrix of $(i \times j)$ zeroes, I_r be the r-dimensional identity matrix, and $\otimes$ be the Kronecker product. According to the empirical application, let us assume that $m_1 = m_2 = m_4 = 6$, $m_3 = 2$, $r_h = 4$, and $r_s = 5$. For simplicity, let us assume that all variables are always observed at a monthly frequency.

In this example, the measurement equation, $Y_t = Hh_t + w_t$, with $w_t \sim i.i.d.N(0, R)$, can be expressed as

$$Y_t = \left(\begin{array}{cccccc} y_t^{2nd} & Z_t^{h'} & Z_t^{s'} & l_t & y_t^{1st} & y_t^f \end{array} \right)', \tag{28}$$

$$w_t = 0_{1,r+4}, \tag{29}$$

$$R = 0_{r+4,r+4}, \tag{30}$$

$$h_t = (f_t, ..., f_{t-11}, u_{1t}, ..., u_{1t-5}, v_{1t}, v_{1t-1}, ..., v_{rt}, v_{rt-1}, u_{2t}, ..., u_{2t-5}, e_{1t}, e_{2t})'. \tag{31}$$

The matrix H is in this case

$$H = \begin{pmatrix} H_{11} & 0_{1,6} & H_{12} & 0_{1,8} & 0_{1,10} & 0_{1,6} & 0 & 0 \\ H_{21} & 0_{r_h,6} & 0_{r_h,6} & H_{22} & 0_{r_h,10} & 0_{r_h,6} & 0_{r_h,1} & 0_{r_h,1} \\ H_{31} & H_{31} & 0_{r_s,6} & 0_{r_s,8} & H_{32} & 0_{r_s,6} & 0_{r_s,1} & 0_{r_s,1} \\ H_4 & 0_{1,6} & 0_{1,6} & 0_{1,8} & 0_{1,10} & H_{12} & 0 & 0 \\ H_{11} & 0_{1,6} & H_{12} & 0_{1,8} & 0_{1,10} & 0_{1,6} & 0 & 1 \\ H_{11} & 0_{1,6} & H_{12} & 0_{1,8} & 0_{1,10} & 0_{1,6} & 1 & 1 \end{pmatrix}, \tag{32}$$

where

$$H_{11} = \left(\begin{array}{cccccc} \frac{\beta_1}{3} & \frac{2\beta_1}{3} & \beta_1 & \frac{\beta_1}{3} & \frac{2\beta_1}{3} & 0 \end{array} \right), \tag{33}$$

$$H_{12} = \left(\begin{array}{cccccc} \frac{1}{3} & \frac{2}{3} & 1 & \frac{1}{3} & \frac{2}{3} & 0 \end{array} \right), \tag{34}$$

$$H_{22} = I_{r_h} \otimes \left(\begin{array}{cc} 1 & 0 \end{array} \right), \tag{35}$$

$$H_{32} = I_{r_s} \otimes \left(\begin{array}{cc} 1 & 0 \end{array} \right), \tag{36}$$

$$H_4 = \left(\begin{array}{cccccc} \frac{\beta_4}{3} & \frac{2\beta_4}{3} & \beta_4 & \frac{\beta_4}{3} & \frac{2\beta_4}{3} & 0 \end{array} \right), \tag{37}$$

H_{21} is a $(r_h \times 6)$ matrix of zeroes whose first column is β_2, and H_{31} is a $(r_s \times 6)$ matrix whose columns are β_3.

Using the assumptions of the underlying example, the transition equation can be stated as follows. Let Q be a diagonal matrix in which the entries inside the main diagonal are determined by the vector

$$q = \left(\begin{array}{ccccccccccccc} \sigma_f^2 & 0_{1,11} & \sigma_{u_1}^2 & 0_{1,5} & \sigma_{v_1}^2 & 0 & \ldots & \sigma_{v_r}^2 & 0 & \sigma_{u_2}^2 & 0_{1,5} & \sigma_{e_1}^2 & \sigma_{e_2}^2 \end{array} \right)', \qquad (38)$$

The matrix F becomes

$$F = \begin{pmatrix} a & 0_{12,6} & 0_{12,8} & 0_{12,10} & 0_{12,6} & 0 & 0 \\ 0_{6,12} & b & 0_{6,8} & 0_{6,10} & 0_{6,6} & 0 & 0 \\ 0_{8,12} & 0_{8,6} & c_h & 0_{8,10} & 0_{8,6} & 0 & 0 \\ 0_{10,12} & 0_{10,6} & 0_{10,8} & c_s & 0_{10,6} & 0 & 0 \\ 0_{6,12} & 0_{6,6} & 0_{6,8} & 0_{6,10} & d & 0 & 0 \\ 0_{1,12} & 0_{1,6} & 0_{1,8} & 0_{1,10} & 0_{1,6} & 0 & 0 \\ 0_{1,12} & 0_{1,6} & 0_{1,8} & 0_{1,10} & 0_{1,6} & 0 & 0 \end{pmatrix}, \qquad (39)$$

where

$$a = \begin{pmatrix} a_1 & \ldots & a_6 & \ldots & 0 & 0 \\ 1 & \ldots & 0 & \ldots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots \\ 0 & \ldots & 0 & \ldots & 1 & 0 \end{pmatrix}, \qquad (40)$$

$$b = \begin{pmatrix} b_1 & \ldots & b_5 & b_6 \\ 1 & \ldots & 0 & 0 \\ \vdots & \ddots & \vdots & \vdots \\ 0 & \ldots & 1 & 0 \end{pmatrix}, \qquad (41)$$

$$c_i = \begin{pmatrix} c_{11} & c_{12} & \ldots & 0 & 0 \\ 1 & 0 & \ldots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \ldots & c_{r1} & c_{r2} \\ 0 & 0 & \ldots & 1 & 0 \end{pmatrix}, \qquad (42)$$

34

$$d = \begin{pmatrix} d_1 & ... & d_5 & d_6 \\ 1 & ... & 0 & 0 \\ \vdots & \ddots & \vdots & \vdots \\ 0 & ... & 1 & 0 \end{pmatrix}. \tag{43}$$

Appendix B

All the indicators used in the forecasting analysis can be found at the following links:

1. EuroCoin:

   http://www.cepr.org/data/eurocoin/

2. DG_ECFIN:

   http://ec.europa.eu/economy_finance/indicators/euroareagdp_en.htm

3. EC_Macroeconomic_Forecast:

   http://ec.europa.eu/economy_finance/about/activities/activities_keyindicatorsforecasts_en.htm

4. IFO_INSEE_ISAE:

   http://www.cesifo-group.de/portal/page/portal/ifoHome/a-winfo/d2kprog/30kprogeeo

5. OCDE:

   http://www.oecd.org/department/0,3355,en_2649_34109_1_1_1_1_1,00.html

References

[1] Altavilla C. and Ciccarelli M. 2007. Information Combination and Forecast (St)Ability: Evidence from Vintages of Time-Series Data. ECB working paper No. 846.

[2] Altissimo, F., Cristadoro, R., Forni, M., Lippi, M. and Veronese, G. 2006. New Eurocoin: Tracking economic growth in real time. CEPR discussion paper No. 5633.

[3] Angelini, E., Camba-Mendez, G., Giannone, D., Reichlin, L., and Runstler, G. 2008. Short-term forecasts of Euro area GDP growth. CEPR discussion paper No. 6746.

[4] Aruoba, B. 2008. Data revisions are not well behaved. *Journal of Money, Credit and Banking* 40: 319-340.

[5] Aruoba, B., Diebold, F., and Scotti, C. 2009. Real-time measurement of business conditions. *Journal of Business and Economic Statistics*, forthcoming.

[6] Bai, J., and Ng, S. 2008. Forecasting economic time series using targeted predictors. *Journal of Econometrics* 148: 304-317.

[7] Barhoumi, K., Benk, S., Cristadoro, R., Reijer, A., Jakaitiene, A., Jelonek, P., Rua, A. Rünstler, G., Ruth K, Van Nieuwenhuyze Ch. 2008. Short-term forecasting of GDP using large monthly datasets a pseudo real-time forecast evaluation exercise. Occasional Paper Series 84, European Central Bank.

[8] Banbura, M., and Runstler, G. 2007. A look into the model factor black box. Publication lags and the role of hard and soft data in forecasting GDP. ECB working paper No. 751.

[9] Boivin, J., and Ng, S. 2006. Are more data always better for factor analysis? *Journal of Econometrics* 132: 169-194.

[10] Coenen, C., Levin, A., and Wieland, V. 2005. Data uncertainty and the role of money as an information variable for monetary policy. *European Economic Review* 49: 975-1006.

[11] Cubadda, G. and Heck. A. 2008. The best business cycle indicator: A medium dimension approach with application to EU countries. University of Rome, Mimeo.

[12] Croushore, D., and Stark, T. 2001. A real time data set for macroeconomists. *Journal of Econometrics* 105: 111-130.

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

[14] Diebold, F., and Rudebusch, A. 1991. Forecasting output with the composite leading index: a real-time analysis. *Journal of the American Statistical Association* 86: 603–610.

[15] Diebold, F., and Rudebusch , G. 1996. Measuring business cycles: A modern perspective. *Review of Economics and Statistics* 78, 67-77.

[16] Diron, M. 2006. Short-term forecasts on Euro area real GDP growth. An assessment of real-time performance based on vintage data. ECB working paper No. 622.

[17] Engle R. 2007. High dimension dynamic correlations. NYU working paper FIN-07-045.

[18] European Commission. 2006. The joint harmonized EU programme of business and consumer surveys. Special Report No. 5/2006.

[19] Evans, M. 2005. Where are we now? Real time estimates of the macroeconomy. *International Journal of Central Banking* 1: 127-175.

[20] Frale, C., Marcellino, M., Mazzi GL and Proietti, T. 2008a. A Monthly Indicator of the Euroa Area GDP. Mimeo.

[21] Frale, C., Marcellino, M., Mazzi GL and Proietti, T. 2008b. Survey Data as Coincident or Leading Indicators. Presented in the 5th Colloquium on Modern Tools for Business Cycle Analysis.

[22] Giannone, D., Reichlin, L., and Small, D. 2006. Nowcasting GDP and inflation: the real-time informational content of macroeconomic data releases. ECB working paper No. 633.

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

[24] Harvey, A., and Shephard, N. 1993. Structural Time Series Models. In G S Maddala et. al. Handbook of Statistics Volume 11 Elsevier Science Publishers.

[25] Inklaar, R., Jacobs, J., Romp, W. 2004. Business cycle indexes: Does a heap of data help?. *Journal of Business Cycle Measurement and Analysis* 1 (3): 309-336.

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

[27] Orphanides, A., and van Norden, S. (2002). The unreliability of output-gap estimates in real time. *The Review of Economics and Statistics* 4: 569-583.

[28] Proietti, T., and Moauro, F. 2006. Dynamic factor analysis with nonlinear temporal aggregation constraints. *Applied Statistics* 55: 281-300.

[29] Runstler, G., and Sedillot, F. 2003. Short-term estimates of Euro area real GDP by means of monthly data. ECB working paper No. 276.

[30] Stark, T., and Croushore, D. 2002. Forecasting with a real-time data set for macroeconomists. *Journal of Macroeconomics* 24: 507-531.

[31] Stock, J., and Watson, M. 1991. A probability model of the coincident economic indicators. In *Leading Economic Indicators: New Approaches and Forecasting Records*, edited by K. Lahiri and G. Moore, Cambridge University Press.

[32] Swanson, N., and van Dijk, D. 2006. Are statistical reporting agencies getting it right? Data rationality and business cycle asymmetry. *Journal of Business and Economic Statistics* 24, 24-42.

Table 1. Data description

Euro area Indicator Variables (a) (b)

	Name	Definition	Observations	Reporting lag (c)
Quarterly hard indicators	Flash GDP	Euro area GDP	19	45 days
	First GDP	Euro area GDP	37	60 days
	Second GDP	Euro area GDP	66	102 days
	Employment	Euro area Total Employment	66	102 days (d)
Monthly hard indicators	IPI	Euro area Industrial Production Index (ex. construction)	200	42 days
	Sales	Euro area Total Retail Sales Volume	155	35 days
	INO	Industrial New Orders Index. Total manuf. work on orders	154	52 days
	Exports	Extra- Euro area Exports	200	45 days
Monthly soft indicators	BNB	Belgium Overall Business Indicator	202	-8 days
	ESI	Euro-area Economic Sentiment Indicator	202	0 days
	IFO	Germany IFO Business Climate Index	202	-8 days
	PMI Manuf.	Euro area Manufacturing Purchasing Managers Index	128	1 day
	PMI Services	Euro area Services Purchasing Managers Index	115	5 days

Notes:

(a) All hard indicators (indicators of real activity) are growth rates of the seasonally adjusted series.
Soft indicators (based on opinion surveys) are seasonally adjusted series in levels.

(b) Euro area refers to EMU-12 until December 2006 and EMU-13 (includes Slovenia) after that date.

(c) Approximately. It can change slightly due to weekends or the number of days of the month.

(d) Starting in 2007.1 the reporting lag is 45 days

Table 2. Data set available on 02/11/08

	Second	IPI	Sales	INO	Export	ESI	BNB	IFO	PMIM	PMIS	Employment	First	Flash
2007.06	0.31	0.01	0.67	4.41	2.52	111.10	5.50	107.00	55.56	58.33	0.58	0.35	0.34
2007.07	na	0.65	0.35	-3.09	-0.57	110.40	4.20	106.40	54.90	58.34	na	na	na
2007.08	na	1.15	-0.03	0.92	2.73	109.40	3.30	105.70	54.34	58.04	na	na	na
2007.09	0.76	-0.86	0.17	-1.17	-1.35	106.30	1.50	104.10	53.21	54.22	0.33	0.71	0.71
2007.10	na	0.54	-0.64	2.54	1.22	105.40	-0.10	103.90	51.52	55.81	na	na	na
2007.11	na	-0.45	-0.66	2.72	0.26	104.10	1.40	104.20	52.80	54.14	na	na	na
2007.12	na	na	-0.09	na	na	103.40	-1.90	103.00	52.56	53.14	na	na	na
2008.01	na	na	na	na	na	101.70	-0.80	103.40	52.77	50.56	na	na	na
2008.02	na	na	na	na	na	na	na	na	na	na	na	na	na
2008.03	na	na	na	na	na	na	na	na	na	na	na	na	na
2008.04	na	na	na	na	na	na	na	na	na	na	na	na	na
2008.05	na	na	na	na	na	na	na	na	na	na	na	na	na
2008.06	na	na	na	na	na	na	na	na	na	na	na	na	na

Notes. See Table 1 for acronyms. Figures labelled as "na" refer to either missing data or data that are not available on 02/11/08.

Table 3. Factor loadings

Second	IPI	Sales	INO	Exports	ESI	BNB	IFO	PMIM	PMIS	Employment
0.15	0.21	0.06	0.19	0.12	0.05	0.06	0.05	0.07	0.07	0.10
(0.03)	(0.04)	(0.03)	(0.04)	(0.01)	(0.01)	(0.01)	(0.01)	(0.04)	(0.02)	(0.04)

Notes. See Table 1 for acronyms. Standard errors are in parentheses. Data set ends on 02/11/08.

Table 4. Last day forecast (February 11th 2008)

Panel A				Panel B	
Series	2007.4	2008.1	2008.2	Series	Next month
				IPI	0.41
FLASH	0.40	0.38	0.38	Retail Sales	0.247
	(0.05)	(0.06)	(0.07)	INO	-1.825
				Exports	0.716
FIRST	0.39	0.36	0.36	ESI	100.712
	(0.07)	(0.07)	(0.08)	BNB	-2.978
				IFO	102.565
SECOND	0.41	0.37	0.37	PMI Man	52.567
	(0.10)	(0.11)	(0.14)	PMI Serv	50.844
				Employment	0.194

Notes. See Table 1 for acronyms. Standard errors are in parentheses. Data set ends on 02/11/08.

Table 5. Cumulative weights

	Second	IPI	Sales	INO	Exports	ESI	BNB	IFO	PMIM	PMIS	Employment	First	Flash
2007.06	1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
2007.07	0.00	0.32	0.06	0.29	0.09	0.04	0.04	0.02	0.09	0.06	0.00	0.00	0.00
2007.08	0.00	0.32	0.06	0.29	0.09	0.04	0.04	0.02	0.09	0.06	0.00	0.00	0.00
2007.09	1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
2007.10	0.00	0.32	0.06	0.29	0.09	0.04	0.04	0.02	0.09	0.06	0.00	0.00	0.00
2007.11	0.00	0.32	0.06	0.29	0.09	0.04	0.04	0.02	0.09	0.06	0.00	0.00	0.00
2007.12	0.00	0.00	0.19	0.00	0.00	0.13	0.12	0.05	0.30	0.21	0.00	0.00	0.00
2008.01	0.00	0.00	0.00	0.00	0.00	0.15	0.15	0.06	0.37	0.27	0.00	0.00	0.00
2008.02	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
2008.03	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
2008.04	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
2008.05	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
2008.06	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00

Notes. See Table 1 for acronyms. Data set ends on 02/11/08.

Table 6. Accuracy of preliminary announcements

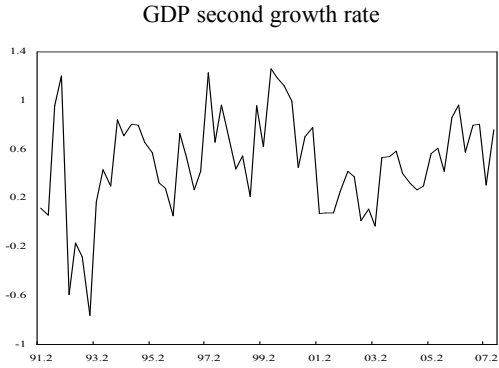| Indicator	Day of forecast	MSE
Flash estimator		0.024
First Estimator		0.025
Euro-STING	Day before Flash	0.027
Euro-STING	Day after Flash	0.022
Euro-STING	Day before First	0.022
Euro-STING	Day after First	0.014

Note. Entries are mean squared errors in forecasting the last revised values of second GDP (vintage 01/09/08). The first two rows refer to preliminary announcements and last four rows are forecasts from the Euro-STING model which are computed the days before and after flash and first releases.

Table 7. Real-time forecasting evaluation

	1 Month lag	2 Month lag	3 Month lag
Eurocoin	0.047	0.046	0.042
Euro-STING	0.042	0.023	0.016
MDM	0.773	0.089	0.037
Encompass	0.323	0.055	0.126
	1 Month lag	2 Month lead	5 Month lead
IFO-INSEE-ISAE	0.060	0.071	0.107
Euro-STING	0.044	0.048	0.037
MDM	0.430	0.279	0.177
Encompass	0.437	0.450	0.491
	3 Month lag	6 Month lead	9 Month lead
European Commission	0.055	0.086	0.068
Euro-Sting	0.028	0.071	0.033
MDM	0.131	0.817	0.206
Encompass	0.588	0.751	0.936
	3 Month lag	0 Month	3 Month lead
OECD	0.019	0.054	0.056
Euro-Sting	0.019	0.052	0.043
MDM	0.999	0.215	0.433
Encompass	0.789	0.067	0.182
	1 Month Lag	2 Month lead	5 Month lead
DG ECFIN	0.045	0.047	0.052
Euro-Sting	0.046	0.035	0.036
MDM	0.933	0.503	0.348
Encompass	0.153	0.298	0.159

Notes. Entries are mean squared errors. Forecasting period: 2003.4 to 2007.3. Last column is the simple average. See Appendix B for data description.

Figure 1. Time series used in the model

GDP second growth rate

Sample 91.2-07.3. Vintage 01/09/08

GDP first growth rate
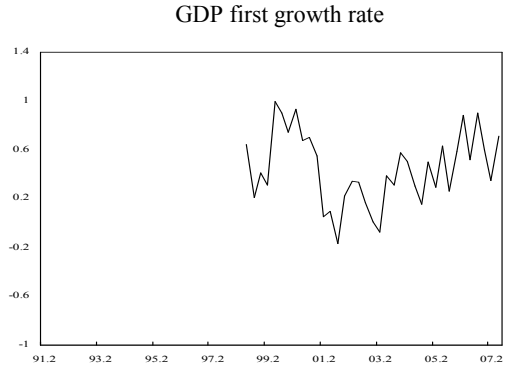
Sample 98.2-07.3. Vintage 11/30/07

GDP flash growth rate
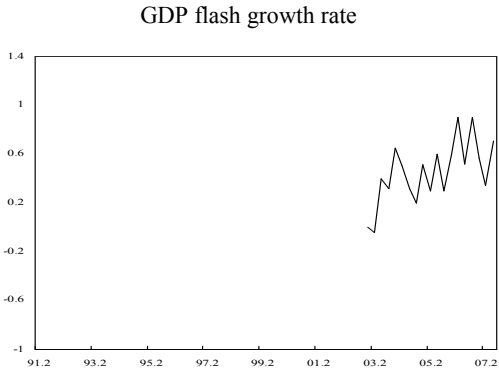
Sample 03.1-07.3. Vintage 11/14/07

Industrial production
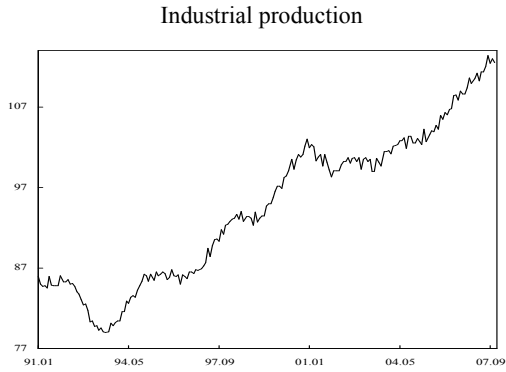
Sample 91.01-07.11. Vintage 01/14/08

Retail Sales

Sample 95.01-07.12. Vintage 02/05/08

Industrial new orders

Sample 95.01-07.11. Vintage 01/23/08

42

Figure 1. Time series used in the model (continued)

Exports
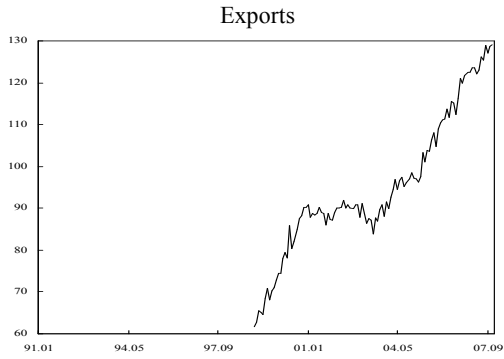
Sample 99.01-07.11. Vintage 01/17/08

Employment
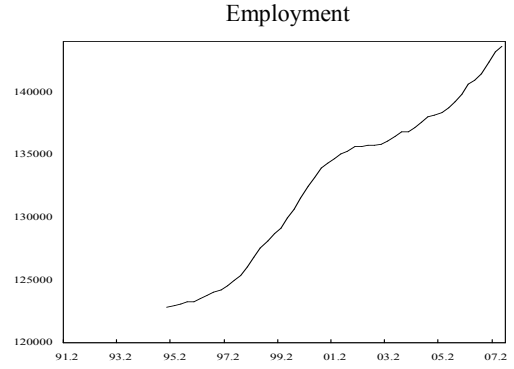
Sample 95.1-07.3. Vintage 01/09/08
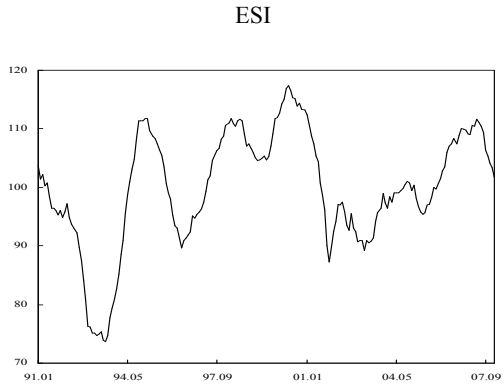
ESI

Sample 91.01-08.01. Vintage 01/31/08
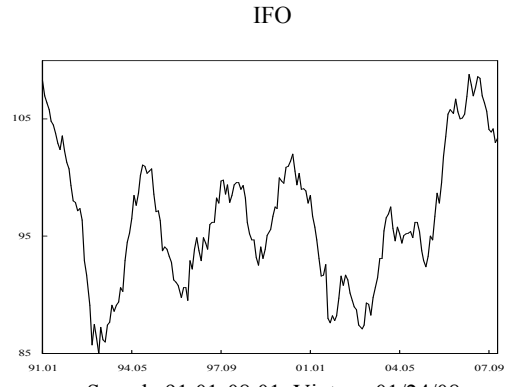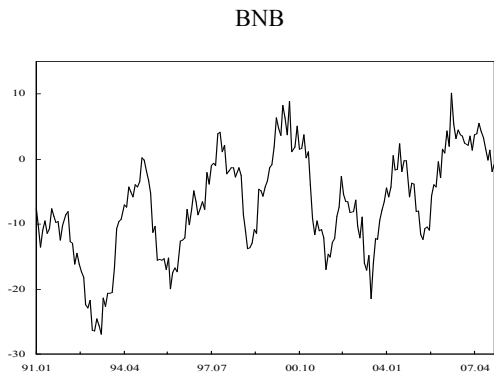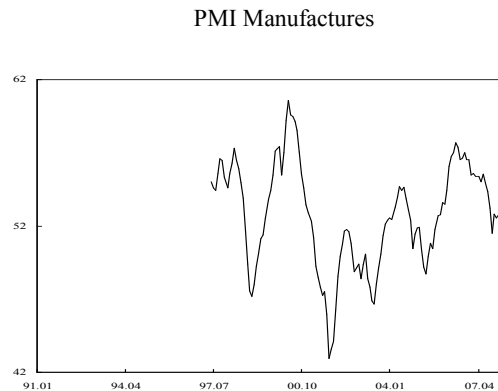
IFO

Sample 91.01-08.01. Vintage 01/24/08

BNB

Sample 91.01-08.01. Vintage 01/24/08

PMI Manufactures

Sample 97.06-08.01. Vintage 02/01/08

43

Figure 1. Time series used in the model (continued)

PMI Services

Sample 97.06-08.01. Vintage 02/05/08
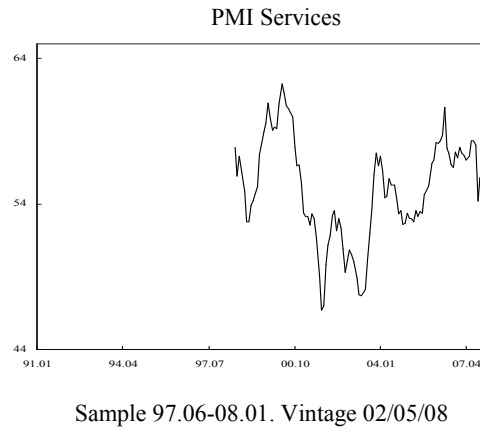
Notes. See Table 1 for acronyms. Charts refer to data available on 02/11/08.

Figure 2. Common factor and Eurocoin
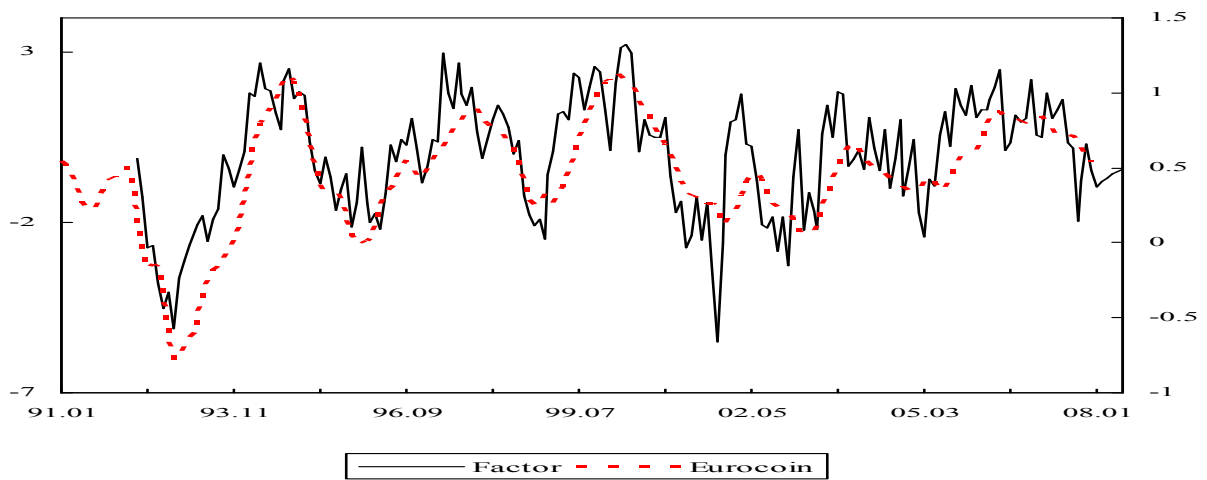
Notes. The factor is estimated from 92.04 to 08.06 with information on 02/11/08. The last vintage for the Eurocoin contains data until 07.12.

Figure 3. GDP second growth rate: actual and estimates
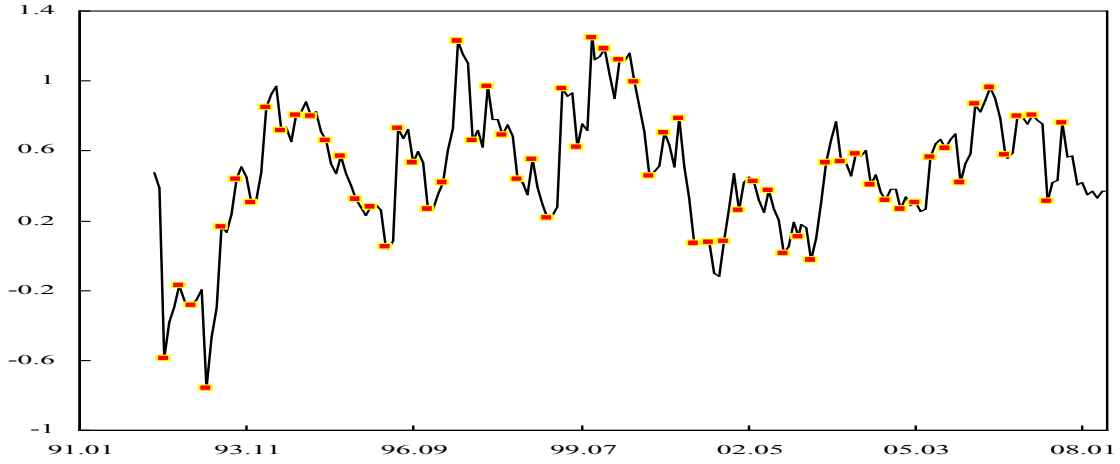
Notes. GDP growth rates are estimated from 1992.04 to 2008.06 with information on 02/11/08. Plotmarks on the line refer to actual data (third month of each quarter; last one in 2007.3).

Figure 4. GDP forecast in 2008.1 and ESI potential releases
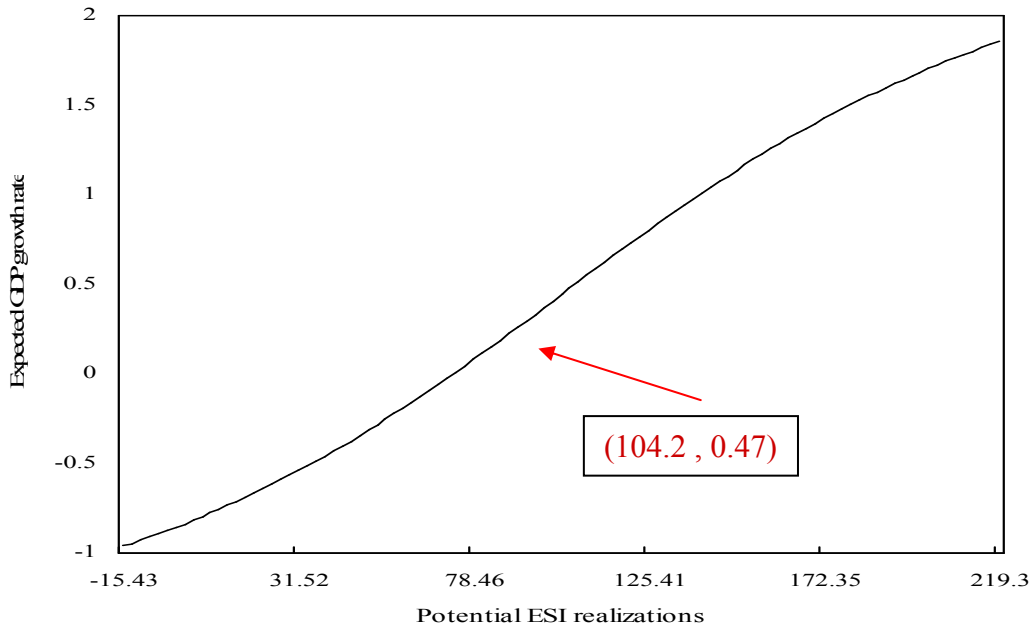
Notes. Expected ESI values and growth rates before the last ESI updates are inside the box. Actual ESI was 104.2, which corresponds an expected growth rate of 0.47.

45

Figure 5. GDP second growth rate in real time

GDP second forecasts for 2007.4

GDP second forecasts for 2006.3
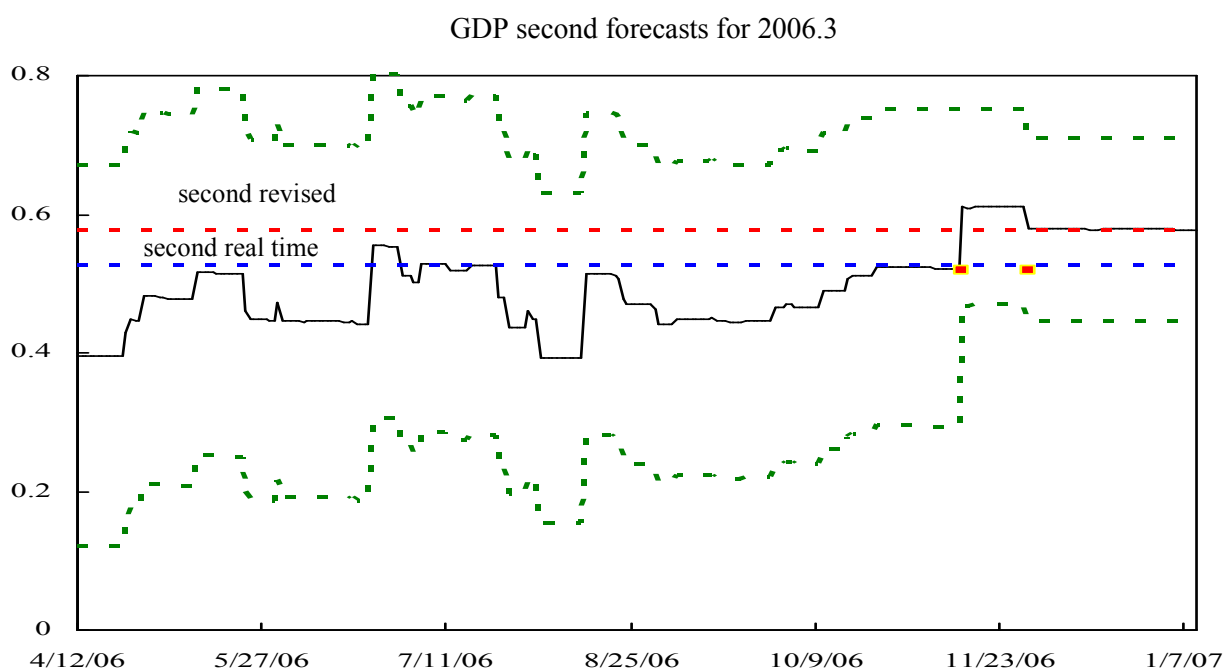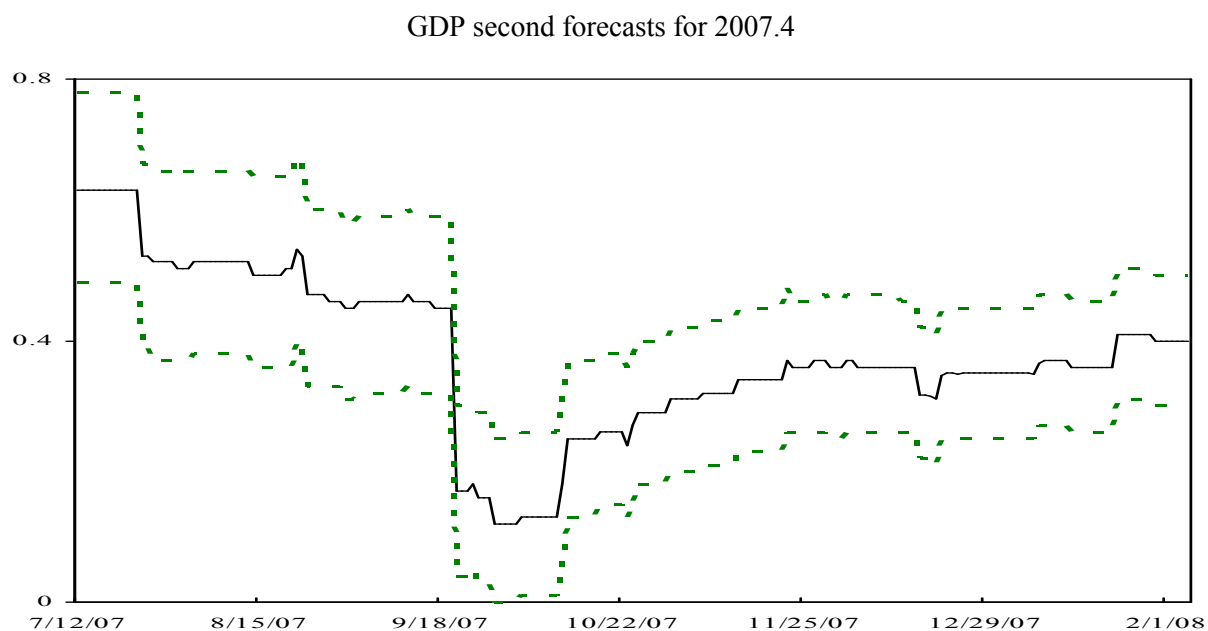
Notes. Floating points refer to flash and first announcements. They are calculated for the nine-month forecasting periods described in the text.

46

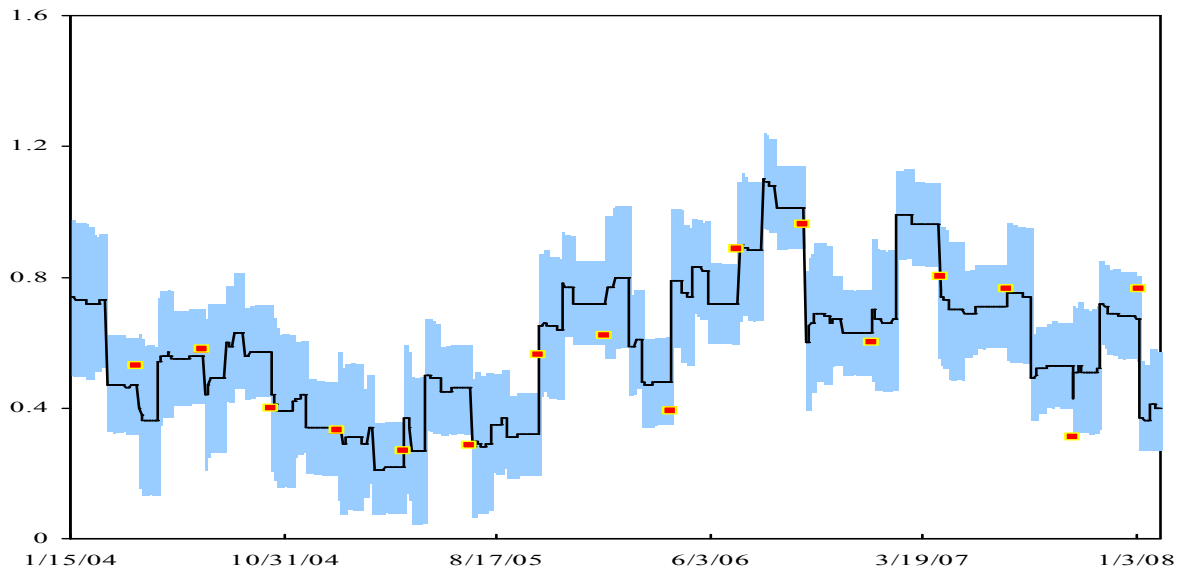Figure 6. Averaged standard errors over the days of the forecasting period
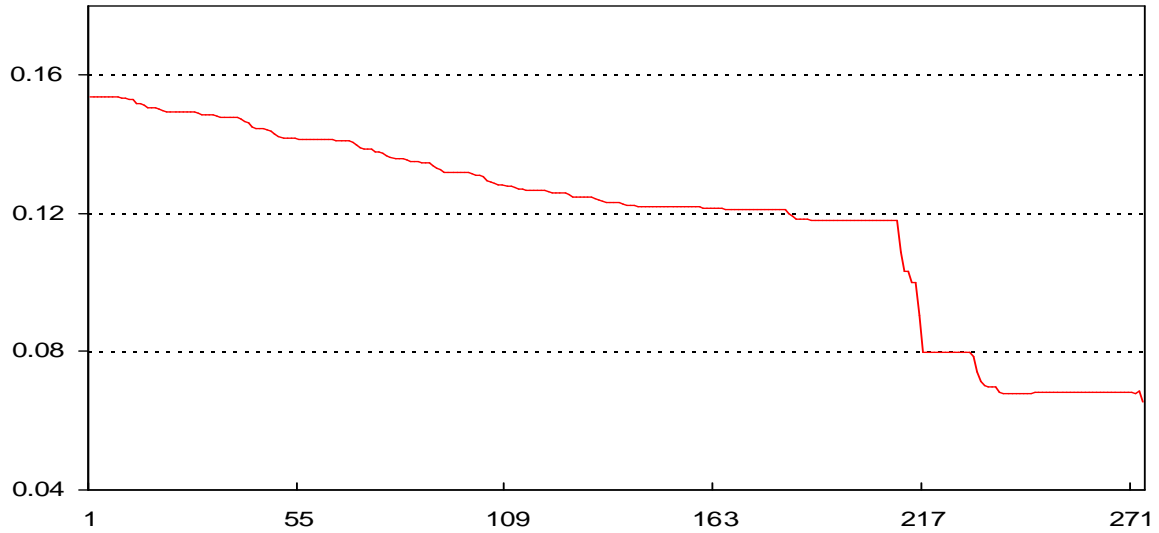
Notes. Real time sample covers from 01/15/04 to 02/11/08.

Figure 7. Cumulative forecast weights for BNB in real-time
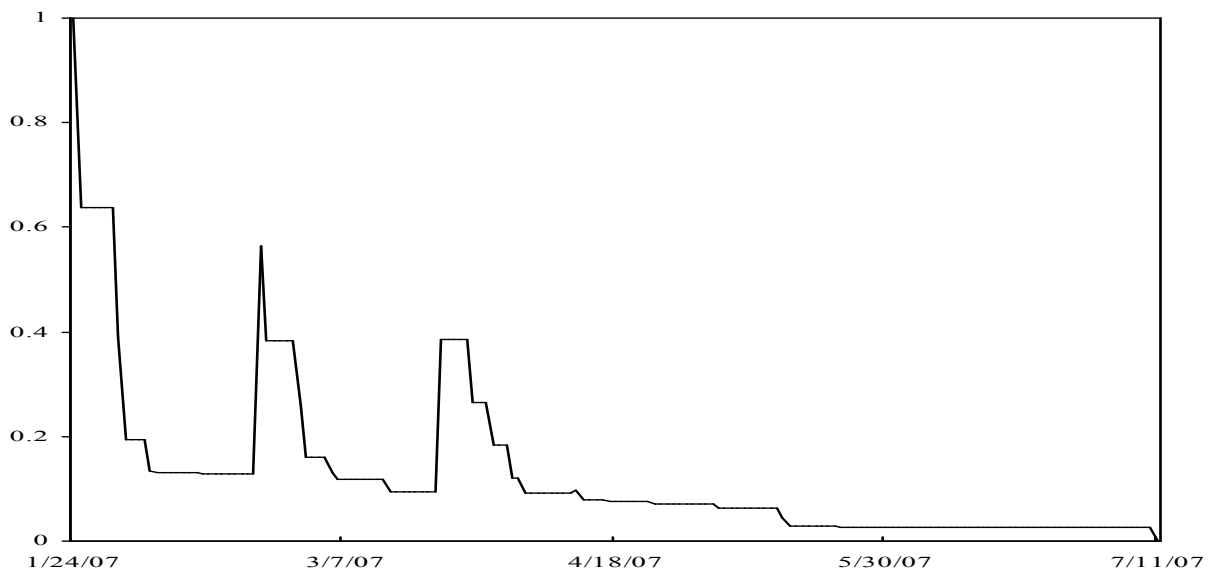
Notes. Weights are computed daily from 01/24/07 to 07/12/07 and refer to GDP forecasts for the first quarter of 2007.

47

Figure 8. Real-time lagged forecasts of GDP

Notes. Real time sample covers from 01/15/04 to 02/11/08. Shaded area is the two standard error band. Floating points refer to revised values of second GDP (Vintage 01/09/08) which are dated their real-time announcement.