Forecasting Euro Area Real GDP: Optimal Pooling of Information

Oliver Hülsewig* Johannes Mayr* Timo Wollmershäuser†

June 25, 2008

Preliminary Version

Abstract

This paper proposes a new method of forecasting euro area quarterly real GDP that uses area–wide indicators, which are derived by optimally pooling the information contained in national indicator series. We construct the area–wide indicators by utilizing weights that minimize the variance of the out–of–sample forecast errors of the area–wide target variable. In an out–of–sample forecast experiment we find that our optimal pooling of information approach performs well compared to alternative forecasting methods in terms of forecast accuracy.

JEL classifications: C13, C51, C53, C82, E37

Key words: Forecasting, Aggregation, Model Averaging, Real Time Experiment

*Ifo Institute for Economic Research, Poschingerstr. 5, 81679 München, Germany. *Email:* <Huelsewig@ifo.de> and <Mayr@ifo.de>

†CESifo and Ifo Institute for Economic Research, Poschingerstr. 5, 81679 München, Germany. *Email:* <Wollmershaeuser@ifo.de>

1 Introduction

Since the first official release of euro area quarterly real GDP is published by Eurostat several weeks after the end of each quarter, a prompt assessment of the current state of the economy is appreciable. Timely information is contained in business cycle indicators – e.g. industrial production, confidence surveys or composite indicators – that are more promptly available. Forecasts of euro area quarterly real GDP are frequently derived by means of bridge models that explicitly incorporate such business cycle indicators.

In the euro area, business cycle indicators are typically collected at a national level by national statistical agencies or national survey institutes. In such a data–rich environment professional forecasters who aim at predicting euro area quarterly real GDP, can choose between two forecast strategies: pooling of forecasts and pooling of information (Diebold and Lopez, 1996). Pooling of forecasts uses national indicator series as predictors in the bridge model equations. One strategy is to produce a large number of forecasts of euro area real GDP growth rates by employing various parsimonious models and to combine them to a single forecast of the area–wide target variable. The optimal weighting scheme thereby takes the correlations of the forecast errors of each model into account (Bates and Granger, 1969). Alternatively, real GDP growth rates of each euro area member country can be forecasted separately and then be aggregated to a single euro area real GDP growth rate by using the relative economic weight of each member country (Marcellino, Stock, and Watson, 2003). Both strategies, however, entail a trade–off between the informational gain of using disaggregate national data and the risk of higher specification errors due to the need to estimate a large number of regression models.

Pooling of information generates a projection of euro area real GDP growth rates by using area–wide indicators as predictors that combine the information of the national indicators. It thereby reduces the number of regressions to one. The simplest strategy is to employ the area–wide indicators which are provided by Eurostat or other institutions – e.g. the European Commission or the OECD – and which are economically weighted averages of national indicators. Alternatively, professional forecasters frequently combine the set of national information by extracting common dynamic factors or principal components (Forni, Hallin, Lippi, and Reichlin, 2000, and Stock and Watson 2002).

This paper proposes a new method of forecasting euro area quarterly real GDP that uses area–wide indicators, which are derived by optimally pooling the information contained in national indicator series. We construct the area–wide indicators by utilizing weights that minimize the variance of the out–of–sample forecast errors of the aggregate target variable. By allowing a pre–aggregation of individual information to national indicator series, the optimal pooling of information problem is reduced to a manageable number of variables, which avoids the construction of a *super model* whose computation is often deemed to be prohibitively costly or even impossible (Timmermann, 2005).

To evaluate the forecast performance of our optimal pooling of information approach, we focus on three euro area business cycle indicators, which are all available at both the area–wide and the national level: the Industrial Production Index (IPI), the Economic Sentiment Indicator (ESI) of the European Commission and the CESifo World Economic Survey (WES) indicator for the euro area. The forecast models are specified as Autoregressive Distributed Lag (ADL) models, which are estimated by employing a model averaging strategy in order to reduce the problems associated with selecting a certain lag length. In a first step, we evaluate the full potential gain of the optimal pooling of information approach and analyze the weighting scheme. Our main result is that, compared to economically weighted indicators, optimally pooled area–wide indicators significantly reduce the out–of–sample mean squared forecast errors (MSE) for euro area quarterly real GDP growth by 40% on average. Furthermore, we find that the optimal weights derived from shorter optimization windows are almost identical to those derived from the entire out–of–sample window. These results indicate a certain stability of the optimal weights and support the application of our approach in real–time.

In a second step, we evaluate the applicability of the optimal pooling of information approach in real–time by employing a pseudo out–of–sample forecast experiment, in which optimally pooled area–wide indicators are computed using only ex–post information that would have been available in real–time. The optimized weights are thereby derived from a recursive growing optimization window, which is then excluded from the forecast evaluation process. The performance of the optimal pooling of information approach is thereby compared to a number of alternative forecasting methods, which include pooling of forecast strategies, i.e. optimal combination of area–wide GDP forecasts and aggregation of national GDP forecasts, as well as competing pooling of information strategies that rest on economic weights and OLS weights in addition to constructed area–wide indicators that are derived from principal components analysis and dynamic factor models. We find that the optimal pooling of information approach generally performs well compared to the competing forecasting methods in terms of forecast accuracy as measured by the out–of–sample forecast MSE.

The remainder of the paper is structured as follows. Section 2 reviews the traditional forecast strategies. In Section 3 we introduce the optimal pooling of information approach. In Section 4 we present our forecast experiment. We describe the forecast models applied, introduce the data set and discuss the empirical results, which refer to (i) the use of ex–ante information and (ii) to the use of ex–post information. Section 5 summarizes and concludes.

2 Review of Traditional Forecast Strategies

For an overview of the traditional forecast strategies we introduce the following notations. Suppose we forecast the aggregate target variable Y_t – i.e. euro area quarterly real GDP growth – using a broad set of disaggregate information variables, denoted by $X_{i,t}$, where t is time and i refers to the disaggregate unit, i.e. the member states of the currency area. The number of disaggregate units is given by K. The data sample that is available for the forecast experiment ranges from $t = 1, \ldots, \Theta_2$. The forecast model is estimated recursively over the estimation window $[1, T]$, with T gradually increasing from Θ_0 to $\Theta_2 - 1$, where $1 < \Theta_0 < \Theta_2 - 1$.

The one–step–ahead out–of–sample forecasts of the area–wide target variable, denoted by $\widehat{Y}_{T+1|T+1}$, are computed for $T + 1$ using the national information already available at $T + 1$.[1] As T increases from Θ_0 to $\Theta_2 - 1$, the number of

[1] Since in our set–up the current quarter is estimated, the literature often uses the notion 'nowcast' instead of forecast (Domenico, Reichlin, and Small (2006)).

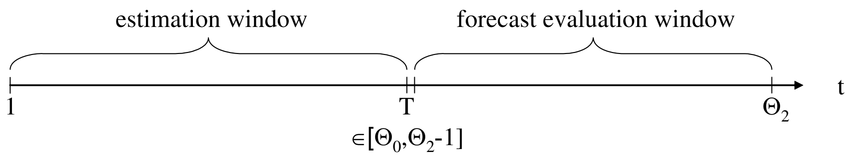

Figure 1: Time structure of the estimation and forecasting procedures

out–of–sample forecasts is given by $\Theta_2 - \Theta_0$. The performance of the different forecast strategies is evaluated by computing the MSE for each model over the forecast evaluation window $[\Theta_0 + 1, \Theta_2]$ on the basis of the out–of–sample forecast errors $\widehat{\varepsilon}_{T+1|T+1} = Y_{T+1} - \widehat{Y}_{T+1|T+1}$. Figure 1 outlines the time structure of the estimation and forecast evaluation window.

Notice that in the following we use a static structure of the forecasting models to keep the review as simple as possible. Later in the empirical part of the paper, we allow for more dynamics.

2.1 Pooling of Forecasts

Pooling of forecasts summarizes the combination of two or more individual forecasts to generate one single, pooled forecast. The idea of improving the accuracy of predictions regarding a certain target variable by combining the forecasts of different models was developed by Bates and Granger (1969) and Granger and Ramanathan (1984) and mainly follows the ideas of portfolio optimization and diversification gains. A large number of theoretical and empirical studies – see e.g. Timmermann (2005) and Stock and Watson (2004) – have shown the superiority of combined model based predictions.

In the context of forecasting euro area quarterly real GDP, three strategies have been proposed for combining single forecasts, which are derived from national indicator series using a multiple equation set–up. The crucial issue in all strategies is the determination of an adequate weighting scheme.

2.1.1 Optimal Combination of Area–Wide GDP Forecasts

In the first strategy the following forecasting model is estimated for each of the K national indicators over the period $t = 1, ..., T$:

$$Y_t = \delta + c_i X_{i,t} + \varepsilon_{i,t}. \tag{1}$$

The K forecasts resulting from the models are then linearly combined to a single forecast for the area–wide target variable according to:

$$\widehat{Y}_{T+1|T+1} = \sum_{i=1}^{K} \omega_i \widehat{Y}^i_{T+1|T+1}, \tag{2}$$

where the superscript i attached to $\widehat{Y}_{T+1|T+1}$ denotes the forecast of the area–wide target variable obtained from the model using the national indicator $X_{i,t}$.

The optimal weights ω_i of the single forecasts, and hence the weights attributed to each model, depend on the model's out–of–sample performance. Under the assumption that the forecasts are unconditionally unbiased, the $\Theta_2 - \Theta_0$ out–of–sample forecast errors of model i, $\widehat{\varepsilon}^i_{T+1|T+1} = Y_{T+1} - \widehat{Y}^i_{T+1|T+1}$ with $T = \Theta_0, \ldots, \Theta_2 - 1$, are normally distributed around zero with variance σ_i^2 and covariance $\rho_{ij}\sigma_i\sigma_j$ for $j = 1, ..., K$. Defining $\boldsymbol{\omega}$ as the $K \times 1$ vector containing the weights of each model and $\boldsymbol{\Sigma}_{\widehat{\varepsilon}}$ as the $K \times K$ variance–covariance matrix of the out–of–sample forecast errors, the optimal weights are obtained from minimizing the variance of the combined out–of–sample forecast error:

$$\boldsymbol{\omega}^{opt} = \arg\min_{\omega} \left[\boldsymbol{\omega}' \boldsymbol{\Sigma}_{\widehat{\varepsilon}} \boldsymbol{\omega} \right] \text{ s.t. } \sum_{i=1}^{K} \omega_i = 1. \tag{3}$$

Assuming linear relationships, optimal weights can be derived by ordinary least squares, regressing realizations of the target variable Y_t on the K–vector of forecasts $\widehat{Y}^i_{T+1|T+1}$ and a constant term. The weights are thereby assumed to be non–negative and to sum up to unity. A major benefit of the combination of forecasts approach is the possibility of including a large number of candidate regressors in forecasting a certain target series without of running into the problem of overparametrization or overfitting. However, as the data generating process is typically unknown, the need to specify a large number of parsimonious regression models may lead to high specification errors (Lütkepohl, 1987). Another challenge of the approach is the estimation of the variance–covariance matrix $\mathbf{\Sigma}_{\widehat{\varepsilon}}$. If the number of models K increases, the computation of optimal weights may become very complex. Furthermore, the optimal weights may not be estimable if the optimization window is too short.

2.1.2 Equally Weighted Combination of Area–Wide GDP Forecasts

A simplification of the optimal combination approach is the use of equal weights, which particularly solves the computation problem. Concerning the forecast performance of equally weighted combinations, Timmermann (2005) derives conditions, under which the simple average of a number of forecasts outperforms single model based forecasts as well as more elaborated weighting schemes. Among others, Stock and Watson (2004) provide evidence for the superiority of the equal weighting scheme in a broad empirical application, thereby confirming the so–called *forecast combination puzzle*.

2.1.3 Aggregation of National GDP Forecasts

According to Marcellino, Stock, and Watson (2003) the third strategy is to aggregate national real GDP forecasts to a single euro area real GDP forecast. The following forecasting model is estimated for each member country of the monetary union $i = 1, \ldots, K$ over the period $t = 1, \ldots, T$:

$$Y_{i,t} = \delta_i + c_i X_{i,t} + \varepsilon_{i,t}. \tag{4}$$

Building on these equations, forecasts of euro area real GDP growth are generated by computing weighted averages of the national predictions:

$$\widehat{Y}_{T+1|T+1} = \sum_{i=1}^{K} \omega_i \widehat{Y}_{i,T+1|T+1} = \sum_{i=1}^{K} \omega_i \widehat{\delta}_i + \sum_{i=1}^{K} \omega_i \widehat{c}_i X_{i,T+1}, \tag{5}$$

where ω_i are economic weights (e.g. GDP shares), reflecting the relative importance of country i in the monetary union.

In contrast to the optimal combination approach, the weighting of national information $X_{i,t}$ is not derived from the minimization of the variance of the out–

of–sample forecast error, but is influence by both, the in–sample fit of the disaggregate model for country i and the economic weight of country i (see equation (5)). As before, the approach hardly suffers from the problem of overfitting. However, due to the need to specify of a large number of parsimonious models, it faces the drawback of larger specification errors when the data generating process is unknown.

2.2 Pooling of Information

Pooling of information generates a projection of euro area quarterly real GDP by using area–wide indicators as predictors that combine all national information. In contrast to the multi–equation approaches of forecast pooling, pooling of information thereby reduces the number of regressions to one and as a consequence the problem of running into specification errors. The crucial issue of the pooling of information approach is again the weighting scheme applied to derive area–wide indicators from the national indicator series.

2.2.1 Economic Weights

A straightforward strategy is to use area–wide indicators – that are calculated as economically weighted averages of the national indicator series – as regressors of the forecasting model:

$$Y_t = \delta + cX_t + \varepsilon_t, \tag{6}$$

with $t = 1, \dots, T$. The area–wide indicator X_t is computed as a weighted average of national information variables:

$$X_t = \sum_{i=1}^{K} \omega_i X_{i,t}, \tag{7}$$

where the ω_i's typically reflect country i's relative economic weight in the currency area.

Employing economic weights to construct a single aggregate indicator series implies that these weights are exogenously given. Thus, any correlation between the national indicator series is ignored. Furthermore, the approach does not take any correlations between the resulting indicator series and the area–wide target variable into account.

2.2.2 OLS Weights

The use of OLS weights circumvents this drawback. Estimating the forecast model:

$$Y_t = \delta + \sum_{i=1}^{K} c_i X_{i,t} + \varepsilon_t \tag{8}$$

over the period $t = 1, \ldots, T$, the weighting of national information is given by the point estimates for c_i, which are derived from the minimization of the in–sample residuals. Thus, the in–sample fit of this approach with respect to the aggregate target variable must be superior to a multiple equation approach (see Section 2.1.3). The problem of this approach is, however, that with an increasing number of disaggregate information variables K, the regression model more likely suffers from overfitting. As overparametrization leads to higher estimation uncertainty in finite samples, the out–of–sample performance of the OLS weighting approach is likely to worsen.

2.2.3 Factor Models

The use of Factor models attempts to mitigate the problem of parameter proliferation. While the forecasting model has the same structure as in equation (6), it is preceded by a factor model that pools disaggregate information over the estimation window $[1, T]$ to a common factor X_t, which is used to forecast the target variable Y_t.

The intuition behind factor models in the context of macroeconomic forecasting is that the co–movement in economic time series, in our case the co–movement in the national indicator series, is arising largely from a small set of common factors or even from a single common factor. A number of estimation techniques have been applied in the literature. The simplest method of constructing latent factors proposed by Stock and Watson (2002) is the static principal components analysis (PCA). In our case, the single common factor thereby corresponds to the first principal component, which accounts for as much of the variability in the disaggregate indicators as possible. The weights ω_i are the squared elements of the eigenvector, which is associated with the first principal component. If the resulting common factor explains a large part of the variance of $X_{i,t}$, then $X_{i,t}$ is attributed a high weight.

In the context of business cycle analysis a useful extension of the static version of the factor model is the generalized dynamic factor model of Forni, Hallin, Lippi, and Reichlin (2000), which takes into account phase differences between disaggregate indicator time series by appropriately weighting leading and lagging variables. Kapetanios and Marcellino (2006) propose a state–space model as an alternative and flexible technique to estimate the dynamic common factors. The advantage of factor models is that information of a possibly large set of indicators is pooled by taking into account the in–sample covariances between the candidate regressors. The main drawback of the factor model is that the construction of the common factor ignores any correlation between the common factor X_t and the aggregate target variable Y_t. Thus, the weighting of national information only reflects in–sample correlation patterns between disaggregate indicators and is independent of the forecasting model.

3 Optimal Pooling of Information

In the optimal pooling of information approach forecasts of euro area quarterly real GDP are derived from area–wide indicators that are constructed from national indicator series by using optimal weights, which minimize the variance of the out–of–sample forecast errors of the aggregate target variable. The procedure involves a non–linear numerical optimization routine, which accounts for correlations between both, the disaggregate indicator series and the aggregate target series.

The determination of the optimal weights involves the following steps. We begin with an initial guess for the weights $\boldsymbol{\omega} = (\omega_1, ..., \omega_K)'$. We then compute the area–wide indicator X_t according to equation (7) and estimate equation (6) over the period $t = 1, ..., T$. Finally, we compute the out–of–sample forecasts:

$$\widehat{Y}_{T+1|T+1}(\boldsymbol{\omega}) = \widehat{\delta} + \widehat{c}\sum_{i=1}^{K} \omega_i X_{i,T+1|T+1} = \widehat{\delta} + \widehat{c}\boldsymbol{X}'_{T+1|T+1}\boldsymbol{\omega}, \tag{9}$$

with $\boldsymbol{X}_{T+1|T+1}$ being a $K \times 1$ vector containing the national indicators at time

$T+1$, and the related out–of–sample forecast error:

$$\widehat{\varepsilon}_{T+1|T+1}(\boldsymbol{\omega}) = Y_{T+1} - \widehat{Y}_{T+1|T+1}(\boldsymbol{\omega}). \tag{10}$$

The optimal weights then result from the minimization of the variance of the out–of–sample forecast error:

$$\boldsymbol{\omega}^{opt} = \arg\min_{\omega} \left[\left(\widehat{\varepsilon}_{T+1|T+1}(\boldsymbol{\omega}) \right)^2 \right] \text{ s.t. } \sum_{i=1}^{K} \omega_i = 1. \tag{11}$$

Following the idea of sparse and stable portfolio optimization – see e.g. Brodie, Daubechies, De Mol, and Giannone (2007) – we regularize our objective function by restricting the weights to be non–negative and to sum up to unity. Introducing theses conditions leads to a stabilization of the optimization problem and promotes sparse portfolios by attributing a weight of zero to a number of national indicators.

The main advantage of the optimal pooling of information approach is that it takes into account correlations between both, predictors as well as predictors and the target variable. In contrast to other pooling of information strategies these correlations refer to the out–of–sample performance of the forecast model. Thus, the approach can be considered as a solution to the bias–variance trade–off one is confronted with when specifying a forecast model. A major drawback of the approach is that with an increasing number of disaggregate information K the computation of optimal weights may become prohibitively costly or even impossible. One way to circumvent the construction of such a so–called *super model* is to pre–aggregate individual information to national indicator series, which reduces the optimal pooling of information problem to a manageable number of variables (Timmermann, 2005).[2]

[2] In case of consumer or business surveys, the number of disaggregate information K can be very large as the approach could in principle be tracked down to the level of single individuals.

4 Empirical Results

4.1 Forecast Model Specification

Following Banerjee, Marcellino, and Masten (2005), we generate forecasts of euro area quarterly real GDP by using bridge models that are specified as:

$$A(L)Y_t = \delta + B(L)X_t + \varepsilon_t, \tag{12}$$

where Y_t denotes real GDP expressed in quarterly growth rates, δ is a constant term, X_t describes the quarterly values of a business cycle indicator, $A(L)$ and $B(L)$ are lag polynomials and ε_t denotes the error terms.[3] Quarterly projections of real GDP growth are derived by exploiting the timely information contained in the contemporaneous business cycle indicator in addition to the information provided by past realizations.

An important issue in specifying a time series model for forecasting purposes is the choice of the number of lags of the endogenous and exogenous variables included as regressors. In order to avoid the drawbacks of the lag selection approaches available – i.e. in–sample and out–of–sample criteria – and to account for model selection uncertainty, we do not restrict the forecasting models to a certain lag length but employ a model averaging strategy that allows for different lag orders. Thus we follow the idea that it is a priori impossible to discard a certain lag order from the forecast exercise. Since simple pooling schemes thereby perform comparably well,[4] we derive forecasts from a business cycle indicator within each forecasting model by allowing for a certain maximum number of lags of the exogenous and the endogenous variables. The different model specifications are then built by permutating the candidate regressors and imposing that the contemporaneous value of the business cycle indicator forms part of each model. One–step ahead forecasts of each model specification are derived and combined using equal weights.

[3]Notice that in cases where national information enters the bridge model equation (12) and/or if national real GDP growth rates are used as dependent variables, the following model specification applies: $A(L)Y_{i,t} = \delta + B(L)X_{i,t} + \varepsilon_{i,t}$.

[4]Recall the discussion in Section 2.1.2.

4.2 Data Set

Our data set includes real GDP and several business cycle indicators. The data is collected for both, the euro area and the member states, over the period from 1990Q1 to 2007Q2. Real GDP from 1995Q1 on is taken from the OECD's Main Economic Indicators Original Release Data and Revisions Database that comprises vintage data, which is published each month since January 2000.[5] In order to get a balanced panel of real GDP data, the period from 1990Q1 to 1994Q4 was completed with real GDP data for the member countries from the final vintage of the OECD database and real GDP data for the euro area from the Area Wide Model of the Euro Area Business Cycle Network (EABCN). Real GDP is seasonally adjusted and converted into quarterly growth rates to satisfy stationarity conditions.

For a business cycle indicator to be selected the following criteria had to be met: (1) It is published both at the area–wide and at the national level. (2) It is a leading or a coincident indicator of economic activity and therefore suited to forecast real GDP growth. (3) The indicator is published quarterly or at a higher frequency. (4) It covers a sufficient time span, starting at least in 1990. (5) It is either not revised or vintage data is available covering the total time span. Keeping these guidelines in mind, we end up with three business cycle indicators, namely the Industrial Production Index (IPI), the Economic Sentiment Indicator (ESI) of the European Commission and the CESifo World Economic Survey (WES).

The IPI provides a measure of the volume of value added generated by production units classified under the industrial sectors, i.e. C (Mining), D (Manufacturing) and E (Electricity, gas and water) of the International Standard Industrial Classification of all Economic Activities (ISIC Rev.3). It is released on a monthly basis so that the quarterly value is derived from the monthly average. In the euro area data are collected by the national statistical offices and aggregated by Eurostat to an area–wide index. The country weights used for the aggregation are

[5]Since real time data for Ireland, Luxembourg and Greece starts considerably later in the OECD database, we excluded these three countries from our data set. The nine remaining countries cover almost 95% of area–wide economic activity.

value added at factor costs; they are revised every five years (Eurostat, 2006). As the indicator is subject to data revisions, vintage data is provided by the OECD's Main Economic Indicators Original Release Data and Revisions Database from 1990 onwards.

The ESI combines the weighted information contained in confidence indicators of different sectors – namely industry, services, construction, retail trade and consumers – that are in turn constructed from survey data. Since the indicator is published on a monthly basis, the quarterly value is computed as an average of the monthly releases within the survey quarter. The ESI is built in two steps. In a first step, the area–wide confidence indicators of each sector are derived by aggregating the individual country sector confidence indicators. The weights are the shares of each of the member states in an area–wide reference series – here GDP growth – and are smoothed by calculating a two year moving average. In a second step, the area–wide confidence indicators are combined by using survey weights, which are based on two criteria: (i) the importance of the corresponding sector in the overall economy, and (ii) the ability of tracking the movements of the reference series (European Commission, 2007).

Finally, the WES summarizes the judgement of economic experts about the current economic situation by revealing their appraisals and expectations. It is exclusively based on qualitative information and is timely released within the survey quarter on a quarterly basis. The WES is collected for each member state of the euro area, whereby the aggregate area–wide index is calculated as a weighted average of the individual country indices. The weighting scheme adopted refers to the share of a single country in total world trade (Stangl, 2007).[6]

4.3 Forecast Experiment Using Ex–Ante Information

We generate forecasts of euro area quarterly real GDP by estimating bridge models for each business cycle indicator recursively. We focus on the entire forecast evaluation window that ranges from $\Theta_0 = 1999Q4$ to $\Theta_2 = 2006Q2$. The projections are derived as nowcasts for every quarter following the end of the estimation

[6]The calculation of the national trade volumes is based on the foreign trade statistic published by the United Nations. The weighting scheme is readjusted once a year.

window T, which is gradually extended from 1999Q3 to 2006Q1.[7]

Since we seek to evaluate the full forecast potential of the optimal pooling of information approach here, the optimization of the weights draws on the $\Theta_2 - \Theta_0 = 27$ out–of–sample forecast errors of the entire forecast evaluation window. As the forecast evaluation window and the optimization window coincide, we explicitly use so–called ex–ante information to optimize the weights, which means that we use information that would not have been available in real time.

For a comparison of the forecast performance of different forecast strategies in a real–time experiment, the optimization window should be separated from the evaluation window in order to avoid any informational advantages. We are not aiming at conducting such a competition here, but rather seek to compare the forecast performance of the optimally pooled area–wide indicator series with that of the corresponding economically weighted indicator series in order to understand the attribution of the weights that result from the optimization algorithm and to quantify the potential gain of optimal pooling of information.

Table 1: Forecast performance of optimally pooled indicators relative to economically weighted indicators

		MSE ratio	HLN p–value
Industrial Production	IPI	0.69	0.05
Economic sentiment	ESI	0.47	0.01
CESifo Economic Climate	WES	0.63	0.01

Notes: The MSE ratios are calculated as the MSE resulting from optimally pooled area–wide indicators relative to the MSE resulting from economically weighted area–wide indicators. The HLN p–value was calculated from a Student's t–distribution with $\Theta_2 - \Theta_0 - 1 = 26$ degrees of freedom.

Analyzing the full forecast potential of optimal pooling of information, the results in Table 1 indicate that forecast accuracy in terms of the out–of–sample MSE calculated over the entire forecast evaluation window is on average improved

[7]Following Zarnowitz and Braun (1992) and Batchelor (2001) we use the release of real GDP, available one year after the end of the respective quarter as the relevant realization for computing the forecast errors. As our data set ranges from 1990Q1 to 2007Q2 the last projection is generated for 2006Q2.

by around 40% compared to the economically weighted indicators. The test of forecast accuracy by Harvey, Leybourne, and Newbold (1997) – denoted HLN hereafter – clearly confirms the significance of the improvement.[8]

For an insight into the composition of the optimally pooled area–wide indicators, Table 2 depicts the weights that the optimization algorithm attributes to the single national indicator series. For the IPI almost all national indicators are considered – the only exception is the Portuguese indicator – while for the ESI and the WES a smaller number of national indicator series are selected.[9] In the IPI high weights are attributed to Germany, France, Italy and Spain, which also constitute the largest economies in the currency area. Likewise in the ESI and the WES a large weight is assigned to Germany, but also to a subset of indicators of smaller countries, such as the Netherlands and Portugal. Surprisingly, for the ESI the Dutch indicator series obtains a weight that lies far above the Dutch share in euro area economic activity, which is currently around 5%. For the WES the same holds for the Portuguese indicator series. Even more surprisingly, despite the eminent economic role of France, Italy and Spain within the euro area, neither in the ESI nor in the WES the French, Italian or Spanish indicators are of prime importance.

In order to analyze why certain national indicators enter the optimally pooled area–wide indicators, we calculated the out–of–sample MSE resulting from area–wide models using only a single national indicator as predictor relative to the MSE resulting from an area–wide model using the economically weighted area–wide indicators. The results are shown in Table 3 in which the best–performing national indicators are marked in bold. A comparison of the relative MSEs with those reported in Table 2 shows that the optimization algorithm attributes the highest weights to those national indicators that exhibit the highest degree of forecast accuracy.

Apart from looking one–dimensionally at the mean forecast error, the theory

[8]The null hypothesis of the HLN test is that the difference between the squared out–of–sample forecast error resulting from optimally pooled area–wide indicators and the squared out–of–sample forecast error resulting from economically weighted area–wide indicators is not less than zero.

[9]Notice that for the ESI information on the Austrian indicator series is not available.

Table 2: Optimal weighting schemes

National Indicator Series	IPI	ESI	WES
Austria	0.04	–	0.00
Belgium	0.04	0.00	0.00
Finland	0.05	0.00	0.00
France	0.19	0.01	0.00
Germany	0.17	0.69	0.52
Italy	0.35	0.00	0.00
Netherlands	0.06	0.30	0.24
Portugal	0.00	0.00	0.19
Spain	0.10	0.00	0.04

Notes: The weights are derived by minimizing the out–of–sample MSE resulting from 27 one–step ahead forecasts.

Table 3: Area–wide ADL–models using single national indicators as predictors

National Indicator Series	IPI	ESI	WES
	MSE ratio		
Austria	1.63	–	1.04
Belgium	1.34	1.42	1.27
Finland	1.75	1.35	1.69
France	1.59	1.55	1.27
Germany	2.11	**0.59**	**0.73**
Italy	**1.11**	1.61	1.70
Netherlands	1.70	0.84	1.23
Portugal	2.06	1.56	1.56
Spain	1.32	2.26	1.45

Notes: The MSE ratios are calculated as the MSE resulting from national indicators relative to the MSE resulting from economically weighted area–wide indicators. MSE ratios in bold label the best performing nation indicators.

Table 4: Correlations of forecast errors of the area–wide models using single national indicators as predictors

IPI	Aus	Bel	Fin	Fra	Ger	Ita	Net	Por	Spa
Aus	1.00	0.68	0.84	0.76	0.46	**0.81**	0.71	0.86	0.87
Bel	0.68	1.00	0.68	0.58	0.53	**0.74**	0.60	0.66	0.70
Fin	0.84	0.68	1.00	0.74	0.36	**0.79**	0.63	0.82	0.83
Fra	0.76	0.58	0.74	1.00	0.57	**0.75**	0.75	0.87	0.82
Ger	0.46	0.53	0.36	0.57	1.00	**0.59**	0.50	0.58	0.58
Ita	0.81	0.74	0.79	0.75	0.59	1.00	0.81	0.93	0.88
Net	0.71	0.60	0.63	0.75	0.50	**0.81**	1.00	0.78	0.76
Por	0.86	0.66	0.82	0.87	0.58	0.93	0.78	1.00	0.91
Spa	0.87	0.70	0.83	0.82	0.58	**0.88**	0.76	0.91	1.00

ESI	Aus	Bel	Fin	Fra	Ger	Ita	Net	Por	Spa
Aus									
Bel		1.00	0.55	0.69	0.61	0.66	0.70	0.63	0.41
Fin		0.55	1.00	0.83	0.58	0.82	0.60	0.70	0.79
Fra		0.69	0.83	1.00	**0.55**	0.87	0.59	0.76	0.85
Ger		0.61	0.58	0.55	1.00	0.65	0.47	0.58	0.48
Ita		0.66	0.82	0.87	0.65	1.00	0.71	0.78	0.85
Net		0.70	0.60	0.59	**0.47**	0.71	1.00	0.70	0.53
Por		0.63	0.70	0.76	**0.58**	0.78	0.70	1.00	0.77
Spa		0.41	0.79	0.85	**0.48**	0.85	0.53	0.77	1.00

WES	Aus	Bel	Fin	Fra	Ger	Ita	Net	Por	Spa
Aus	1.00	0.92	0.92	0.91	0.71	0.92	0.90	0.89	0.88
Bel	0.92	1.00	0.87	0.93	0.74	0.88	0.81	0.85	0.88
Fin	0.92	0.87	1.00	0.94	0.63	0.97	0.96	0.91	0.93
Fra	0.91	0.93	0.94	1.00	0.76	0.96	0.90	0.88	0.94
Ger	0.71	0.74	0.63	0.76	1.00	0.66	0.62	0.55	0.66
Ita	0.92	0.88	0.97	0.96	0.66	1.00	0.94	0.91	0.95
Net	0.90	0.81	0.96	0.90	**0.62**	0.94	1.00	0.86	0.87
Por	0.89	0.85	0.91	0.88	**0.55**	0.91	0.86	1.00	0.86
Spa	0.88	0.88	0.93	0.94	**0.66**	0.95	0.87	0.86	1.00

Notes: Figures in bold label the national indicator series that enter the newly constructed area–wide indicators.

of portfolio optimization highlights the role of correlations for the determination of the optimal weighting scheme. An analysis of the correlations of the forecast errors resulting from area–wide models that only use a single national indicator as predictor, might in particular be helpful in explaining why some of the rather poorly–performing national indicator series enter the optimally pooled indicators in addition to the best performing national indicators. Table 4 reveals that the optimization algorithm attributes a weight larger than zero to those national indicator series whose forecast errors are only little correlated with the best–performing national indicators. Consider the Dutch and the Portuguese WES indicators as an example. Although they perform rather poorly when it comes to forecasting euro area real GDP growth – i.e. their MSE ratios are greater than the MSE ratio for the Austrian indicator – they obtain a weight that is far greater than the relative economic share of their economies in the euro area, simply because the correlation between their forecast errors and those resulting from the forecasting model using the German indicator as a predictor are among the lowest.

4.4 Forecast Experiment Using Ex–Post Information

The critical point of the optimal pooling of information approach is the use of the out–of–sample MSE as the target function of the optimization algorithm since this requires to rely on ex–ante information. By exploiting information stemming from the forecast evaluation window, the approach is advantaged compared to competing forecasting methods in a real–time forecast experiment. To overcome this drawback it has to be shown that the optimal weights reported in Table 2 remain stable over time. In this context stability means that the weights attributed to each national indicator series are robust against variations of the length and the initial date of the optimization window.

In the following we derive optimal weights by focusing on shorter optimization windows that are strictly separated from the evaluation window. Figure 2 presents for an overview of the timing of events. Table 5 shows the optimal weights that are computed from rolling optimization windows with 10 and 15 forecast errors. Given that the number of potential out–of–sample forecasts in the experiment is equal to $\Theta_2 - \Theta_0 = 27$, we end up with 18 and 13 fixed–length optimization windows, which can be used to derive the weights. A comparison of Table 5 with Table 2 shows that the mean of the weights is similar to the weights computed from the complete optimization window and the variation, measured in terms of standard deviations, decreases with an increasing optimization window.

Table 5: Optimal weighting schemes derived from rolling optimization windows with 10 and 15 forecast errors

	window size 10			window size 15		
	IPI	ESI	WES	IPI	ESI	WES
Aus	0.03	–	0.00	0.00	–	0.00
	(0.04)		(0.00)	(0.00)		(0.00)
Bel	0.15	0.03	0.03	0.15	0.04	0.00
	(0.13)	(0.09)	(0.07)	(0.05)	(0.09)	(0.00)
Fin	0.03	0.01	0.00	0.03	0.05	0.00
	(0.04)	(0.04)	(0.00)	(0.03)	(0.09)	(0.00)
Fra	0.17	0.05	0.00	0.26	0.01	0.00
	(0.13)	(0.07)	(0.00)	(0.12)	(0.02)	(0.00)
Ger	0.17	0.51	0.45	0.10	0.50	0.44
	(0.09)	(0.29)	(0.26)	(0.09)	(0.13)	(0.16)
Ita	0.26	0.02	0.00	0.34	0.00	0.00
	(0.10)	(0.03)	(0.00)	(0.06)	(0.00)	(0.00)
Net	0.12	0.15	0.19	0.09	0.24	0.27
	(0.08)	(0.16)	(0.09)	(0.04)	(0.14)	(0.08)
Por	0.02	0.17	0.26	0.00	0.14	0.24
	(0.03)	(0.22)	(0.22)	(0.01)	(0.14)	(0.16)
Spa	0.05	0.07	0.07	0.02	0.03	0.05
	(0.09)	(0.18)	(0.06)	(0.04)	(0.05)	(0.04)

Notes: The Table shows the mean of 18 and 13 optimal weights derived from rolling optimization windows with 10 and 15 forecast errors. The figures in parentheses denote the standard deviations around the mean.

Although economic weights pose the most popular aggregation scheme, a number of alternative benchmark models are at disposal. In order to take our optimal pooling approach to a tougher test we compare its forecast accuracy in the following with the competing economic and econometric weighting schemes and prediction approaches presented in Section 2 in the real–time experiment. In addition, we also derive forecasts from an univariate forecast model. The competing forecast models are thereby estimated using the same area–wide and national

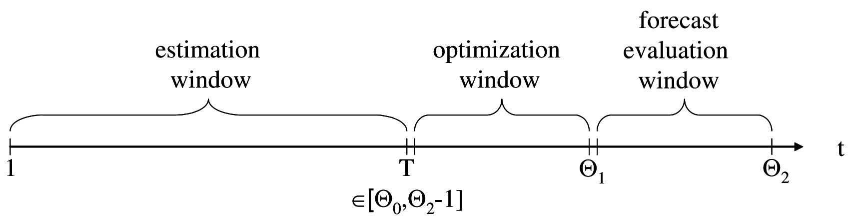

Figure 2: Time structure of the estimation and forecasting procedures

business cycle indicators at disposal. The optimized weights are derived from a recursively growing optimization window, which is excluded from the forecast evaluation process.[10] For the first iteration, the optimized indicator is calculated by minimizing the sum of the first 10 out–of–sample squared forecast errors and the forecast of euro area real GDP growth for second quarter 2002 is generated. At each iteration, the optimization window is expanded one quarter and the weights are updated using a recursive approach. The same setting is used to derive the weights for the optimal pooling of area–wide forecasts as described in Section 2.1.1 in detail. Again, the weighting scheme is solely based on ex–post information. Furthermore, at each iteration a static factor model as well as a dynamic factor model are employed to extract an area–wide indicator which is used to forecast current quarter's real GDP growth for the euro area. The area–wide indicator thereby corresponds to the first common factor extracted.[11] For the aggregation of national GDP forecasts, we employ economic weights based on the relative nominal GDP within the euro area.

Table 6 reports the forecast MSE of the optimal pooling of information approach relative to those of the alternative forecast approaches. The results can be summarized as follows: 1) The optimal pooling of information approach results in general in a lower forecast error, i.e. the MSE ratios are below unity. Only in two cases – for the IPI and the WES – the MSE ratios are above unity. 2) The optimal

[10]Note that we used the first estimate of real GDP growth as realization to calculate the loss function that is minimized to generate the optimized indicators.

[11]The number of common factors extracted as well as the lag–window size used in the dynamic factor model are optimized regarding the ex–post forecast performance of the resulting area–wide indicator.

Table 6: Forecast performance of optimal pooling of information relative to traditional forecast strategies

		MSE ratio	HLN p–value
IPI	Univariate approach	0.61	0.03
	Pooling of information		
	Economic weights	0.95	0.38
	OLS weights	0.37	0.03
	Principal component analysis	1.12	0.82
	Dynamic factor model	0.67	0.09
	Pooling of forecasts		
	Optimal weighting of area–wide forecasts	0.74	0.07
	Equal weighting of area–wide forecasts	0.82	0.13
	Aggregation of national forecasts	0.92	0.36
ESI	Univariate approach	0.44	0.03
	Pooling of information		
	Economic weights	0.63	0.04
	OLS weights	0.66	0.14
	Principal component analysis	0.67	0.05
	Dynamic factor model	0.89	0.33
	Pooling of forecasts		
	Optimal weighting of area–wide forecasts	0.90	0.16
	Equal weighting of area–wide forecasts	0.65	0.07
	Aggregation of national forecasts	0.70	0.11
WES	Univariate approach	0.48	0.06
	Pooling of information		
	Economic weights	0.64	0.03
	OLS weights	1.11	0.60
	Principal component analysis	0.60	0.01
	Dynamic factor model	0.71	0.07
	Pooling of forecasts		
	Optimal weighting of area–wide forecasts	0.91	0.30
	Equal weighting of area–wide forecasts	0.60	0.08
	Aggregation of national forecasts	0.71	0.13

Notes: The MSE ratios are calculated as the MSE resulting from the optimal pooling of information approach relative to the MSE resulting from respective forecast strategy. The HLN p–value was calculated from a Student's t–distribution with $\Theta_2 - \Theta_0 - 1 = 16$ degrees of freedom.

pooling of information approach outperforms the economic weighting schemes for all three indicators under consideration, which confirms the results obtained in Section 4.3 where we allowed for the use of ex–ante information. However, the improvement is significant only in the cases of the WES and the ESI. Likewise, the optimal pooling of information approach dominates the univariate forecast model significantly in all cases. 3) The HLN test shows at the 10% significance level that the optimal pooling of information approach significantly dominates 5 of the competing forecast approaches in the case of the IPI and the ESI, and 4 of the competing forecast approaches in the case of the WES. 4) In those cases where the MSE ratio is above unity, the optimal pooling of information approach is not systematically beaten by the competing forecast method.

5 Conclusion

This paper proposes a new method of forecasting euro area quarterly real GDP that uses area–wide indicators, which are derived by optimally pooling the information contained in national indicator series. The area–wide indicators are computed by applying weights that minimize the variance of the out–of–sample forecast error of the aggregate target variable. We evaluate the forecast performance of our optimal pooling of information approach by focusing on three business cycle indicators, namely the Industrial Production Index (IPI), the economic sentiment indicator (ESI) of the European Commission and the CESifo World Economic Survey (WES) indicator for the euro area, which are all available at the area–wide and country–specific level.

Our results show that short–term forecasts of euro area quarterly real GDP are improved by using area–wide indicators based on optimal weights rather than economic weights. The optimally pooled area–wide indicators reduce the out–of–sample MSE by 40% on average. We also demonstrate that the optimal weights are relatively robust against changes in the length of the optimization window, which promotes a certain stability of the optimal weighting scheme lending support to the practicability of our approach in real–time.

In an out–of–sample forecast experiment we compare the forecast performance of the optimal pooling of information approach with that of a number of competing forecasting approaches. The optimally pooled area–wide indicators are constructed using only information that would have been available in real–time. We find that our method of forecasting performs well compared to the competing forecasting methods in terms of forecast accuracy.

References

Banerjee, A., M. Marcellino, and I. Masten (2005): "Leading Indicators for Euro-area Inflation and GDP Growth," *Oxford Bulletin of Economics & Statistics*, 67, 785–813.

Batchelor, R. (2001): "How Useful are the Forecasts of Intergovernmental Agencies? The IMF and OECD versus the Consensus," *Applied Economics*, 33, 225–235.

Bates, J., and C. W. J. Granger (1969): "The Combination of Forecasts," *Operations Research Quarterly*, 20(4), 451–468.

Brodie, J., I. Daubechies, C. De Mol, and D. Giannone (2007): "Sparse and Stable Markowitz Portfolios," CEPR Discussion Papers 6474.

Diebold, F. X., and J. A. Lopez (1996): "Forecast Evaluation and Combination," in *Handbook of Statistics*, ed. by G. S. Maddala, and C. R. Rao, vol. 14, pp. 241–268. Kluwer Academic Publishers.

Domenico, G., L. Reichlin, and D. Small (2006): "Nowcasting GDP and Inflation: The Realtime Informational Content of Macroeconomic Data Releases," ECB Working Paper Series 633.

European Commission (2007): "The Joint Harmonised EU Programme of Business and Consumer Surveys: User Guide," Economic studies and research, European Commission.

Eurostat (2006): "Methodology of Short-term Statistics: Interpretation and Guidelines," Methods and nomenclatures, Eurostat.

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(4), 540–554.

Granger, C. W. J., and R. Ramanathan (1984): "Improved Methods for Combining Forecasts," *Journal of Forecasting*, 3(2), 197–204.

Harvey, D., S. Leybourne, and P. Newbold (1997): "Testing the Equality of Prediction Mean Squared Errors," *International Journal of Forecasting*, 13(2), 281–291.

Kapetanios, G., and M. Marcellino (2006): "A Parametric Estimation Method for Dynamic Factor Models of Large Dimensions," CEPR Discussion Papers 5620.

Lütkepohl, H. (1987): *Forecasting Aggregated Vector ARMA Processes*. Springer Verlag.

Marcellino, M., J. Stock, and M. Watson (2003): "Macroeconomic Forecasting in the Euro Area: Country Specific versus Area-Wide Information," *European Economic Review*, 47(1), 1–18.

Stangl, A. (2007): "World Economic Survey," in *Handbook Of Survey-Based Business Cycle Analysis*, ed. by G. Goldrian, pp. 57–67. Ifo Economic Policy Series, Edward Elgar.

Stock, J., and M. Watson (2002): "Macroeconomic Forecasting Using Diffusion Indexes," *Journal of Business and Economic Statistics*, 20(2), 147–62.

——— (2004): "Combination Forecasts of Output Growth in a Seven-country Data Set," *Journal of Forecasting*, 23(6), 405–430.

Timmermann, A. (2005): "Forecast Combinations," in *Handbook of Economic Forecasting*, ed. by G. C. W. J. Elliott, G., and A. Timmermann. North Holland.

Zarnowitz, V., and P. Braun (1992): "Twenty–Two Years of the NBER–ASA Quarterly Outlook Surveys: Aspects and Comparisons of Forecast Performance," NBER Working Papers 3965.