LIETUVOS BANKAS

WORKING PAPER SERIES

No 1 / 2008

SHORT-TERM FORECASTING OF GDP USING LARGE MONTHLY DATA SETS: A PSEUDO REAL-TIME FORECAST EVALUATION EXERCISE

by G. Rünstler, K. Barhoumi, S. Benk, R. Cristadoro, A. Den Reijer, A. Jakaitienė, P. Jelonek, A. Rua, K. Ruth, C. Van Nieuwenhuyze

WORKING PAPER SERIES
No 1 / 2008

SHORT-TERM FORECASTING OF GDP USING LARGE MONTHLY DATA SETS: A PSEUDO REAL-TIME FORECAST EVALUATION EXERCISE*

by G. Rünstler, K. Barhoumi, S. Benk, R. Cristadoro, A. Den Reijer, A. Jakaitienė, P. Jelonek, A. Rua, K. Ruth, C. Van Nieuwenhuyze**

*The project described in this paper was conducted under the auspices of the Eurosystem working groups on Econometric Modelling and on Forecasting. The authors would like to thank the members of the two working groups, in particular L. Reichlin, A. d'Agostino, D. Giannone and C. Schumacher, for their helpful comments.

**G. Rünstler (European Central Bank), K. Barhoumi (Banque de France), R. Cristadoro (Banca d'Italia), S. Benk (Magyar Nemzeti Bank), A. Den Reijer (De Nederlandsche Bank), A. Jakaitienė (Lietuvos Bankas), P. Jelonek (Narodowy Bank Polski), A. Rua (Banco de Portugal), K. Ruth (Deutsche Bundesbank), C. Van Nieuwenhuyze (National Bank of Belgium).

Address
Totorių 4
LT-01121 Vilnius
Lithuania
Telephone +370 5 2680132
Fax +370 5 2124423

Internet
http://www.lb.lt

Statement of purpose
Working Papers describe research in progress by the author(s) and
are published to stimulate discussion and critical comments.

The Series is managed by Economic Research Division of Economics
Department.

All Working Papers of the Series are refereed by internal and external experts.

ISSN 2029-0446 (ONLINE)

Contents

ABSTRACT **4**
NON-TECHNICAL SUMMARY **5**
1. INTRODUCTION **6**
2. MODELS **8**
2. 1 QUARTERLY MODELS 8
2.1.1 Recursive mean and quarterly autoregressive model (AR) *8*
2.1.2 Quarterly vector autoregressive models (VAR) – forecast averages *8*
2.2 BRIDGING MONTHLY DATA WITH QUARTERLY GDP 9
2.2.1 Bridge equations (BE) – forecast average across indicators *9*
2.2.2 Bridging with factors *10*
2.2.3 Generalised principal components *11*
3. PSEUDO REAL-TIME FORECAST DESIGN **12**
3. 1. FORECAST DESIGN 12
4. DATA **13**
5. RESULTS **14**
5.1 FORECAST ACCURACY 14
5.2 ENCOMPASSING TESTS 17
6. CONCLUSIONS **18**
REFERENCES **20**
TABLES **23**

Abstract

This paper evaluates different models for the short-term forecasting of real GDP growth in ten selected European countries and the euro area as a whole. Purely quarterly models are compared with models designed to exploit early releases of monthly indicators for the nowcast and forecast of quarterly GDP growth. Amongst the latter, we consider small bridge equations and forecast equations in which the bridging between monthly and quarterly data is achieved through a regression on factors extracted from large monthly datasets. The forecasting exercise is performed in a simulated real-time context, which takes account of publication lags in the individual series. In general, we find that models that exploit monthly information outperform models that use purely quarterly data and, amongst the former, factor models perform best.

Keywords: Bridge models, Dynamic factor models, real-time data flow.
JEL classification: E37, C53.

Santrauka

Šiame straipsnyje lyginami realiojo bendrojo vidaus produkto (BVP) augimo trumpojo laikotarpio prognozavimo modeliai, taikomi dešimčiai Europos šalių ir visai euro zonai. BVP augimo einamojo ir trumpojo laikotarpio prognozavimo modeliai, kuriems naudojami tik ketvirtiniai duomenis, lyginami su modeliais, sudarytais naudojant išankstinę mėnesinių rodiklių informaciją. Pastarieji modeliai – tai kelių kintamųjų jungiamosios lygtys (*bridge equations*) ir prognozei taikomos lygtys, kuriomis mėnesiniai ir ketvirtiniai duomenys susiejami į regresiją įtraukiant faktorius, apskaičiuotus iš didelės mėnesinių duomenų matricos. Prognozavimas atliekamas imituojant realųjį laiką, kai atsižvelgiama į atskirų rodiklių paskelbimo datą. Pagrindinė išvada yra ta, kad modeliai, pagrįsti mėnesiniais duomenimis, yra tikslesni negu modeliai, kuriems naudojama ketvirčių informacija. Iš pastarųjų tiksliausi yra faktoriniai modeliai.

Non-technical summary

Official estimates of GDP growth are released with a considerable delay. For the euro area as a whole, the first official number is a flash estimate, which is published six weeks after the end of the quarter. Meanwhile, economic analysis must rely on monthly indicators which arrive within the quarter such as, e.g. industrial production, retail sales and trade, surveys, and monetary and financial data.

This paper performs a forecasting evaluation of models used in central banks for computing early estimates of current quarter GDP and short-term forecasts of next-quarter GDP. The models are designed to “bridge” early releases of monthly indicators with quarterly GDP. In providing the starting point for a longer-term analysis, the assessment of the current state of the economy is certainly an important element in macro-economic forecasting.

The paper considers a range of models for this purpose, including traditional bridge equations and dynamic factor models. The key features of the evaluation study presented in this paper are as follows. First, we examine the forecast performance under the real-time flow of data releases, taking account of the non-synchronous release of monthly information throughout the quarter. Second, we use ten large datasets. In addition to the euro area as whole we consider datasets from six euro area Member States and three new Member States of the European Union. Third, we examine a wider range of models than previous studies and consider, beside euro aggregate data, individual country datasets.

The main finding obtained for the euro area countries is that bridge models, which timely exploit monthly releases, fare considerably better than quarterly models. Amongst those, dynamic factor models, which exploit a large number of releases, do generally better than averages of traditional bridge equations. Results for the new Member States, on the other hand, are difficult to interpret. All models perform quite badly with respect to naïve benchmarks, but, given the short evaluation sample, it is hard to understand what drives the results.

1. Introduction

This paper performs a forecasting evaluation of models used in central banks for computing early estimates of current quarter GDP and short-term forecasts of next-quarter GDP. These models are designed to "bridge" early releases of monthly indicators with quarterly GDP. Official estimates of GDP growth are released with a considerable delay. For the euro area as a whole, the first official number is a flash estimate, which is published six weeks after the end of the quarter. Meanwhile, economic analysis must rely on monthly indicators which arrive within the quarter such as, e.g. industrial production, retail sales and trade, surveys, and monetary and financial data.

In providing the starting point for a longer-term analysis, the assessment of the current state of the economy is certainly an important element in macroeconomic forecasting. This holds even more so as the longer-term predictability of quarterly GDP growth has declined since the 1990s (D'Agostino, Giannone and Surico, 2006).

A key feature of this paper is that we examine the forecast performance taking into account the real-time data flow, that is, the non-synchronous release of monthly information throughout the quarter. To this end, we replicate the design of the forecast exercise proposed by Rünstler and Sédillot (2003) for the euro area and by Giannone, Reichlin and Sala (2004) and Giannone, Reichlin and Small (2005) for the United States, which has also been applied for euro area aggregate data by Angelini et al. (2008a) and Angelini, Banbura and Rünstler (2008b). We examine a wider range of models than previous studies and consider, beside euro aggregate data, individual country datasets.

Macroeconomic indicators are subject to important differences in publication lags. Monthly industrial production data, for instance, are released about six weeks after the end of the respective month for the euro area, while survey and financial data are available right at the end of the month. Our forecast evaluation exercise is designed to replicate the data availability situation that is faced in real-time application of the models. In addition, the models are re-estimated only from the information available at the time of the forecast. However, our design differs from a perfect real-time evaluation insofar as we use final data vintages and hence ignore revisions to earlier data releases.

In order to understand the importance of timely monthly information, the paper considers purely quarterly models and bridge equations developed to link monthly releases with quarterly GDP growth. Bridge equations are used by many institutions and have been studied in various papers (Baffigi, Golinelli and Parigi, 2004; Diron, 2006; Rünstler and Sédillot, 2003).

Traditional bridge equations can only handle few variables. To exploit information in the releases of several indicators, the standard approach is to average equations using different regressors. Recently, Giannone, Reichlin and Sala (2004) and Giannone, Reichlin and Small (2008) have proposed to use factors extracted from large monthly datasets to perform bridging which exploit a large number of indicators within the same model (bridging with factors). They propose to use the Kalman filter to estimate the factors and handle missing data.[1] When bridging with factors, however, one can consider alternative estimation methods for the factors than that based on the Kalman filter. Methods that have been used in the Eurosystem include the principal component estimator of the factors (Stock and Watson, 2002b) and the frequency domain-based two-step estimator of Forni et al. (2005). It is therefore natural for this study to consider these estimators in the bridging with factors framework. However, these methods have to be complemented with some tool to handle missing data. We will fill the missing data of each series on the basis of univariate forecasts following common practice with bridge equations.

It is important to stress that while there are several studies that apply factor models for forecasting euro area data (Marcellino et al. (2003) for euro area data, Artis et al. (2005) for the United Kingdom, Bruneau et al. (2007) for France, Den Reijer (2007) for the Netherlands, Duarte and Rua (2007) for Portugal, Schumacher (2007) for Germany, and Van Nieuwenhuyze (2005) for Belgium, among others), this paper considers the bridge version of these models which is appropriate for real-time short-term forecasting and can be meaningfully compared with traditional bridge equations.

Our model comparison is performed for the euro area as a whole as well as for six euro area countries. Moreover, we also assess the above-mentioned models for three new members of the European Union. We end up with ten large monthly datasets, with an average dimension of more than one hundred series for each country. Hence, we provide some cross-country evidence regarding the relative performance of the different models considered.

The paper is organised as follows. Section 2 presents the models that we consider in our exercise. Section 3 discusses the pseudo real-time forecast design, while section 4 presents the data. In section 5 the empirical results are discussed. Finally, section 6 concludes.

---

1 Beside the US and euro area applications cited above, the method is also used at Norges Bank (Aastveit and Trovik, 2007) and the Reserve Bank of New Zealand (Matheson, 2007).

2. Models

This section describes several models that may be used for forecasting GDP growth in the presence of large datasets. We consider models that rely solely on quarterly data as well as models that exploit the monthly nature of the available data with models ranging from the simple autoregressive process to the more sophisticated dynamic factor models proposed in the literature.

2. 1 QUARTERLY MODELS

2.1.1 Recursive mean and quarterly autoregressive model (AR)

As benchmarks we use two univariate time series models for quarterly GDP growth y_t^Q, i.e.

a) average GDP growth, i.e. the naïve model $y_t^Q = \mu + \varepsilon_t^Q$, and
b) a first-order autoregressive model,

$$y_t^Q - \mu = \rho(y_{t-1}^Q - \mu) + \varepsilon_t^Q, \qquad (1)$$

where μ is a constant and ε_t^Q is quarterly white noise, $\varepsilon_t^Q \sim N(0, \sigma_\varepsilon^2)$.

The forecasting performance of these two models will serve as a reference point in forecast evaluation. Given the differences in the statistical properties of GDP growth across countries, absolute measures of forecast performance are of limited use. We use the performance relative to the above models instead.

2.1.2 Quarterly vector autoregressive models (VAR) – forecast averages

Another forecast that uses purely quarterly data can be obtained from vector autoregressive models. This approach has been reported to perform well, for example, for the United Kingdom (see Camba-Mendez et al., 2001). We run bivariate VARs including quarterly GDP and the quarterly aggregate of a single monthly indicator, and average the forecasts across indicators.

1. We consider a set of *k* monthly indicators from the dataset and calculate their quarterly aggregates $\{x_{1,t}^Q, x_{2,t}^Q, \ldots, x_{k,t}^Q\}$.
2. For each indicator $x_{i,t}^Q$, we run a quarterly bivariate VAR, which includes the indicator and GDP growth,

$$z_{i,t}^Q = \mu_i + \sum_{s=1}^{p_i} A_s z_{i,t-s}^Q + \varepsilon_{i.t}^Q, \quad i = 1, \ldots, k, \qquad (2)$$

with $z_{i,t}^Q = (y_t^Q, x_{i,t}^Q)$; from this VAR, we produce forecasts $y_{i,t+h|t}^Q$ of GDP growth. The lag length (p_i) of each VAR is determined from the Schwartz information criterion (SIC).

3. We form the average of the k forecasts $y^Q_{i,t+h|t}$ from the individual indicators,

$$y^Q_{t+h|t} = k^{-1}\sum_{i=1}^{k} y^Q_{i,t+h|t} \ .$$

These forecasting methods do not exploit early monthly releases and hence they do not deal with ragged edges due to the non-synchronous flow of data releases.

2.2 BRIDGING MONTHLY DATA WITH QUARTERLY GDP

2.2.1 Bridge equations (BE) – forecast average across indicators

Bridge equations are a widely used method to forecast quarterly GDP from monthly data (see, for example, Baffigi, Golinelli and Parigi, 2004). Two steps are involved: (i) the monthly indicators are forecast over the horizon; (ii) the quarterly aggregates of the obtained forecasts are used to predict GDP growth. In averaging across a large number of indicators we follow the same bivariate approach as in section 2.2 (see also Kitchen and Monaco, 2003).

1. We consider a set of monthly indicators $\{x_{1,t}, x_{2,t}, \ldots, x_{k,t}\}$ and forecast the individual indicators $x_{i,t}$ over the relevant horizon from univariate autoregressive models,

$$x_{i,t} = \sum_{s=1}^{p_i} \rho_s x_{i,t-s} + u_{i,t} \ \ , \ i = 1,\ldots,k \ , \tag{3}$$

with coefficients ρ_s and white noise term $u_{i,t} \sim N(0,\sigma_i^2)$.

2. For each indicator $x^Q_{i,t}$, we consider the bridge equation

$$y^Q_t = \mu_i + \sum_{s=0}^{q_i} \beta_{is} x^Q_{i,t-s} + \varepsilon^Q_{i,t} \ , \tag{4}$$

   which relates quarterly GDP growth to the quarterly aggregate of the monthly indicator, evaluated in the third month of each quarter (see Mariano and Murasawa, 2003). Again, lag lengths p_i and q_i in the equations (3) and (4) are determined from the SIC. We produce a forecast of GDP growth, $y^Q_{i,t+h|t}$, by inserting the quarterly aggregates $x^Q_{i,t+h|t}$ of the forecasts $x_{i,t+h|t}$ into equation (4).

3. We form the average of the k resulting forecasts $y^Q_{i,t+h|t}$ from the individual indicators, as in step 3 in section 2.2.

2.2.2 Bridging with factors

Giannone, Reichlin and Sala (2004) and Giannone, Reichlin and Small (2005) propose the idea of bridging with factors. They consider the bridge equation

$$y_t^Q = \mu + \beta' f_t^Q + \varepsilon_t \quad , \tag{5}$$

where f_t^Q is a quarterly aggregate of common factors driving all the monthly indicators. Given a large set of monthly time series $x_t = (x_{1t},...,x_{nt})'$, we consider the following factor structure

$$x_t = \Lambda f_t + \xi_t \tag{6}$$

which relates the $n \times 1$ vector of monthly time series x_t to the $r \times 1$ vector of common factors $f_t = (f_{1t},...,f_{rt})'$ via a matrix of factor loadings Λ and to the idiosyncratic component $\xi_t = (\xi_{1t},...,\xi_{nt})'$. The number of static factors r is typically much smaller than the number of series n.

The procedure works in two steps. First the factors are extracted from the monthly indicators. We will consider two different approaches for extracting the factors.

1. Simple principal components (PC) following Stock and Watson (2002).
2. Two-step approach (KF) based on principal components and Kalman filtering (Doz, Giannone and Reichlin, 2007). In this approach the common factors f_t are assumed to follow vector autoregressive process which is driven by a vector of innovations $u_t = (u_{1t},...,u_{qt})'$ which are called the common shocks:[2]

$$f_t = \sum_{s=1}^{p} A_i f_{t-s} + Bu_t \tag{7}$$

The estimation by PC requires the setting of the number of common factors r only. The lag length p and the number of common shocks q need not be specified since the PC estimator does not take into account the dynamic properties of the common factors. The latter is explicitly taken into account by the KF approach, for which all the three parameters must be set.

The forecast of GDP is obtained in a second step. The Kalman filter delivers the forecasts of the common factors needed for predicting GDP, since it takes into account their dynamic properties. The forecast of GDP growth $y_{t+h|t}^Q$ is obtained by inserting into the bridge equation the quarterly aggregates of the estimated common factors and their forecast $f_{t+h|t}^Q$. Forecasts of the factors are not directly obtained when factors are extracted using PC, since in this procedure

[2] For more details on the generality of such representation, see Forni, Giannone, Lippi and Reichlin (2007).

the dynamics of the common factors are not explicitly considered. For this reason, the h-steps ahead forecast for GDP growth is computed with a direct approach, from the bridge equation $y^{Q}_{t+h} = \mu + \beta' f^{Q}_{t} + \varepsilon_{t}$, where GDP appears with a lead of h periods and there is hence no need to forecast monthly factors.

It remains to specify how to deal with ragged edges due to the non-synchronous flow of data releases. The KF estimator deals efficiently with ragged edges by replacing the missing observations with optimal predictions based on the entire set of monthly indicators. Concerning PC we deal with ragged edges by filling the missing monthly indicators with predictions based on univariate autoregressions, as done for the traditional bridge equations. Again, the lag length is determined from the SIC. Alternative methods are also studied for robustness (see section 5).

The factors extracted using the KF are appropriate combinations of present and past observations with weights derived by taking into consideration the persistence of the common factors and the heterogeneity in the informational content of every monthly indicator relative to the common factors. On the other hand, the factors extracted by PC are linear combinations only of the most recent observations since the PC estimator does not take into consideration the persistence of the common factors. Moreover, in PC all monthly indicators are considered to be equally informative about the common factors.

2.2.3 Generalised principal components

Another factor model that accounts for factor dynamics is given by the generalised principal components model (GPC) as put forward by Forni et al. (2005). Within this framework, no specific model is postulated for the factors. Therefore they can not be predicted directly, as it is the case with the KF approach.

In this paper, we deal with this issue by effectively running a quarterly model. We combine GDP growth and the quarterly aggregates of the monthly series in our dataset, from which factors are estimated. The GDP forecast is then obtained as a forecast of the common component of GDP, as provided by the factor model.[3]

Again, as with bridge equations and model PC, we deal with ragged edges by filling the missing monthly observations with predictions based on univariate autoregressions. We do so

[3] Possible alternative solutions – which are not considered in this paper – include: (i) using a monthly interpolation of GDP among the variables in x_t and taking the projection of the common component of this variable for the quarterly GDP forecast (Altissimo et al, 2001); (ii) extracting monthly "smooth" factors and regressing GDP growth on their appropriately transformed values (Altissimo et al. 2007).

While one may add a forecast of the idiosyncratic component, D'Agostino and Giannone (2006) report some evidence that this component is highly unforecastable.

before aggregating the data to quarterly frequency. Further, parameters r and q are to be specified. They are determined from the recursive minimum RMSE measure.

3. Pseudo real-time forecast design

In this section, the general principles underlying the forecasting exercise, which are applied to all models, are described.

3. 1. Forecast design

The forecast evaluation exercise is designed to predict quarterly GDP growth from monthly indicators, which are published within the quarter. While flash estimates of GDP growth are released around six weeks after the end of the quarter, a considerable amount of monthly data on real activity within the same quarter is published earlier. There may be gains in making use of this information when producing short-term forecasts for GDP.

With our forecast design, we aim at replicating the real-time application of the models as closely as possible. We do not have real-time datasets at hand. However, following Rünstler and Sédillot (2003) and Giannone et al. (2005) we take account of publication lags in the individual monthly series and consider a sequence of forecasts to replicate the flow of monthly information that arrives within a quarter.

More precisely, we consider a sequence of eight forecasts for GDP growth in a given quarter, obtained in consecutive months. The timing is illustrated in Table 2 and is best explained using an example. Assume that our objective is to forecast GDP growth in the second quarter of 2007. We start forecasting in January 2007: this forecast refers to next quarter GDP and we denote it as the first month *one quarter ahead* forecast. In moving forward in time we produce a forecast in each month, and – with the GDP flash estimate being published in mid-August – run the final forecast on 1 August. We denote the latter as the second month *preceding quarter* "forecast", which is actually a backcast. This sequence of forecasts is applied to each quarter of our out-of-sample period.

Another issue concerns the "unbalancedness" of the available data. The individual monthly series are published with different delays. As a result, the number of missing observations at the end of the sample differs across series. Survey and financial data, for instance, are available right at the end of the month, but industrial production data are published, for example, with a delay of six weeks for the euro area. Similar lags are found for other official statistics. In this respect, Giannone, Reichlin and Small (2005) and Banbura and Rünstler (2007) have shown that ignoring unbalancedness in the data may have strong effects on the results.

In this paper, we fully account for unbalancedness. We download our datasets at the beginning of the month, when most of the survey and financial market data for the previous month are already available. For each forecast, we apply in a recursive way the data release pattern that we find in our datasets to the time at which the forecasts are made. Formally, our pseudo real-time datasets X_t are defined as follows: given our main set of monthly observations, T x n matrix X_T, as downloaded on a certain day of the month, we define with t x n matrix X_t the observations from the original data X_T up to period t, but with elements $X_t(t\text{-}h,i)$ eliminated, if observation $X_T(T\text{-}h,i)$ is missing in X_T (for $i = 1,..,n$, and $h \geq 0$).

A forecast $y^Q_{t+h|t}$ made in period t is based on information set X_t. In all cases, we also re-estimate and re-specify the models in each point in time based on information set X_t. Given the absence of well agreed information criteria, the specification of factor models, i.e. the choices of the numbers of static (r) and dynamic factors (q) and the number of lags p in equation (6), is based on a recursive minimum RMSE criterion. In each month of the evaluation period, we simply select the specification that has provided the best forecasts in the past. More precisely, we calculate the average RMSE across all horizons and select the specification with minimum average RMSE. We repeat this in each individual month of the evaluation period. We limit the specification search to values of $r \leq 8$, $q \leq r$, and $p \leq 3$. In addition, we consider forecast averages across all specifications.

For those models that use only quarterly data, the same rules can be applied. At each point in time, we form the quarterly aggregates $x^Q_{i,t}$ of individual series $x_{i,t}$ from pseudo real-time datasets X_t and treat an observation in $x^Q_{i,t}$ as missing if the monthly data are not complete. Naturally, the forecasts then remain unchanged for three consecutive months, and are updated only once new quarterly data arrives, depending on publication lags.

4. Data

The data used in this paper comprise ten large datasets that have been compiled for the euro area as a whole as well as for six euro area countries (Belgium, Germany, France, Italy, Netherlands, Portugal) and three new Member States (Lithuania, Hungary and Poland). The datasets were downloaded in either early July or August 2006.

The datasets have an average dimension of more than one hundred series for each country and all series are available from January 1991 up to mid-2006, apart from the new Member States where the sample period is shorter (see Table 1 for details on the datasets). Additionally, quarterly real GDP series were also collected for the corresponding sample period.

All data are seasonally adjusted. For the analysis, the data are differenced to be stationary. For trending data (such as industrial production, employment, retail sales) we take logarithms beforehand, which amounts to calculating rates of change, while survey and financial data are not logarithmised. We use three-month differences of the monthly data, i.e. the rates of change against the same month of the previous quarter, $(x_t - x_{t-3})/3$.[4] This implies that the quarterly aggregate of the series is given by $x_t^Q = (x_t + x_{t-1} + x_{t-2})/3$ from a log-linear approximation.

In application, data X_t are standardised to mean zero and variance one in a recursive manner. For the factor models, we also clean the data from outliers in a recursive manner.[5]

5. Results

Concerning the out-of-sample period, for the euro area countries, we evaluate the forecast performance of the various models over the period from 2000 Q1 to 2005 Q4. For new Member States, the short samples require truncating the evaluation period to 2002 Q1 to 2005 Q4.[6]

5.1 Forecast accuracy

Taking into account the number of models considered and the different model selection criteria, balancing methods, etc. we end up with almost forty specifications for each country. In order to make the presentation of the results tractable, we narrow the number of specifications to be presented by focusing on the specifications that performed better while discussing the sensitivity of the results obtained.[7]

First, regarding quarterly VARs and traditional bridge equations, we considered two alternative sets of indicators. The first set comprises all indicators in the dataset. The second contains only those indicators that experts in central banks regard as being the most important when monitoring economic activity. In the first case, we average forecasts across all series in the dataset while in the second case we only average across a narrow dataset. Although the differences are minor, since the results of the latter are slightly better, we report only for those models (labelled as VAR_n and BE_n respectively in Table 3).

4 From a theoretical perspective, month-on-month differences, $x_t - x_{t-1}$ may be preferred as they allow for a more precise modelling of dynamics by avoiding a moving average structure of the residuals. From a practical perspective, using three-month differences has the advantage that noise in the data is reduced and data irregularities are smoothed out. We find that three-month differences tend to give better forecasts. The results are available from the authors upon request.

5 Outlier detection was based on a simple rule applied to the differenced series: we identified those observations as outliers, which were five times larger in absolute value than the 20% quantile of the series' distribution. We either set these outliers as missing values (model KF) or replace them with the value of the cut-off point.

6 When using recursive RMSE criterion for the factor model specifications, we use a "burning in" phase starting in 1998 Q1 to find the initial specification.

7 All the results are available from the authors upon request.

Second, as concerns factor models, we have considered alternative ways to specification search in addition to the recursive RMSE criterion as described in section 3.1. As one alternative option, we have combined information criteria proposed by Bai and Ng (2002, 2007) to determine the number of static and dynamic factors with the SIC to determine lag length p in equation (6). In addition, we have considered unweighted forecast averages across all specifications. Again, we find the differences to be rather small, but for all factor models, the recursive RMSE selection slightly outperforms the alternatives considered.

Third, for the PC and GPC estimation method we have also considered alternative methods to deal with ragged edges owing to the synchronicity of data releases. Precisely, in addition to the univariate models, we consider alternatives in which the predictions are obtained from multivariate models. First we shift the series with missing observations forward in time: if the last m observations are missing in series i, lagged series $\tilde{x}_{i,t} = x_{i,t-m}$ is used in place of $x_{i,t}$. Moreover, for the PC estimates we have also considered the EM algorithm developed by Stock and Watson, 2002a to handle missing observations. The differences are, on average, small, but the results of univariate models reported here tend to fare slightly better, in particular for PC.[8]

The main results for the preferred specifications are shown in Table 3. We report the RMSE of each model relative to the naïve benchmark of constant growth. A number lower than one indicates that the model's forecasts are more accurate than the average growth over the past sample. We report measures for individual countries and the euro area. We also report in the right panel the mean MSFE across the euro area countries (excluding the euro area as a whole) and new Member States. In the bottom panel we report the rank across models and, in the last two columns, the mean rank for euro area countries and new Member States.

The findings differ qualitatively among the euro area countries and the new Member States. The two groups of countries are therefore discussed separately.

The results for the euro area countries included in the study might be summarised as follows:

a. Models that use monthly data tend to outperform those models that use purely quarterly data. Bridge equation and factor models, that incorporate early releases, produce forecasts that are more accurate than those based on quarterly models. These results highlight the importance of exploitation of monthly releases.
b. Factor-based estimates are in general more accurate than forecasts based on simple bridge equations. With the exception of the Netherlands (and one minor exception in the case of Italy), the three factor models rank ahead of the alternative models. This indicates that

[8] The PC-EM algorithm estimates the factors from the available observations and uses these estimates to predict missing observations. This procedure is iterated until convergence.

bridging with factors extracted from many monthly time series is preferable to the average of many small bridge equations each constructed with individual monthly series.

c. Among the factor models the most accurate forecasts are those based on factors extracted by the KF proposed by Giannone, Reichlin and Small (2005). The KF methods attain rank one for all countries but France and the Netherlands. For France, model PC fares slightly better, while for the Netherlands the quarterly VAR performs best. [9]

d. Estimates of GDP growth at euro area aggregate level are more accurate than the estimates of GDP growth in individual Member States. The estimates based on the common factors extracted by the KF improve upon the naïve forecast by 25 percent in the euro area. The accuracy relative to the naïve model is much less pronounced for individual countries and for several countries we find little improvement over the naïve constant growth model.

The differences in the average RMSE across countries are small. However, one can establish significant differences from considering the cross-country perspective. Assume that the ranks of the individual models are independent across countries and consider the null hypothesis that two models perform equally well. Under the null hypothesis, the probability that model 1 is found to perform better than model 2 in k of n countries is found from the binomial distribution with

$$0.5^n \sum_{j=1}^{k} \binom{n}{j}.$$

For $n=7$ one can establish that the probability that model 1 performs better than model 2 in six or all seven cases amounts to $p=0.063$ and $p=0.008$ respectively. Hence, we can establish from the rank statistics that the improvement of factor models extracted by KF and PC over the bridge equations, quarterly VARs and the factors extracted by GPC is significant. Equivalently, the forecasts based on factors extracted using KF are significantly more accurate than those based on factors extracted by PC.

As regards the three new Member States, in general the model-based forecasts are not uniformly better than the naïve forecasts. These findings may be related to the short samples at hand (data start only in 1995-1998), the rapid transition of the economies, which implies unstable relationships among series, and possibly other issues regarding the quality of the data

[9] Although not reported in this paper, for the Netherlands, the KF model based on information criteria performs best across all specifications including the quarterly VARs.

(for example, a lack of seasonally adjusted monthly data means it is necessary to use 12-month differences of the data).

Tables 4a to 4c show the corresponding measures for averages of the RMSE over the individual quarters of the forecast horizon. One can see that the relative performance of the models remains stable across horizons. The factor models, in particular, continue to outperform the quarterly models and bridge equations, with a model based on factors extracted by the KF performing best for the preceding and current quarter forecasts. The differences across methods are less pronounced for the one-quarter-ahead forecasts when the relative RMSE tends to one, which represents non-forecastability.

5.2 Encompassing tests

Forecast encompassing tests are another means to assess the relative performance of models. The encompassing test between two alternative models 1 and 2 is based on a regression of the actual data y_t^Q on forecasts $f_{1,t}^Q$ and $f_{2,t}^Q$ from two models (see, e.g. Clements and Hendry, 1998: 228ff),

$$y_t^Q = \lambda f_{1,t}^Q + (1-\lambda) f_{2,t}^Q + u_t\,, \qquad 0 \le \lambda \le 1. \tag{8}$$

Parameter λ gives the optimal weight of model 1 in the combined forecast. In the extreme case, a value of $\lambda = 1$ indicates that model 1 dominates model 2, i.e. forecasts $f_{2,t}^Q$ from model 2 do not contain any information beyond the information contained in forecasts $f_{1,t}^Q$. Hence, forecasts from model 2 can be disregarded. Equivalently, a value of $\lambda = 0$ implies that forecasts from model 1 can be disregarded. In the intermediate case of $0 < \lambda < 1$, combinations of forecasts from the two models might be considered.

Table 5 shows encompassing tests of the models shown in Table 3 against the best-performing one, KF. Here, a large value of λ means that a model based on factors estimated by the KF dominates the alternative model. The tests are shown for the forecasts obtained in the second month of the current quarter, which represents the centre of our forecast horizon.

For the euro area countries, the results indicate some dominance of estimates based on the factor model with KF against models AR, VAR and bridge equations. Estimates of λ always exceed a value of 0.5 and are in many close to one. The hypothesis of $\lambda = 0$, i.e. that the estimates based on factors extracted by the KF would not add information to forecasts from these alternative models is uniformly rejected. The opposite hypothesis of $\lambda = 1$, i.e. that models AR(1), BE and VAR do not add information to forecasts from the KF-based factor model is rejected only in the case of Germany. Furthermore, the KF-based estimates of the factor model also tend to attain high weights against the alternative factor models. With the exception of model GPC in case of Belgium, λ is estimated larger than 0.5, while the hypothesis of $\lambda = 0$ is rejected in most cases.

We have also performed encompassing tests for other forecast horizons. With one exception, the findings remain reasonably robust across horizons. The exception is that the dominance of estimates based on the KF against the estimates based on PC is lost for higher horizons, i.e. the one-quarter-ahead quarter forecasts. A possible reason is related to the efficiency of model KF in dealing with unbalanced data. While this advantage may be particularly important for the very short horizons, it may become less important for the next quarter forecasts.[10]

For the new Member States, the ranking among forecasts methods cannot be established. This is expected given that the evaluation and estimation samples are both very short.

6. Conclusions

This paper has performed a large-scale forecast exercise, involving ten large datasets for ten European countries and one large dataset for the euro area economy. We have compared simple quarterly models with models exploiting more timely monthly data to obtain early estimates and short-term forecasts of quarterly GDP growth. Amongst these models we have considered both traditional bridge equations and factor models adapted to handle unsynchronised data releases. The forecast design has aimed at replicating the real-time application of the models as closely as possible. It deviates from a real-time application only insofar as we had to use final data releases, as such real-time data are not readily available.

The main message of the results obtained for the euro area countries is that models that exploit timely monthly releases fare better than quarterly models. Amongst those, factor models, which exploit a large number of releases, do generally better than averages of bridge equations. This suggests that the idea of using factors to bridge monthly with quarterly information is promising and should be more systematically explored in the Eurosystem. We have also tried to establish a ranking between different estimators and between different methods to handle unbalanced data at the end of the sample. Differences between different approaches were found to be small, with the exception of the experiment based on the euro area aggregate dataset where the Kalman-filter-based procedure proposed by Giannone, Reichlin and Sala (2004) and Giannone, Reichlin and Small (2005) gives significantly better results.

[10] The results that the gains from using the KF are less pronounced for longer horizons are in line with findings based on the Monte Carlo exercise performed by Doz, Giannone and Reichlin (2007).

Results for the new Member States, on the other hand, are difficult to interpret. All models perform quite badly with respect to naïve benchmarks, but, given the short evaluation sample, it is hard to understand what drives the results.

On the basis of this first evaluation we can outline an agenda for more detailed studies on short-term forecasting methods:

1. Evaluate the design of bridge equations which are routinely used in some institutions.
2. The bridge models can be further extended and refined both in terms of identifying key monthly releases and extending the class of models. Bayesian VARs extended to handle the bridge problem, for example, should be given further consideration.
3. For factor-based bridge equations, further thought should be given to variables selection (size of the dataset) and data transformations.
4. Our evaluation does not clearly distinguish between methods of estimation and methods of filling missing observations at the end of the sample. This could be the subject of a more detailed evaluation although our results do suggest that differences between methods are minimal.
5. Models that handle the data flow problem of short-term forecasting in a unified framework can be extended to provide an interpretation of the contributions of data releases to the forecast and to the uncertainty around the forecast along the lines suggested by Angelini et al. (2008), Banbura and Rünstler (2007) and Giannone, Reichlin and Small (2005).
6. Results for the new Member States should be further evaluated. In order to perform the evaluation and the comparison, the present study is based on very short estimation samples which make the results unreliable. However, at present it is possible to use at least ten years of data for the new Member States. Results should be revaluated using the longer sample.

References

Aastveit, K. A. and T. G. Trovik (2007), Nowcasting Norwegian GDP: The role of asset prices in a small open economy. Norges Bank working paper 2007/09.

Altissimo, F., A. Basanetti, R. Cristadoro, M. Forni, M. Hallin, M. Lippi, L. Reichlin and G. Veronese (2003), A real-time coincident indicator for the euro area business cycle. CEPR discussion paper No 5633.

Altissimo, F., R. Cristadoro, M. Forni, M. Lippi and G. Veronese (2007), New Eurocoin: Tracking economic growth in real time. Bank of Italy, Tema di Discussione, No 631.

Angelini, E., G. Camba-Mendez, D. Giannone, L. Reichlin and G. Rünstler (2008a), Short-term forecasts of euro area GDP. ECB Working Paper series, forthcoming.

Angelini, E., M. Banbura, and G. Rünstler (2008b), Estimating and forecasting the euro area monthly national accounts from a dynamic factor model. ECB Working Paper series, forthcoming.

Artis, M.J., A. Banerjee and M. Marcellino (2005), Factor forecasts for the UK. *Journal of Forecasting* 24, 279-298.

Bai, B.J. and S. Ng (2002), Determining the number of factors in approximate factor models. *Econometrica* 71(1), 191-221.

Bai, B.J. and S. Ng (2007), Determining the number of primitive shocks in factor models. *Journal of Business and Economic Statistics* 25(1), 52-60.

Baffigi, A., R. Golinelli, and G. Parigi (2004), Bridge models to forecast the euro area GDP. *International Journal of Forecasting* 20(3), 447-460.

Banbura, M. and G. Rünstler (2007), A look into the factor model black box: publication lags and the role of hard and soft data in forecasting GDP. ECB Working Paper No 715.

Bruneau, C., O. de Bandt, A. Flageollet, and E. Michaux (2007), Forecasting inflation using economic indicators: the case of France. *Journal of Forecasting* 26 (1), 1-22.

Camba-Mendez, G., G. Kapetianos, M. Smith and R. Weale (2001), An automatic leading indicator of economic activity: forecasting GDP growth for European countries. *Econometrics Journal* 4(1), 56-80.

Clements, M. and D. Hendry (1998), Forecasting Economic Time Series. Cambridge: Cambridge University Press.

D'Agostino, A., D. Giannone and Surico, F. (2006), (Un)predictability and macroeconomic stability. ECB Working Paper No 605.

D'Agostino, A. and D. Giannone (2006), Comparing alternative predictors based on large-panel factor models. ECB Working Paper No 680.

Den Reijer, A.H.J. (2007), Forecasting Dutch GDP using large scale factor models. *mimeo*.

Doz, C., D. Giannone and L. Reichlin (2007), A Two-Step Estimator for Large Approximate Dynamic Factor Models Based on Kalman Filtering. CEPR discussion paper 6043.

Duarte, C. and A. Rua (2007), Forecasting inflation through a bottom-up approach: How bottom is bottom? *Economic Modelling* 24, 941-953.

Durbin, J. and S.J. Koopman (2001), Time Series Analysis by State Space Methods. Oxford: Oxford University Press.

Forni, M., D. Giannone, M. Lippi and L. Reichlin (2007), Opening the black box – structural factor models with large gross-sections. ECB Working Paper No 712.

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

Forni, M., M. Hallin, M. Lippi, and L. Reichlin (2005), The generalized dynamic factor model: one-sided estimation and forecasting. *Journal of the American Statistical Association*, 100, 830-840.

Giannone, D., L. Reichlin and L. Sala (2004), Monetary Policy in Real Time. Mark Gertler and Kenneth Rogoff, editors, NBER Macroeconomics Annual 2004, pages 161-200. MIT Press.

Giannone, D., L. Reichlin and D. Small (2005), Nowcasting GDP and inflation: the real-time informational content of macroeconomic data releases. Federal Reserve Bank Finance and Economics Discussion Series 2005-42, *Journal of Monetary Economics* forthcoming.

Kitchen, J. and R. Monaco (2003), Real-time forecasting in practice: The U.S. Treasury staff's real-time GDP forecast system. *Business Economics* 38 (4), 10-28.

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

Marcellino, M., Stock, J.H. and M. Watson (2003), Macroeconomic forecasting in the euro area: country specific versus euro wide information. *European Economic Review* 47, 1-18.

Matheson, T. (2007), An analysis of the informational content of New Zealand data releases: the importance of business opinion surveys. Reserve Bank of New Zealand Discussion Paper DP2007/13.

Rünstler, G. and F. Sédillot (2003), Short-term estimates of euro area real GDP by means of monthly data. ECB Working Paper No 276.

Schumacher, C. (2007), Forecasting German GDP using alternative factor models based on large data sets. *Journal of Forecasting* 26(4), 271-302.

Stock, J.H. and M. Watson (1999), Forecasting inflation. *Journal of Monetary Economics* 44, 293-335.

Stock, J.H. and M. Watson (2002a), Macroeconomic forecasting using diffusion indices. *Journal of Business and Economics Statistics* 20, 147-162.

Stock, J.H. and M. Watson, (2002b), Forecasting using principal components from a large number of predictors. *Journal of the American Statistical Association* 97, 1167-1179.

Van Nieuwenhuyze, C. (2005), A generalised dynamic factor model for the Belgian economy: identification of the business cycle and GDP growth forecasts. *Journal of Business Cycle Measurement and Analysis*, 2(2): 213-247.

Tables

Table 1: Datasets

		No of series	*of which* Production and sales	Surveys	Financial	Prices	Other	Sample start
Euro area	EA	85	25	25	24	0	11	1991 M1
Belgium	BE	393	25	262	50	42	14	1991 M1
Germany	DE	111	55	19	32	4	1	1991 M1
France	FR	118	19	96	0	2	1	1991 M1
Italy	IT	84	27	24	10	20	3	1991 M1
Netherlands	NL	76	8	33	8	23	4	1991 M1
Portugal	PT	141	32	78	12	10	9	1991 M1
Lithuania	LT	103	35	21	12	33	1	1995 M1
Hungary	HU	80	33	9	12	11	15	1998 M1
Poland	PL	81	16	30	10	11	14	1997 M1

Table 2: Timing of forecast exercise

(Example: forecasts for second quarter)

Quarter to be forecast		Forecast made on first day of
One quarter ahead	1	January
	2	February
	3	March
Current	1	April
	2	May
	3	June
Preceding	1	July
	2	August

Table 3: Results overview

Forecasts 2000 Q1 – 2005 Q4 for euro area countries and 2002 Q1– 2005 Q4 for NMS

Average RMSE for preceding, current and one-quarter-ahead forecasts relative to the naive forecast

	EA	BE	DE	FR	IT	NL	PT	LT	HU	PL	EuroA	NMS
AR	0.92	0.99	1.04	0.99	0.91	1.03	1.00	1.07	0.96	0.82	0.99	0.95
VAR	0.90	1.10	1.06	0.95	0.94	0.94	0.95	0.92	1.07	0.71	0.99	0.90
BEQ	0.87	0.94	1.04	0.94	0.96	1.01	0.93	1.01	1.05	0.82	0.97	0.96
KF	0.75	0.89	0.95	0.87	0.87	0.94	0.84	1.07	1.07	1.01	0.89	1.05
PC	0.85	0.90	0.95	0.85	0.90	1.01	0.85	1.12	1.12	1.02	0.91	1.09
GPC	0.93	0.90	0.99	0.91	0.93	0.97	0.91	1.07	0.95	0.80	0.94	0.94

Ranks of models according to the RRMSE measure

	EA	BE	DE	FR	IT	NL	PT	LT	HU	PL	EuroA	NMS
AR	5	5	5	6	3	6	6	5	2	3	5.2	3.3
VAR	4	6	6	5	5	1	5	1	5	1	4.7	2.3
BEQ	3	4	4	4	6	4	4	2	3	4	4.3	3.0
KF	1	1	1	2	1	2	1	4	4	5	1.3	4.3
PC	2	3	2	1	2	5	2	6	6	6	2.5	6.0
GPC	6	2	3	3	4	3	3	3	1	2	3.0	2.0

AR denotes a univariate autoregressive model for GDP; VAR and BEQ denote the quarterly bivariate VAR and bridge equation models respectively. KF, PC and GPC denote the 3 versions of factor models, based on the Kalman filter, principal components and generalised principal components respectively.

See Table 1 for an explanation of country abbreviations; EA denotes data for the euro area aggregate, while EuroA and NMS denote averages of the various measures across the six euro area Member States and the three new Member States included in the investigation respectively.

Table 4a: Results overview – preceding quarter

Forecasts 2000 Q1 – 2005 Q4 for euro area countries and 2002 Q1 – 2005 Q4 for NMS

Average RMSE for preceding quarter forecasts relative to the naive forecast

	EA	BE	DE	FR	IT	NL	PT	LT	HU	PL	EuroA	NMS
AR	0.82	1.00	1.06	1.02	0.81	1.02	1.01	1.05	0.97	0.72	0.99	0.91
VAR	0.81	1.10	1.08	0.96	0.85	0.89	0.95	0.93	1.11	0.81	0.97	0.95
BEQ	0.84	0.87	1.02	0.90	0.93	0.98	0.91	0.99	0.95	0.87	0.93	0.94
KF	0.71	0.77	0.96	0.72	0.86	0.93	0.73	1.14	1.08	1.20	0.83	1.14
PC	0.78	0.86	0.95	0.68	0.90	1.03	0.74	1.28	1.08	1.36	0.86	1.24
GPC	0.91	0.86	0.97	0.89	0.89	0.96	0.86	1.04	0.91	0.76	0.91	0.90

Rank of models according to the RRMSE measure

	EA	BE	DE	FR	IT	NL	PT	LT	HU	PL	EuroA	NMS
AR	4	5	5	6	1	5	6	4	3	1	4.7	2.7
VAR	3	6	6	5	2	1	5	1	6	3	4.2	3.3
BEQ	5	4	4	4	6	4	4	2	2	4	4.3	2.7
KF	1	1	2	2	3	2	1	5	5	5	1.8	5.0
PC	2	2	1	1	5	6	2	6	4	6	2.8	5.3
GPC	6	3	3	3	4	3	3	3	1	2	3.2	2.0

AR denotes a univariate autoregressive model for GDP; VAR and BEQ denote the quarterly bivariate VAR and bridge equation models respectively. KF, PC and GPC denote the 3 versions of factor models, based on the Kalman filter, principal components and generalised principal components respectively.

See Table 1 for an explanation of country abbreviations; EA denotes data for the euro area aggregate, while EuroA and NMS denote averages of the various measures across the six euro area Member States and the three new Member States included in the investigation respectively.

Table 4b: Results overview – current quarter

Forecasts 2000Q1 – 2005 Q4 for euro area countries and 2002 Q1 – 2005 Q4 for NMS

Average RMSE for current quarter forecasts relative to the naive forecast

	EA	BE	DE	FR	IT	NL	PT	LT	HU	PL	EuroA	NMS
AR	0.91	0.99	1.03	0.98	0.92	1.03	1.00	1.09	0.95	0.82	0.99	0.95
VAR	0.89	1.09	1.05	0.93	0.95	0.94	0.95	1.09	1.03	0.70	0.99	0.94
BEQ	0.85	0.93	1.03	0.92	0.95	1.00	0.93	1.04	1.06	0.85	0.96	0.98
KF	0.76	0.90	1.00	0.88	0.86	0.91	0.84	1.08	1.06	1.03	0.90	1.06
PC	0.86	0.86	0.97	0.90	0.89	1.02	0.87	1.14	1.05	0.99	0.92	1.06
GPC	0.91	0.84	1.02	0.87	0.92	0.94	0.90	1.12	0.92	0.78	0.92	0.94

Rank of models according to the RRMSE measure

	EA	BE	DE	FR	IT	NL	PT	LT	HU	PL	EuroA	NMS
AR	6	5	5	6	3	6	6	4	2	3	5.2	3.0
VAR	4	6	6	5	6	2	5	3	3	1	5.0	2.3
BEQ	2	4	4	4	5	4	4	1	5	4	4.2	3.3
KF	1	3	2	2	1	1	1	2	6	6	1.7	4.7
PC	3	2	1	3	2	5	2	6	4	5	2.5	5.0
GPC	5	1	3	1	4	3	3	5	1	2	2.5	2.7

AR denotes a univariate autoregressive model for GDP; VAR and BEQ denote the quarterly bivariate VAR and bridge equation models respectively. KF, PC and GPC denote the 3 versions of factor models, based on the Kalman filter, principal components and generalised principal components respectively.

See Table 1 for an explanation of country abbreviations; EA denotes data for the euro area aggregate, while EuroA and NMS denote averages of the various measures across the six euro area Member States and the three new Member States included in the investigation respectively.

Table 4c: Results overview – one quarter ahead

Forecasts 2000 Q1 – 2005 Q4 for euro area countries and 2002 Q1 – 2005 Q4 for NMS

Average RMSE for one-quarter-ahead forecasts relative to the naive forecast

	EA	BE	DE	FR	IT	NL	PT	LT	HU	PL	EuroA	NMS
AR	1.00	0.99	1.03	0.98	0.99	1.04	0.99	1.11	0.97	0.90	1.00	0.99
VAR	0.98	1.10	1.05	0.95	1.01	0.97	0.96	1.09	1.09	0.67	1.01	0.95
BEQ	0.90	0.99	1.05	1.01	0.99	1.03	0.95	1.06	1.12	0.77	1.00	0.98
KF	0.78	1.07	0.89	0.97	0.89	0.98	0.92	1.07	1.08	0.87	0.95	1.01
PC	0.90	0.99	0.93	0.95	0.92	0.98	0.90	1.13	1.23	0.86	0.94	1.07
GPC	0.96	0.99	0.97	0.97	0.97	1.01	0.95	1.13	1.01	0.83	0.98	0.99

Rank of models according to the RRMSE measure

	EA	BE	GE	FR	IT	NL	PT	LT	HU	PL	EuroA	NMS
AR	6	2	4	5	4	6	6	4	1	6	4.5	3.7
VAR	5	6	6	2	6	1	5	3	4	1	4.3	2.7
BEQ	3	4	5	6	5	5	4	1	5	2	4.8	2.7
KF	1	5	1	4	1	2	2	2	3	5	2.5	3.3
PC	2	3	2	1	2	3	1	6	6	4	2.0	5.3
GPC	4	1	3	3	3	4	3	5	2	3	2.8	3.3

AR denotes a univariate autoregressive model for GDP; VAR and BEQ denote the quarterly bivariate VAR and bridge equation models respectively. KF, PC and GPC denote the 3 versions of factor models, based on the Kalman filter, principal components and generalised principal components respectively.

See Table 1 for an explanation of country abbreviations; EA denotes data for the euro area aggregate, while EuroA and NMS denote averages of the various measures across the six euro area Member States and the three new Member States included in the investigation respectively.

Table 5: Encompassing tests against model KF (selected models)

Forecasts 2000 Q1 – 2005 Q4 for euro area countries and 2002 Q1 – 2005 Q4 for NMS

Point estimate of parameter λ in the encompassing regression $y_t^Q = \lambda f_{1,t}^Q + (1\text{-}\lambda) f_{2,t}^Q + u_t$

Second month current quarter forecasts

	EA	BE	DE	FR	IT	NL	PT	LT	HU	PL
AR	0.92	0.78	0.63	0.65	0.95	1.06	0.86	-0.73	0.11	0.02
VAR	0.87	0.89	0.65	0.53	1.02	0.73	0.81	-0.68	0.26	0.05
BEQ	0.82	0.91	0.65	0.62	0.90	1.05	0.78	-0.86	0.54	0.04
PC	1.28	0.57	0.67	1.26	0.68	1.10	0.89	0.10	0.53	-0.23
GPC	1.03	0.55	0.72	0.27	0.93	0.78	0.72	-0.42	-0.08	-0.08

Test of the null hypothesis of λ = 1

	EA	BE	DE	FR	IT	NL	PT	LT	HU	PL
AR			*					**	**	**
VAR			*					**	*	**
BEQ			*	*				**		**
PC								*		**
GPC		*						**	**	**

*** and * denote rejection of the null hypothesis of λ = 1 at the 5% and 10% levels, respectively.*

Test of the null hypothesis of λ = 0

	EA	BE	DE	FR	IT	NL	PT	LT	HU	PL
AR	++	++	++	++	++	++	++			
VAR	++	++	++	++	++	++	++			
BEQ	++	++	++	++	++	++	++			
PC	++			++		++	++			
GPC	++		++		++	++	++			

++ and + denote rejection of the null hypothesis of λ = 0 at the 5% and 10% levels, respectively.

AR denotes a univariate autoregressive model for GDP; VAR and BEQ denote the quarterly bivariate VAR and bridge equation models respectively. KF, PC and GPC denote the 3 versions of factor models, based on the Kalman filter, principal components and generalised principal components respectively.

See Table 1 for an explanation of country abbreviations; EA denotes data for the euro area aggregate.