On the optimal number of indicators – nowcasting GDP growth in CESEE

David Havrlant,
Peter Tóth,
Julia Wörz[1]

We employ principal components and dynamic factor models for nowcasting GDP growth in selected Central, Eastern and Southeastern European (CESEE) economies. Our estimation sample extends from the first quarter of 2000 to the second quarter of 2008, our evaluation period from the third quarter of 2008 to the third quarter of 2014. For this period, we produce quasi out-of-sample forecasts of past-, current- and next-quarter GDP growth for seven CESEE economies. The models differ with respect to the estimation method, model specification, and the number of short-term indicators used. We find, first of all, a clear gain in predictive accuracy from using a nowcasting model with monthly indicators compared to the naïve benchmark. Furthermore, for our sample of small, open economies, we find that models using a smaller set of carefully selected indicators yield lower prediction errors on average than models based on larger information sets. Finally, we identify a clear gain in forecast performance from including foreign or euro area indicators.

JEL classification: C52, C53, E37
Keywords: nowcasting, dynamic factor models, principal components, Central, Eastern and Southeastern Europe

One of the most important indicators of economic activity – GDP – is reported with a considerable time lag and at a rather low frequency. In the EU, a first, so-called "flash" estimate of GDP is not released until six weeks after the end of a quarter; the second estimate (including information on the components of GDP as well) is announced with a delay of almost one quarter (11 weeks). The resulting information gap can be filled by making use of higher frequency indicators in the time between the end of the reporting period and the publication of official GDP figures.

For large economies (U.S.A., the U.K. and the euro area), large-scale models have been developed to make use of this high-frequency information. Since the pioneering work of Evans (2005), Nunes (2005) and Giannone et al. (2008), it has become common to rely on computational estimates of GDP in real time. The menu of available model classes ranges from single-indicator, regression-based bridge equations to highly complex, multi-indicator dynamic factor models (DFMs).

Yet, for Central, Eastern and Southeastern European (CESEE) economies, considerably fewer indicators have traditionally been available, and the transition history of shorter time series has often precluded the use of such computationally intensive models. For instance, Rünstler et al. (2009) report that models for three Eastern European EU Member States (Lithuania, Hungary and Poland) performed rather badly with respect to naïve benchmarks in their analysis. They used data

[1] *European Commission, david.havrlant@ec.europa.eu; Národná Banka Slovenska Research Department, peter.toth@nbs.sk; Oesterreichische Nationalbank, Foreign Research Division, julia.woerz@oenb.at (corresponding author). The opinions expressed by the authors of this study do not necessarily reflect the official viewpoints of the European Commission, the Eurosystem, Národná Banka Slovenska or the Oesterreichische Nationalbank. None of the afore-mentioned institutions or any person acting on their behalf may be held responsible for the use which may be made of the information contained in this publication, or for any errors that may remain. The authors would like to thank Sebastian Beer, Martin Feldkircher, two anonymous referees, participants of the May 2016 Joint Annual Meeting of the Austrian Economic Association and the Slovak Economic Association (NOeG-SEA) in Bratislava and participants of the June 2016 CESEEnet workshop on Macroeconomic Forecasting in Vienna for helpful comments and valuable suggestions.*

starting in January 1995, 1997 and 1998, respectively, and ending in July or August 2006. Hence, they could not assess whether this bad performance resulted from the short dataset or other characteristics specific to the catching-up experience of these countries.

In time, such restraints eased. As a result of EU accession roughly ten years ago and the associated Eurostat reporting commitments, the set of high-frequency indicators that are available for a reasonable time period for these small and open economies has grown rapidly, opening up new possibilities for estimating the current level or growth rate of GDP.

In a related paper, Feldkircher et al. (2015) explored small-scale models ranging from bridge equations to dynamic factor models for nowcasting real GDP growth in selected CESEE economies. The analysis was based on a handful of time series that were selected from all available monthly indicators using both very simple and highly sophisticated selection procedures varying from picking the "usual suspects," such as industrial production or Eurostat's Economic Sentiment Indicator (ESI), to applying a Bayesian model averaging approach to narrow down the set of short-term indicators to fewer than 10. The results suggested that a small dynamic factor model based on about six to eight indicators carefully selected according to their correlation with GDP consistently outperformed the benchmark autoregressive model AR(1).

Factor models are powerful tools for extracting relevant information from large datasets. Large factor models allow researchers to include all potentially important information from a data-rich environment (see Barhoumi at al., 2013, for a survey of dynamic factor models). Yet, it is not clear whether enlarging the number of time series also results in a better forecasting performance. Boivin and Ng (2006) find that a smaller number of time series (40 of a maximum of 147 series available to them) can yield better results in a real-time forecasting exercise. This result arises when idiosyncratic errors show cross-correlations in large datasets or when a highly informative factor dominates a small dataset but is dominated in a larger dataset. Proposing different methods to identify relevant or efficient predictors, Bai and Ng (2008) show that forecasting performance improves when factors are estimated using fewer but informative predictors.

Given the discussion on factor models in the literature, the main focus of this paper lies on exploring the optimal number of predictors to be used in a factor model for forecasting the GDP growth of selected CESEE economies. Hence, our horse race is between different dataset classes distinguished by the number of high-frequency indicators. In the present analysis, we focus on two types of factor models, namely principal component models, or approximate factor models, in the spirit of Stock and Watson (2002a, 2002b) and mixed frequency dynamic factor models for large datasets following Bańbura and Modugno (2014). Both types of models have been applied to Czech data before (see Arnoštová et al., 2011, and Rusnák, 2016) and performed well.

We apply those models to estimate GDP growth with a very short-term horizon (last, current and next quarter) for seven CESEE countries: Bulgaria, the Czech Republic, Hungary, Poland, Romania, Slovakia and Slovenia.[2] To ensure full comparability with Feldkircher et al. (2015), we assess relative forecasting performance over the period 2008 to 2014. Hence, our evaluation period covers the global financial crisis and the subsequent recovery. In contrast to most euro area economies that experienced a double-dip recession during those years, some CESEE economies, in particular Poland, where there was no recession at all, recovered quickly and posted rather sound growth rates, especially toward the end of our observation period. Hence, we cover a period including both recessions and expansions, which is preferable for evaluating the quality of a forecast. In general, both principal component models and dynamic factor models tend to outperform our benchmark AR(1) model. While we cannot distinguish easily between the forecasting performance of different model classes and model specifications, we observe a consistently better performance of models relying on a smaller set of 9 to 14 carefully selected indicators.

Section 1 describes the two competing models. Section 2 defines the different indicator sets and the data sample. Section 3 reports the results for individual models and section 4 concludes.

1 Our horses: principal component models v. dynamic factor models

In our analysis, we rely purely on computational methods to predict GDP growth from higher frequency indicators. More specifically, we use factor models. This type of model makes use of timing properties of the higher frequency indicators and can broadly be attributed to one of two model classes.

Principal component models, also called approximate factor models, make use of static factors. The monthly dataset is first rebalanced by lagging some of the time series appropriately to deal with ragged edges in the data. Principal components are extracted either from the monthly time series or after having aggregated monthly indicators to quarterly frequency. In a second step, the principal components are bridged to GDP in a simple ordinary least squares (OLS) equation. To sum up, the principal components approach requires rebalancing the monthly series by lagging and aggregating them (or their common factors) to the quarterly frequency. However, the process of lagging and aggregating may neglect the true dynamic relationships between the monthly series, their common factors and GDP growth.

The latest generation of dynamic factor models (DFMs) can deal with both mixed frequencies and unbalanced datasets without the need to rebalance and aggregate data. The monthly DFM is cast in a state-space framework and is estimated in an iterated fashion. The starting values of the common factors are initialized by principal components from the balanced subsample of the indicators. Then the next steps iterate between estimating parameters conditional on the factors and estimating the factors conditional on the parameters from previous iterations. Once the estimates converge, the missing values of the indicators and monthly

[2] *Given our focus on obtaining good nowcasts for CESEE countries that are relevant from the viewpoint of the Oesterreichische Nationalbank, we do not include the Baltic states in our sample. We also excluded Croatia, as we encountered some problems in using exactly the same set of indicators due to its late EU accession.*

GDP are estimated via the Kalman smoother until the end of the forecast horizon. This procedure takes into account all available information on the uneven edges of the dataset.

1.1 Principal component models

Forecasting output growth by principal components, or by the approximate factor model, was inspired by the work of Stock and Watson (2002a, 2002b). The authors use such a model to forecast four U.S. macroeconomic variables with more than 200 predictors. The four variables forecast are industrial production, personal income, manufacturing sales and employment, which are all available on a monthly basis. Their approach has been applied also to forecasting GDP. For a European cross-country study, see e.g. Rünstler et al. (2009).
Our principal component model can be described by the following equations:

$$x_{it}^{Q} = \lambda_i PC_t^{Q} + \omega_{it} \tag{1}$$

$$y_t^{Q} = \Phi_h PC_{t-h}^{Q} + \psi_t \tag{2}$$

where x_{it}^{Q} is the quarterly aggregate of monthly indicator i and y_t^{Q} is the quarterly growth rate of real GDP. The quarterly aggregates are transformed to be stationary, have zero means and unit variances. The issue of uneven endpoints of the x_{it} series due to differences in publication lags is resolved by shifting each series appropriately. This means rebalancing the panel of indicators so that the last observations of x_{it}^{Q} and y_t^{Q} correspond. Vector PC_t^{Q} contains J common factors estimated by principal components analysis, and λ_i is a vector of J factor loadings specific to each indicator i. The number of factors J is set to one, two or three in alternative specifications. The principal components are estimated either at a monthly frequency (PC-M model) using the balanced indicator set x^{it} only, or at a quarterly frequency, including both x_{it}^{Q} and y_t^{Q} (PC-Q model) in the estimation of PC_t^{Q}. Once the PC_t^{Q} series has been estimated, equation (2) is fitted by OLS to obtain the vector of J coefficients Φ_h. Given the static nature of principal components, we need to lag PC_t^{Q} in equation (2) by h periods to forecast GDP growth on the horizon of h quarters ahead.

The remaining terms in the equations, ω_{it} and ψ_t, are idiosyncratic shocks, which may be serially correlated. The identification of PC_t^{Q} requires further that the cross-correlations across ω_{it} are not "too strong" when the sample size (in terms of the number of indicators and the time dimension) is large (see Stock and Watson, 2002a). In other words, including many predictors may come at the cost of increasing the cross-correlations of idiosyncratic shocks[3] for some series. Therefore, careful variable selection may improve the identification of the common factors and, potentially, the forecasting performance of the model.

To sum up, we use different model specifications that vary by frequency aggregation and the number of principal components that we extract. According to our choice of frequency aggregation, we distinguish between a monthly principal components and a quarterly principal components specification. We consider

[3] *In practice, the shocks could be correlated for some sectoral disaggregates of the same series, turnover versus production indexes for the same sector, export and import turnover for small open economies, different labor market indicators, etc.*

models with one, two and three factors. Hence, we obtain six versions of the static factor model for each country, each forecast horizon and each indicator set.

1.2 Dynamic factor models

Dynamic factor models extract signals from all available information even when several indicators are highly correlated. The first generation of DFMs was estimated by maximum likelihood or Kalman filters and can handle data irregularities, but is limited to using a set of few variables (see Engle and Watson, 1981). The next generation of models estimates the factors by nonparametric principal components (Chamberlain and Rothschild, 1983; Forni and Reichlin, 1998; Forni et al., 2000; Stock and Watson, 2002a, 2002b). While these models can handle short time series in large cross sections, they cannot deal with ragged ends in the data. The third generation of DFMs again approximates factors by principal components and utilizes them in a state-space framework (see Giannone et al., 2008; Rünstler et al., 2009). These models can handle large datasets with data irregularities that are present in a real-time forecasting setting. Finally, the latest generation of DFMs uses an expectation-maximization algorithm to obtain ML estimates of large models that are able to deal with unbalanced datasets (Schumacher and Breitung, 2008; Bańbura and Modugno, 2014). We follow the approach of Bańbura and Modugno (2014) and use a mixed-frequency DFM for large datasets. Rusnák (2016) applied the same model to Czech data.

Our model is specified for monthly variables, where the indicators x_{it} are transformed to stationary processes with zero means and unit variances. Quarterly GDP growth, y_t^Q, is assumed to be observable only in the third month of each quarter, while its values in the first two months are treated as missing. Using the approximation[4] of Mariano and Murasawa (2003), we can decompose y_t^Q as its lagged (unobserved) monthly growth rates y_t as follows:

$$y_t^Q = y_t + 2y_{t-1} + 3y_{t-2} + 2y_{t-3} + y_{t-4} \tag{3}$$

The monthly DFM is specified in a state-space form as a set of measurement equations:

$$x_{it} = \Lambda_i f_t + \varepsilon_{it} \tag{4}$$

$$y_t^Q = \Lambda_y (f_t + 2f_{t-1} + 3f_{t-2} + 2f_{t-3} + f_{t-4}) + u_t + 2u_{t-1} + 3u_{t-2} + 2u_{t-3} + u_{t-4} \tag{5}$$

where the second line comes from (1) and the expression below:

$$y_t = \Lambda_y f_t + u_t \tag{6}$$

where f_t are the unobserved common factors for the indicators and GDP growth, Λ_i and Λ_y are the respective factor loadings, and ε_{it} and u_t are idiosyncratic shocks, which may be autocorrelated and weakly cross-correlated.

[4] *This follows from assuming that the level of real GDP in quarter τ (Y_τ^Q) equals the geometric mean of its (unobserved) monthly levels (Y_t). Taking logs, the quarterly first difference of this expression becomes $dlogY_\tau^Q = {}^1/_3(logY_t + logY_{t-1} + logY_{t-2} - logY_{t-3} - logY_{t-4} - logY_{t-5})$.*
Adding and subtracting different lags of $logY_t$ in the above parentheses results in expression (3).

Finally, the state equation defines the dynamics of the common factors as an AR(p) process:

$$f_t = A_1 f_{t-1} + A_2 f_{t-2} + \ldots + A_p f_{t-p} + \upsilon_t \tag{7}$$

where υ_t is an idiosyncratic shock.

Again, we employ different model specifications, which vary by the assumption we make on the idiosyncratic component υ_t (serially uncorrelated versus the AR(1) specification) and by the number of extracted factors f_t (up to four). Hence, we obtain eight different model specifications of the dynamic factor model for each country, forecast horizon and indicator set.

2 Horse feed and race track: data sample and forecast horizon

Our set of available indicators comprises 69 country-specific indicators and 6 foreign indicators. The domestic indicators comprise information for the total economy and individual subsectors on industrial production, turnover, business and consumer surveys, economic sentiment, energy supply, prices, unemployment and international trade. The foreign indicators are the ECB commodity price index, the index of world market prices of the Hamburg Institute of International Economics (HWWI), the HWWI crude oil price index, production in euro area industry, the Markit PMI™ (Purchasing Managers' Index™) for the euro area and the CES-Ifo Export Expectations index). All these indicators are at monthly frequency. Overall, a set of 75 indicators is available for each country model. Guided by the consideration that small indicator sets may prove useful also for DFMs when the time series dimension is short (and hence asymptotic properties do not hold) and that the variability of idiosyncratic components is small, and recalling the satisfactory forecasting performance of the small DFMs reported by Feldkircher et al. (2015) for the same dataset, we run the estimation on five different sets of indicators that vary by coverage. Our large indicator set comprises all 75 indicators. Our medium set contains only selected indicators from the main categories (production, turnover, consumer sentiment, etc.). This set includes 31 indicators. Moreover, we use one variant of the medium-sized set that excludes all foreign variables, reducing it to 26 indicators. Finally, we diminish the number of indicators even further based on their correlation with GDP, using the same standard set of indicators for all countries. The small indicator set contains 14 indicators. Again, we differentiate between a small set including foreign variables and a small domestic set based on nine country-specific indicators. Detailed information on the indicator sets is given in annex table A1.

Our sampling period extends from the first quarter of 2000 to the third quarter of 2014. We discarded data prior to 2000 to be able to work with a meaningful number of indicators readily available from Eurostat. In mid-1995, only 7 indicators are available from this data source; in mid-1996, this number jumps to 27, at the beginning of 1998 to 37, in January 2000 to 50 and to finally to 68 in mid-2002. As is standard in the literature, we focus on indicators reflecting real economic activity and economic sentiment and do not include financial or capital flow data.[5] All models are estimated for the period from the beginning of the sample to the second quarter of 2008. Our evaluation period runs from the third quarter of 2008 to the end of the sample period. For this period, we obtain so-called "quasi out-of-sample" forecasts. We measure forecasting accuracy by the root mean square error (RMSE).

Different frequencies for the explanatory variables and the dependent variable result in a total of eight forecast horizons. For every month in a quarter, we produce a backcast for the GDP of the previous quarter, a nowcast of the current quarter's GDP growth and a forecast of the next quarter as represented in table 1.

Table 1

Forecast horizons

Month in which forecast is made	Month 1			Month 2			Month 3	
Quarter for which GDP is predicted	Q_{t-1}	Q_t	Q_{t+1}	Q_{t-1}	Q_t	Q_{t+1}	Q_t	Q_{t+1}
Label of forecast horizon	Back_1	Now_1	For_1	Back_2	Now_2	For_2	Now_3	For_3

Source: Authors' compilations.

Note that we extract monthly data in the middle of every month. We define calendar months according to their position within a quarter, such that January, April, July and October are labeled "month 1." Hence, in the first and second months of a quarter, we do not even know GDP growth of the previous quarter. Therefore, for these months, we predict a backcast, a nowcast and a forecast, respectively. We accordingly label the predictions obtained from information in month 1 Back_1, Now_1 and For_1. For example, a prediction of first-quarter GDP growth made in April is called Back_1, while the prediction of second-quarter GDP growth made in the same month is Now_1. Likewise, in month 2 we obtain the predictions Back_2, Now_2 and For_2. Continuing the above example, the "forecast" (or better backcast) of first-quarter GDP growth which we obtain in May is labeled Back_2, while the estimate of second-quarter GDP growth in May is called Now_2, and so on. In month 3, we already have an official GDP estimate for the previous quarter. Hence, we do not estimate a backcast in these months. This implies that in month 3, we predict only current and next-quarter GDP growth (horizons Now_3, For_3).

3 The race: forecast accuracy of competing models

Having laid out all these preliminaries, we are now ready in this section to report the results. We estimate three different models for each of the seven countries and each indicator set. Beside the principal component model and the DFM, we also estimate a simple AR(1) model of GDP growth for each country; it serves as our benchmark. Furthermore, we run different specifications of each model, as explained in section 1. In total, we obtain 15 model specifications (6 for the prin-

[5] *Moreover, information on financial or capital flows would not be available from a common data source, which would render a frequent and automatized updating routine complicated. As the aim of this analysis is to provide a sound basis for implementing a nowcasting tool at the OeNB, we opted for harmonized and common data sources across all countries.*

cipal components, 8 for the DFM and one benchmark model) for 7 countries and 5 indicator sets. From each of these roughly 500 model specifications, we obtain a prediction for each of the 8 horizons.[6]

We report the forecasting accuracy of the best-performing model specification for each country, indicator set and forecast horizon in chart 1.[7] Forecasting accuracy is measured by the RMSE relative to the benchmark AR(1) model. Since real-time GDP data series are not available for some of the countries in our sample, we use the latest available GDP growth figures to calculate forecasting errors. Thus, in measuring forecasting accuracy of our quasi out-of-sample forecasts, we ignore the impact of different data vintages on the results.

Chart 1 distinguishes between the results obtained by the two model classes, principal component models and DFMs. The results suggest that both models outperform the naïve benchmark, which models GDP as a simple autoregressive process of order 1. Hence, model-based predictions using higher frequency indicators pay off by producing higher forecasting accuracy. This result is in line with Rünstler et al. (2009). However, relative model performance varies by country. Picking the best-performing specification for each estimation method, we obtain the lowest forecast error on average for Bulgaria, followed by the results for the Czech Republic, Romania and Slovenia. Model performance is worst for Poland and Slovakia.

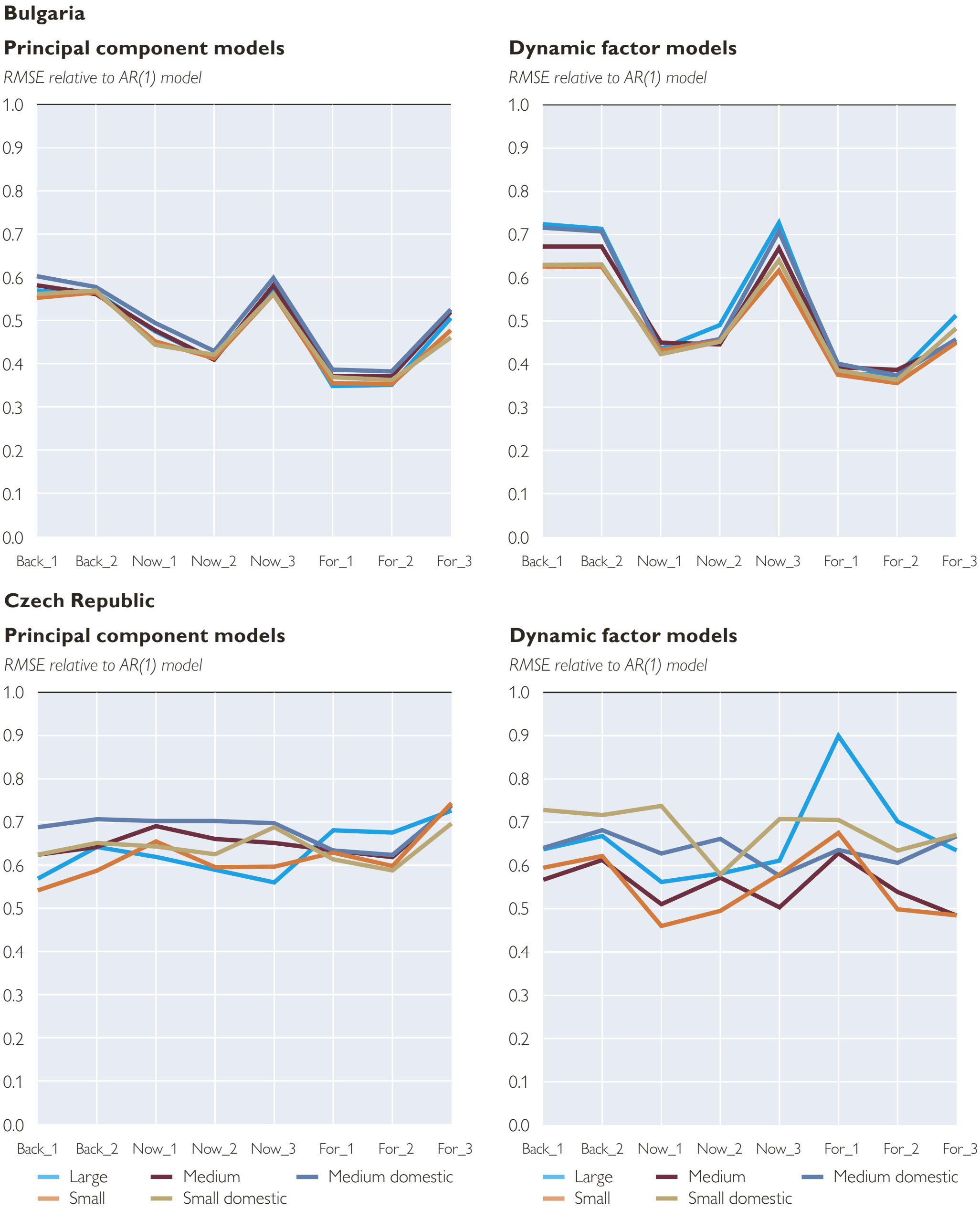
At the same time, forecast accuracy varies considerably across forecast horizons. Not very surprisingly, backcasts show on average smaller RMSEs, while forecasts exhibit the highest RMSEs. For Hungary, Poland, Slovenia and Slovakia, the AR(1) model even outperforms our best model specification for some horizons. Inferior model performance – indicated by a value greater than one in the chart – is observed for forecasts produced by the DFM for Hungary (for horizons For_1 and For_2), for Poland (all nowcasts, For_1 and For_2) and Slovenia (For_1 and For_2). For Slovakia, model performance is rather poor in general; only the small and medium-sized principal component model as well as the small DFM manage to outperform the AR(1) model for some horizons.

[6] *Note that not every DFM specification could be estimated for each country because data availability varied across countries. Thus, the total number of predictions for all countries, indicator sets and horizons is 3,569.*

[7] *Detailed information on all model specifications is available from the authors on request. The best-performing model was chosen as that with the lowest prediction error.*

Chart 1

Relative forecast accuracy by country and model specification



Source: Authors' calculations.

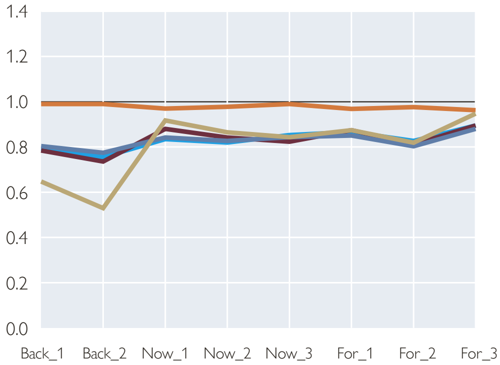
Chart 1 (continued)

Relative forecast accuracy by country and model specification

Hungary

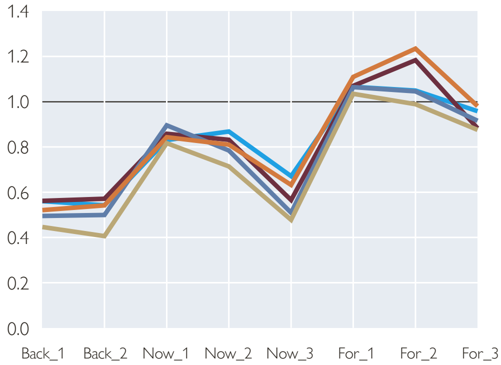
Principal component models

RMSE relative to AR(1) model



Dynamic factor models

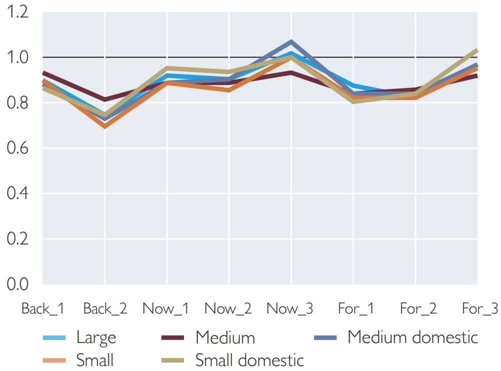
RMSE relative to AR(1) model



Poland

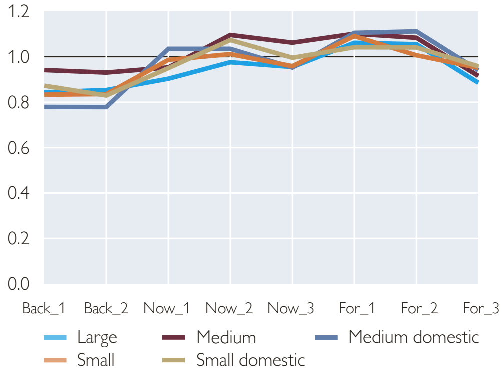
Principal component models

RMSE relative to AR(1) model



Dynamic factor models

RMSE relative to AR(1) model



Source: Authors' calculations.

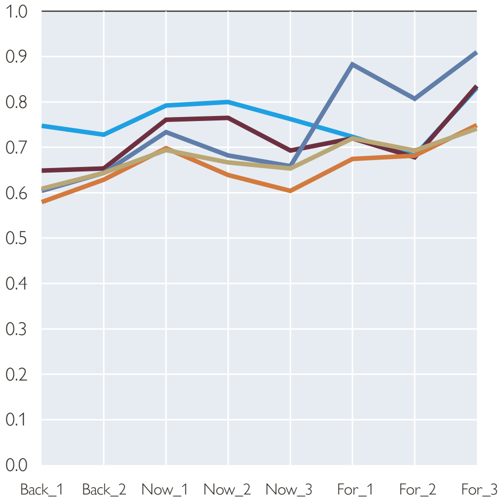
Chart 1 (continued)

Relative forecast accuracy by country and model specification

Romania

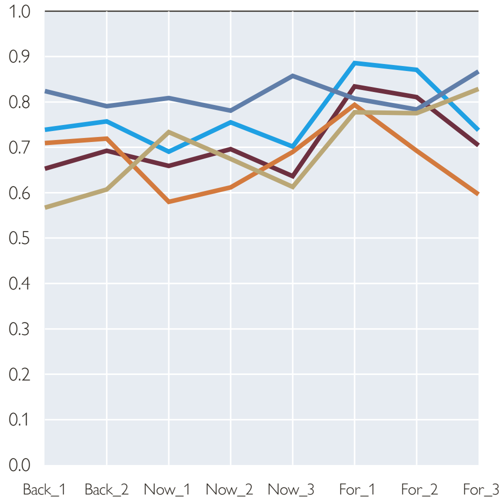
Principal component models

RMSE relative to AR(1) model



Dynamic factor models

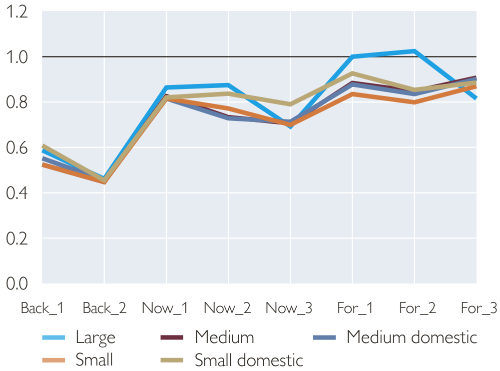
RMSE relative to AR(1) model



Slovenia

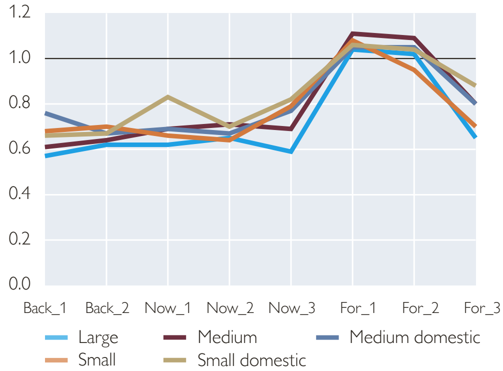
Principal component models

RMSE relative to AR(1) model



Dynamic factor models

RMSE relative to AR(1) model



Source: Authors' calculations.

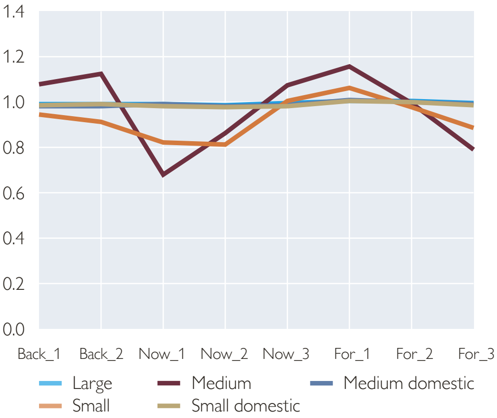
Chart 1 (continued)

Relative forecast accuracy by country and model specification

Slovakia

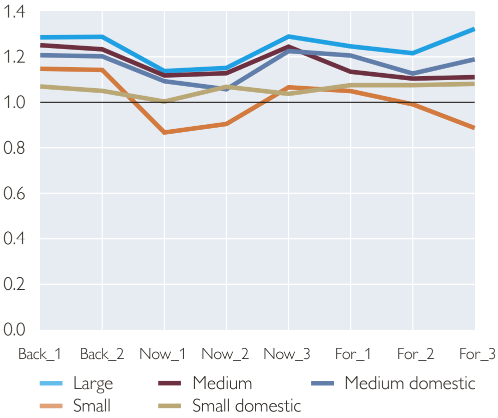
Principal component models

RMSE relative to AR(1) model



Dynamic factor models

RMSE relative to AR(1) model



Source: Authors' calculations.

While we clearly observe differences in model performance across countries and horizons, the match between our two models is less clear-cut. In some cases, especially in Bulgaria, Poland, Romania and Slovakia, the principal component model yields lower RMSEs on average than the DFM. In the Czech Republic and Hungary, the DFM predictions exhibit lower RMSEs, while the results are unclear for Slovenia.

For the practitioner who is confronted with producing a prediction for each country on a monthly basis, the most interesting distinction lies in differences across indicator sets. But knowing whether including more information improves forecast accuracy or rather just adds noise to the forecast is important not just because it has an impact on the amount of data work but also for theoretical reasons. As mentioned above, under certain assumptions,[8] more information always results in more accurate forecasts. However, these assumed conditions are often not met in practice. Chart 1 suggests that the gains in forecast accuracy from varying the size of the indicator set are modest. Eyeball inspection even suggests a slightly better performance of the smallest indicator set of 14 indicators including foreign variables.

We tested this observation by applying a Wald test on the equality of RMSEs across indicator sets. In order to obtain a reasonable test setting, we regressed the absolute RMSE of each model specification on a set of dummy variables, including dummies for the size of the indicator set. Equation 8 reports our test regression:

$$RMSE_{i,c,h,m} = \sum_{I} \alpha_i DIND_i + \sum_{C} \beta_c DCOUNTRY_c + \sum_{H} \gamma_h DHOR_h + \sum_{M} \delta_m DMODEL_m + \varepsilon_{i,c,h,m} \quad (8)$$

[8] *These assumptions mean that asymptotic properties of the indicators must hold, i.e. time series must tend to infinity in terms of number and length. Furthermore, idiosyncratic components must not be strongly correlated, and the variability of the common component needs to be large.*

We include the following dummy variables (labeled D combined with a descriptor for the respective variable) to control for variation in the RMSE that arises from differences in horizon, country, method and model specification. The DIND dummies capture the indicator set used, whereby subscript *i* stands for the size of the variable set (large, medium, medium domestic, small, and small domestic). DCOUNTRY is a set of dummies for each of the seven countries in our sample. DHOR is a set of dummies for each of the eight horizons. Finally, the dummies DMODEL capture the model specification (i.e. six dummies for each variant of the PC model and eight dummies for the DFM models). Equation 8 is estimated by least square dummy variables (LSDV). We then apply Wald tests on

Table 2

Comparison of model RMSEs

		Coefficient	Robust standard error	t–value	P>|t|	95% confidence interval	
						lower bound	upper bound
Indicator set	large	–0.190	0.089	–2.13	0.033	–0.364	–0.015
	medium	–0.281	0.089	–3.16	0.002	–0.456	–0.107
	medium domestic	–0.223	0.088	–2.53	0.011	–0.395	–0.050
	small	–0.369	0.089	–4.14	0.000	–0.543	–0.194
	small domestic	–0.267	0.089	–2.99	0.003	–0.442	–0.092
Country	BG	0.356	0.018	19.88	0.000	0.321	0.391
	CZ	–0.345	0.011	–31.80	0.000	–0.366	–0.323
	HU	–0.296	0.012	–24.08	0.000	–0.320	–0.272
	PL	–0.815	0.012	–65.69	0.000	–0.839	–0.790
	RO	0.587	0.017	35.60	0.000	0.555	0.620
	SK	1.095	0.019	57.62	0.000	1.058	1.132
Horizon	Back_1	1.536	0.082	18.66	0.000	1.375	1.697
	Back_2	1.505	0.082	18.27	0.000	1.343	1.666
	Now_1	1.822	0.084	21.82	0.000	1.659	1.986
	Now_2	1.797	0.083	21.55	0.000	1.634	1.961
	Now_3	1.590	0.082	19.32	0.000	1.429	1.752
	For_1	1.947	0.085	22.84	0.000	1.780	2.115
	For_2	1.917	0.085	22.44	0.000	1.750	2.085
	For_3	1.913	0.084	22.90	0.000	1.749	2.077
PC model specification	m1	0.177	0.037	4.73	0.000	0.104	0.251
	m2	0.088	0.034	2.57	0.010	0.021	0.156
	m3	0.078	0.036	2.19	0.028	0.008	0.147
	q1	–0.224	0.030	–7.37	0.000	–0.284	–0.165
	q2	–0.204	0.031	–6.50	0.000	–0.266	–0.143
	q3	–0.212	0.033	–6.46	0.000	–0.276	–0.147
DFM model specification	id11	0.006	0.042	0.15	0.878	–0.076	0.088
	id12	– omitted –					
	id22	–0.088	0.032	–2.76	0.006	–0.151	–0.026
	id32	–0.128	0.032	–3.97	0.000	–0.191	–0.065
	id42	–0.103	0.033	–3.11	0.002	–0.168	–0.038
	sm22	–0.043	0.031	–1.39	0.164	–0.104	0.018
	sm32	–0.089	0.031	–2.86	0.004	–0.150	–0.028
	sm42	–0.100	0.032	–3.16	0.002	–0.161	–0.038
	no. of obs.	3,496					
	R^2	0.974					

Source: Authors' calculations.

Note: OLS regression on dummy variables for different models (AR(1), dynamic factor, principal components), model specifications, countries, horizons and indicators sets, dependent variable = RMSE, robust standard errors.

restrictions, including the coefficients of the dummy variables for the five indicator sets, to test for statistically significant differences between the RMSEs based on the large, medium or small set of predictors. Since we also include the AR(1) results in the regression, the significance of the DIND dummies in the LSDV regression directly indicates whether any of the models including monthly indicators outperforms the AR(1) benchmark.

Table 2 reports the results of the LSDV regression including all specifications using both broad model classes and the benchmark. The models for Slovenia are omitted in the regression below, as is the DFM specification based on the assumption of an AR(1) idiosyncratic component, extracting one factor and using 2 lags ("id12"). Since we estimate without a constant, we can read the average RMSE for each horizon from the coefficients of the DHOR dummies. The negative and significant coefficients on the dummies for the five indicator sets clearly demonstrate the superiority of model-based predictions using monthly indicators over the AR(1) model.

In the next step, we test the restriction that $\alpha_j = \alpha_k$ for any $j,k \in I$, $j \neq k$ against the alternative that the difference between the two coefficients is greater than zero. The Wald tests in table 3 in combination with the regression results above show that the small indicator set including foreign variables yields the best forecasting accuracy.

Table 3

Wald test on the equality of coefficients for indicator sets

	Large	Medium	Medium domestic	Small
Medium	42.700 *0.000*			
Medium domestic	5.620 *0.018*	18.750 *0.000*		
Small	161.910 *0.000*	41.000 *0.000*	115.510 *0.000*	
Small domestic	25.290 *0.000*	0.850 *0.358*	8.910 *0.003*	45.080 *0.000*

Source: Authors' calculations.

Note: F-values of a two-sided test on the equality of coefficients for different indicator sets are reported, p-values in italics, based on robust standard errors.

Even though we consider the dummies for country, horizon and model specification mainly as control variables, it is interesting to take a quick look at the coefficients of these dummies as well. Supported also by the results of the bilateral Wald tests of all combinations of coefficients (not reported here), we can clearly reject the hypothesis that forecasting performance for different horizons is equal and, in line with our impression from chart 1, we conclude that forecasting performance is significantly better for backcasts, followed by nowcasts. This is also reflected in the low coefficients of the dummies for backcasts, followed by those for nowcasts in table 2. Forecasts show the largest RMSE on average. In addition, while we can clearly distinguish between the respective quarters for which a prediction is made, we do not always find a significant difference between the months in which the prediction is made. The coefficients of the three forecasts made in different months of a quarter are not statistically different from each other. Likewise, the nowcast in the first month cannot be distinguished from the nowcast in the second month of a quarter. Yet the nowcast in the third month is significantly better than in the previous two months. Finally, the distinction between models is less clear-cut, similar to the results in Rünstler et al. (2009). We cannot identify a superior forecasting performance of either the principal component model or the DFM, as the results depend strongly on the specification used. Yet we see that the principal component model based on quarterly aggregation consistently yields the lowest RMSEs controlling for all other factors of variation.

4 Conclusions

We tested the performance of different computational estimates of GDP growth for selected CESEE EU Member States using two competing analysis methods, namely principal component models, also called approximate factor models or static factor models, and dynamic factor models. We use a wide range of 75 monthly indicators to provide an automated real-time solution to predicting past-, current- and next-quarter GDP growth. We put special emphasis on the effect of varying the size of the indicator set for forecasting performance. More specifically, we distinguish between large indicator sets (comprising all 75 indicators including also some foreign variables), medium-sized sets (including only the main component of each indicator) and small sets of 14 indicators that we identify based on careful selection of indicators according to their historic correlation with GDP. For the latter two set types, we also explore whether including indicators that trace foreign economic developments (i.e. global prices and economic activity in the euro area) improves forecast accuracy.

Our results show that forecasting performance – measured by the root mean square error relative to the prediction obtained by the AR(1) model we use as our naïve benchmark – varies significantly between countries, forecast horizons and model specifications. As a first and important result, we are able to obtain more precise forecasts based on computationally intensive models using monthly indicators than a simple extrapolation of GDP using an AR(1) model yields. This holds true for all countries with the exception of Slovakia. Not surprisingly, holding all other factors constant, backcasts are on average more precise than nowcasts, while forecasts are least precise. Interestingly, the precision of one-quarter-ahead forecasts does not improve significantly when new information becomes available during the three months of a quarter, whereas we see a significant gain from the second to the third month for nowcasts and from the first to the second month for backcasts.

More importantly, we can identify a clear gain in forecasting accuracy from selecting indicators based on their lagged and contemporaneous correlation with GDP. Our results suggest that for the CESEE economies in our sample, the inclusion of variables capturing economic developments abroad greatly improves forecasting performance. This is likely to be grounded in the fact that these economies are small and open and hence strongly dependent on external demand and global price developments. Furthermore, we obtain better results when we reduce the set of indicators, even though factor models are generally known for extracting reliable information from large sets of variables. In line with the literature on variable selection in factor models (such as Boivin and Ng, 2006; Bai and Ng, 2008) we attribute this finding to the fact that our large indicator set contains a range of variables from the same category (production in different sectors, different variants of consumer and business sentiment, etc.). Hence, we conclude that the basic conditions that need to be in place for factor models to extract orthogonal factors from the dataset are not met when using the large indicator set for our sample. This may be grounded in a violation of the weak orthogonality assumption as well as in the relative shortness of the time series for these countries. Reducing the set of indicators to fewer indicators clearly improves forecast accuracy.

We thus suggest basing nowcasting models for GDP growth in CESEE economies on carefully selected indicators, including information on foreign economic developments, rather than simply using all available indicators.

References

Arnoštová, K., D. Havrlant, L. Růžička and P. Tóth. 2011. Short-Term Forecasting of Czech Quarterly GDP Using Monthly Indicators. In: Czech Journal of Economics and Finance (Finance a úvěr) 61(6). 566–583.

Bai, J. and S. Ng. 2008. Forecasting economic time series using targeted predictors. In: Journal of Econometrics 146(2). 304–317.

Bańbura, M., D. Giannone and L. Reichlin. 2010. Nowcasting. ECB Working Paper 1275. Frankfurt: European Central Bank.

Bańbura, M. and M. Modugno. 2014. Maximum likelihood estimation of factor models on datasets with arbitrary pattern of missing data. In: Journal of Applied Econometrics 29(1). 133–160.

Barhoumi, K., O. Darné and L. Ferrara. 2013. Dynamic factor models: A review of the literature. In: OECD Journal: Journal of Business Cycle Measurement and Analysis 2013(2). 73–107.

Boivin, J. and S. Ng. 2006. Are more data always better for factor analysis? In: Journal of Econometrics 132(1). 169–194.

Chamberlain, G. and M. Rothschild. 1983. Arbitrage, Factor Structure, and Mean-Variance Analysis on Large Asset Markets. In: Econometrica 51(5). 1281–1304.

Engle, R. and M. Watson. 1981. A One-Factor Multivariate Time Series Model of Metropolitan Wage Rates. In: Journal of the American Statistical Association 76(376). 774–781.

Evans, M. D. D. 2005. Where Are We Now? Real-Time Estimates of the Macroeconomy. In: International Journal of Central Banking 1(2). 127–175.

Feldkircher, M., F. Huber, J. Schreiner, M. Tirpák, P. Tóth and J. Wörz. 2015. Bridging the information gap: small-scale nowcasting models of GDP growth for selected CESEE countries. In: Focus on European Economic Integration Q2/15. 56–75.

Forni, M., M. Hallin, M. Lippi and L. Reichlin. 2000. The Generalized Dynamic-Factor Model: Identification and Estimation. In: The Review of Economics and Statistics 82(4). 540–554.

Forni, M. and L. Reichlin. 1998. Let's get real: a factor analytical approach to disaggregated business cycle dynamics. In: The Review of Economic Studies 65(3). 453–473.

Giannone, D., L. Reichlin and D. Small. 2008. Nowcasting: The real-time informational content of macroeconomic data. In: Journal of Monetary Economics 55(4). 665–676.

Mariano, R. and Y. Murasawa. 2003. A New Coincident Index of Business Cycles Based on Monthly and Quarterly Data. In: Journal of Applied Econometrics 18. 427–443.

Nunes, L. C. 2005. Nowcasting quarterly GDP growth in a monthly coincident indicator model. In: Journal of Forecasting 24(8). 575–592.

Rünstler, G. 2016. On the Design of Data Sets for Forecasting with Dynamic Factor Models. In: E. Hillebrand and S. J. Koopman (eds.). Dynamic Factor Models (Advances in Econometrics 35). Emerald Group Publishing Limited. 629–662.

Rünstler, G., K. Barhoumi, S. Benk, R. Cristadoro, A. Den Reijer, A. Jakaitiene, P. Jelonek, A. Rua, K. Ruth and C. Van Nieuwenhuyze. 2009. Short-term forecasting of GDP using large datasets: a pseudo real-time forecast evaluation exercise. In: Journal of Forecasting 28(7). 595–611.

Rusnák, M. 2016. Nowcasting Czech GDP in real time. Economic Modelling 54(C). 26–39.

Schumacher, C. and J. Breitung. 2008. Real-time forecasting of German GDP based on a large factor model with monthly and quarterly data. In: International Journal of Forecasting 24(3). 386–398.

Stock, J. H. and M. W. Watson. 2002a. Forecasting Using Principal Components From a Large Number of Predictors. In: Journal of the American Statistical Association 97(460). 1167–1179.

Stock, J. H. and M. W. Watson. 2002b. Macroeconomic Forecasting Using Diffusion Indexes. In: Journal of Business & Economic Statistics 20(2). 147–162.

Annex

Table A1

List of monthly indicators and indicator sets

Indicator	Seasonal adjustment	Source	Publication lag (weeks)	Indicator set in which indicator is included		
				small	medium	large
Production in industry						
Industry total	SCA	Eurostat	6	×	×	×
Mining and quarrying	SCA	Eurostat	6			×
Manufacturing	SCA	Eurostat	6			×
Electricity, gas, steam and air conditioning supply	SCA	Eurostat	6			×
Turnover in industry						
Mining and quarrying	SCA	Eurostat	6			×
Manufacturing	SCA	Eurostat	6	×	×	×
Turnover in industry, domestic market						
Mining and quarrying	SCA	Eurostat	6			×
Manufacturing	SCA	Eurostat	6			×
Turnover in industry, nondomestic market						
Mining and quarrying	SCA	Eurostat	6			×
Manufacturing	SCA	Eurostat	6			×
Production in construction						
Production in construction	SCA	Eurostat	7	×	×	×
Turnover in retail trade						
Retail trade, except of motor vehicles and motorcycles	SCA	Eurostat	5	×	×	×
Nights spent at tourist accommodation establishments						
Nights spent at tourist accommodation establishments	SCA	Eurostat	6			×
Business and consumer surveys						
Consumers						
Financial situation over the past 12 months	SA	Eurostat	0			×
Financial situation over the next 12 months	SA	Eurostat	0		×	×
General economic situation over the past 12 months	SA	Eurostat	0			×
General economic situation over the next 12 months	SA	Eurostat	0		×	×
Price trends over the past 12 months	SA	Eurostat	0			×
Price trends over the next 12 months	SA	Eurostat	0		×	×
Unemployment expectations over the next 12 months	SA	Eurostat	0		×	×
The current economic situation is adequate to make major purchases	SA	Eurostat	0			×
Major purchases over the next 12 months	SA	Eurostat	0		×	×
The current economic situation is adequate for savings	SA	Eurostat	0			×
Savings over the next 12 months	SA	Eurostat	0			×
Statement on the financial situation of the household	SA	Eurostat	0			×
Consumer confidence indicator	SA	Eurostat	0		×	×
Industry						
Production development observed over the past three months	SA	Eurostat	0		×	×
Employment expectations over the next three months	SA	Eurostat	0			×
Assessment of order book levels	SA	Eurostat	0		×	×
Assessment of export order book levels	SA	Eurostat	0			×
Assessment of the current level of stocks of finished products	SA	Eurostat	0			×
Production expectations over the next three months	SA	Eurostat	0		×	×
Selling price expectations over the next three months	SA	Eurostat	0			×
Industrial confidence indicator	SA	Eurostat	0		×	×
Construction						
Building activity development over the past three months	SA	Eurostat	0		×	×
Evolution of the current overall order books	SA	Eurostat	0			×
Employment expectations over the next three months	SA	Eurostat	0			×
Price expectations over the next three months	SA	Eurostat	0			×
Construction confidence indicator	SA	Eurostat	0		×	×
Factors limiting building activity – none	SA	Eurostat	0			×
Factors limiting building activity – insufficient demand	SA	Eurostat	0			×
Factors limiting building activity – weather conditions	SA	Eurostat	0			×
Factors limiting building activity – shortage of labor	SA	Eurostat	0			×
Factors limiting building activity – shortage of material and/or equipment	SA	Eurostat	0			×
Factors limiting building activity – other	SA	Eurostat	0			×
Factors limiting building activity – financial constraints	SA	Eurostat	0			×
Retail sale						
Business activity (sales) development over the past three months	SA	Eurostat	0			×
Volume of stocks currently held	SA	Eurostat	0			×
Expectations of the number of orders placed with suppliers over the next three months	SA	Eurostat	0			×
Business activity expectations over the next three months	SA	Eurostat	0			×
Employment expectations over the next three months	SA	Eurostat	0			×
Retail confidence indicator	SA	Eurostat	0		×	×

Table A1 (continued)

Monthly indicators (continued)

Indicator	Seasonal adjustment	Source	Publication lag (weeks)	Indicator set in which indicator is included		
				small	medium	large
Economic Sentiment Indicator						
Economic Sentiment Indicator	SA	Eurostat	0	×	×	×
Services						
Business situation development over the past three months	SA	Eurostat	0		×	×
Evolution of demand over the past three months	SA	Eurostat	0			×
Expectation of demand over the next three months	SA	Eurostat	0			×
Evolution of employment over the past three months	SA	Eurostat	0			×
Expectation of employment over the next three months	SA	Eurostat	0			×
Services Confidence Indicator	SA	Eurostat	0		×	×
Energy supply						
Natural gas	NA	Eurostat	7		×	×
Electricity	NA	Eurostat	7		×	×
Motor spirit	NA	Eurostat	7			×
Diesel oil	NA	Eurostat	7			×
Passenger car registrations						
Passenger car registrations	SCA	ECB	2	×	×	×
Prices						
HICP	NA	Eurostat	2			×
Producer prices in industry	NA	Eurostat	5			×
Labor market						
Unemployment rate	SA	Eurostat	5	×	×	×
International trade						
Imports	NA	Eurostat	6	×	×	×
Exports	NA	Eurostat	6	×	×	×
World commodity prices						
ECB Commodity Price Index	NA	Eurostat	1			×
HWWI index of world market prices	NA	HWWI	1	×	×	×
HWWI index of world market prices, crude oil	NA	HWWI	1	×	×	×
Foreign economic activity						
Production in industry, euro area	SCA	Eurostat	6	×	×	×
Markit Eurozone Manufacturing Purchasing Managers Index (PMI®)	SA	Markit	0	×	×	×
Ifo Export Expectations, Industry	SA	CESifo	0	×	×	×

Source: Authors' compilations.

Note: Seasonal as well as seasonal and calendar-day adjustment of indicators is undertaken by national statistical institutes. SCA stands for seasonal and calendar-day adjusted, SA for seasonally adjusted, NA for not adjusted times series.