EASTERN EUROPEAN ECONOMICS
2020, VOL. 58, NO. 3, 197–220
https://doi.org/10.1080/00128775.2020.1726185

Nowcasting Real GDP Growth: Comparison between Old and New EU Countries

Evžen Kočenda[a,b,c,d] and Karen Poghosyan[e]

[a]Institute of Economic Studies, Faculty of Social Sciences, Charles University, Prague, Czech Republic; [b]Institute of Theory of Information and Automation, Czech Academy of Sciences, Prague, Czech Republic; [c]CESifo, Munich, Germany; [d]IOS, Regensburg, Germany; [e]Economic Research Department, Central Bank of Armenia, Yerevan, Armenia

ABSTRACT

We analyze the performance of a broad range of nowcasting and short-term forecasting models for a representative set of twelve old and six new member countries of the European Union (EU) that are characterized by substantial differences in aggregate output variability. In our analysis, we generate ex-post out-of-sample nowcasts and forecasts based on hard and soft indicators that come from a comparable set of identical data. We show that nowcasting works well for the new EU countries because, although that variability in their GDP growth data is larger than that of the old EU economies, the economic significance of nowcasting is on average somewhat larger.

KEYWORDS

Bayesian VAR; dynamic and static principal components; european OECD countries; factor augmented VAR; nowcasting; real GDP growth; short-term forecasting

JEL CLASSIFICATION

C33; C38; C52; C53; E37; E52

Introduction

Effective economic policy in any country is conditioned by the availability of timely and accurate economic data and forecasts of them (Banbura et al. 2013; Giannone, Reichlin, and Small 2008; Jansen, Jin, and de Winter 2016). A typical case is represented by central banks, in which policy makers, as a rule, have to make decisions in real time with incomplete information on current economic conditions. The issue is even more important in emerging economies, where variation in economic activity is often high, and data availability might be less than perfect (Bragoli and Fosten 2018; Bragoli, Metelli, and Modugno 2015; Giannone, Agrippino, and Modugno 2013; Luciani et al. 2018). Typically, the data are not available in the required time or are incomplete, and the resulting accuracy of forecasts might be plagued by volatility in the input data. We explore the issue of forecast accuracy with a set of old and new member countries of the European Union (EU) for which a comparable set of identical data is available. Specifically, we compare the forecast accuracy of nowcasting[1] and forecasting algorithms based on the use of data on the real economy from eighteen European countries characterized by different output volatility regimes.[2] Our goal is to show which algorithm delivers the most accurate short-term forecasts of the growth in the real gross domestic product (GDP) and how the results differ between old and new EU members.

Our analysis specifically assesses GDP growth forecasts because GDP is one of the most comprehensive macroeconomic indicators of economic activity. Thus, GDP growth is the target for providing important information in the policy-making process. However, for

CONTACT Evžen Kočenda evzen.kocenda@fsv.cuni.cz Institute of Economic Studies, Faculty of Social Sciences, Charles University, Opletalova 26, Prague 110 00, Czech Republic

most EU members, GDP data are available roughly 45 days (1.5 months) after the end of a reference quarter. However, many different higher-frequency economic indicators are available between the beginning of the quarter and the publication of official figures on real GDP. This information includes data on industrial production, prices and exchange rates, external sector indices, financial variables, money aggregates, business surveys, and confidence indicators. This type of data is not often structured and covered in the same way for both emerging and developed economies. However, the set of the EU members represents such an opportunity and various types of high-frequency data could be quite useful for predicting and understanding the dynamics of real GDP in two groups of countries in a single economic region. For this type of data, nowcasting is a suitable tool whose basic principle is to use early published information in order to obtain an early estimate of real GDP growth before the official figure becomes available (Giannone, Reichlin, and Small 2008).

The forecasting literature has recently developed different algorithms for extracting useful information from large datasets to improve the assessment of real GDP growth in a current quarter (Camacho, Perez-Quiros, and Poncela 2013). They include dynamic factor models that provide a framework for the integration of a large number of economic series with mixed frequencies, missing data, and publication lags to exploit all useful information to forecast real GDP growth in a current quarter. Our analysis uses a dynamic factor model proposed by Giannone, Reichlin, and Small (2008), who generated a real GDP growth nowcast for the US economy using around 200 macroeconomic indicators. We use twenty-five macroeconomic indicators (ten hard/coincident and fifteen soft/leading indicators) because Barhoumi, Darne, and Ferrara (2010) and Alvares, Maximo, and Perez-Quiros (2016) have shown that, a small dataset (with ten to thirty variables) suffice to estimate common factors. Our goal is to use a relatively small number of indicators to compare this nowcasting algorithm to alternative short-term forecasting algorithms to see which algorithm delivers the most accurate short-term forecasts of real GDP growth.[3]

In this paper, we compare a broad range of linear statistical models – nine models in all – that have recently been applied for short-term forecasting. In most recent empirical work, nowcasting is considered only for a single country and with a limited number of models (Aastveit and Trovik 2012; D'Agostino, McQuinn, and O'Brien 2012; Kuzin, Marcellino, and Schumacher 2013; Marcellino and Schumacher 2010; Yiu and Chow, 2010). Therefore, we adopt a comprehensive approach and consider nowcasting for a large number of countries: eighteen European countries (twelve old and six new EU economies) that are also members of the Organization for Economic Cooperation and Development (OECD) whose data we study. For all models with additional factors, we use static and dynamic approaches to extract unobserved components (factors). We also conduct an out-of-sample forecast evaluation for different lag lengths and different combinations of static and dynamic factors. Finally, we analyze the performance of nowcasting and short-term forecasting models, when the variability of real GDP growth changes over time and across countries. For that, our sample includes episodes of volatility in the 2008 financial crisis and its aftermath.

We use the same unified dataset for all countries, comprising twenty-five variables (ten hard and fifteen soft indicators). The hard data include an industrial production index, production of total construction, total retail trade, passenger car registration, housing permits issued, import and export growth rates (Table 1). However, hard data are published with a certain delay. For the majority of countries in our analysis, the hard

Table 1. Dataset description.

Hard Indicators
Total industrial production s.a., Index, 2015 = 100
Total manufacturing production s.a., Index, 2015 = 100
Production of electricity, gas, steam, and air conditioning supply s.a., index, 2015 = 100
Total construction production s.a., Index, 2015 = 100
Total retail trade (Volume), s.a., Index, 2010 = 100
Passenger car registrations s.a., Index, 2010 = 100
Housing permits issued s.a., Index, 2010 = 100
Housing construction starts s.a., Index, 2010 = 100
Imports of goods, s.a., growth previous period
Exports of goods, s.a., growth previous period
Soft Indicators
Manufacturing, production trend, balance s.a., percentage
Manufacturing, production future trend, balance s.a., percentage
Manufacturing production employment, future trend, balance s.a., percentage
Manufacturing production confidence indicators, balance s.a., percentage
Construction, business activity, tendency, balance s.a., percentage
Construction confidence indicators, balance s.a., percentage
Construction employment, future trend, balance s.a., percentage
Real trade, business situation activity trend, balance s.a., percentage
Real trade, business situation activity future trend, balance s.a., percentage
Real trade, confidence indicators, balance s.a., percentage
Retail trade employment, future trend, balance s.a., percentage
Services (excluding retail trade), business situation, activity, trend, balance s.a., percentage
Services (excluding retail trade), confidence indicators, balance s.a., percentage
Services (excluding retail trade), employment, trend, balance s.a., percentage
Services (excluding retail trade), employment, future trend, balance s.a., percentage

Source: OECD (http://stats.oecd.org).
Data are seasonally adjusted. The soft indicators are collected by OECD member countries separately with the help of surveys.

data are published at least forty to forty-five days after the end of the reference month. The soft data mainly include business tendency surveys and consumer confidence indicators for different economic sectors (industry, construction, trade, and services). The soft data are presented in Table 1. In contrast, soft data for a given month become available much earlier than many hard indicators. These data are direct related to current conditions in real economy.

In this paper, we contribute to the existing literature in two ways. First, we provide a comprehensive comparison of the nowcasting and short-term forecasting methods on a solid set of twelve old and six new EU economies, along with a broad range of statistical models over the pre-crisis and post-crisis period. Second, we show that nowcasting based on a small number of indicators is an efficient tool for a current quarter, forecasting when variation in real GDP growth increases over time and across countries. Specifically, we show that nowcasting works well in the new EU countries, and although the variation in their GDP growth data is larger than it is in the old EU economies, the economic effect of nowcasting results is on average comparable between the two groups. The outcome of our analysis could be useful for practitioners at central banks, especially those in new EU countries, as we show that nowcasting performs well when uncertainty in real GDP growth is relatively high.

The remainder of the paper is organized as follows. In section 2, we briefly review the literature related to the topic researched. In section 3, we present the methodological details on nowcasting and short-term forecasting models. In section 4, we present the dynamics of real GDP growth and some important descriptive statistics for the countries selected. In this section, we also give a short description of additional explanatory variables that serve as initial variables for extracting the dynamics of unobservable factors. In section 5, we present a recursive regression scheme for our experimental design. In section 6, we present the out-of-sample evaluation results. Section 7 concludes.

Literature Review

In the following literature review, we cover relevant information on methodologies as well as on selected empirical contributions. The literature offers several approaches in terms of nowcasting models: bridge equations, mixed-data sampling (MIDAS) regressions, mixed-frequency VARs, and mixed-frequency dynamic factor models.[4]

Forecasting with a bridge equation is performed in two steps: in the first step, we forecast each high-frequency indicator (e.g., using ARIMA) to deal with ragged ends; in the next step, monthly indicators are averaged (with equal weights) to quarterly frequency and used to forecast GDP growth or its subcomponents via simple bivariate regression. An example of an application related to the European context is in Baffigi et al. (2004), who estimated bridge models for aggregate GDP and components in the euro area. The results show that the performance of a bridge model is always better than that of standard univariate or multivariate benchmark models.

The MIDAS regression represents alternative benchmark model (Ghysels, Santa-Clara, and Valkanov 2007). MIDAS deals with mixed frequencies by employing a polynomial weighting function to link high-frequency and low-frequency data. The main difference between MIDAS and bridge model is that the MIDAS regression is a direct forecasting tool, while in bridge regression, we model the dynamics of each indicators separately and then use expanded indicators to nowcast real GDP growth. Clements and Galvao (2008) used MIDAS regressions to forecast US real GDP growth and show that, compared to standard alternative methods, it is an effective way to exploit monthly data.

In contrast to the MIDAS approach and consistent with a conventional VAR model based on single-frequency data, the mixed-frequency (MF-VAR) model specifies the joint dynamics of monthly GDP, which are obtained from quarterly GDP by time disaggregation, and the monthly indicator. An example of an application related to the European context is in Kuzin et al. (2011), who compare MIDAS and MF-VAR for nowcasting and forecasting quarterly GDP growth in the euro area. Their results show that the two approaches are more complementary than substitutive: MIDAS tends to perform better for shorter horizons, whereas MF-VAR is better for longer horizons.

The large-scale dynamic factor models' approach to nowcasting was proposed by Giannone, Reichlin, and Small (2008). Their methodology combines principal components analysis and a Kalman filter. First, they obtain the common factors from a large set of macroeconomic indicators. In the second step, they smooth these common factors using a Kalman filter. Afterward, they use smoothed common factors as explanatory variables in a simple ordinary least squares (OLS) regression to produce nowcasts of GDP growth.

The mixed-frequency dynamic factor model has been adopted by a number of central banks and it has nowcasted real GDP growth in different countries. For example, real GDP growth was nowcasted with a dynamic factor model for France (Barhoumi, Darne, and Ferrara 2010), Germany (Marcellino and Schumacher 2010), Ireland (D'Agostino, McQuinn, and O'Brien 2012; Liebermann 2012), Norway (Aastveit and Trovik 2012), and Japan (Bragoli 2017). More recently, nowcasting studies have been published on emerging market economies, including large countries, such as China (Giannone, Agrippino, and Modugno 2013; Yiu and Chow 2010), Indonesia (Luciani et al. 2018), Argentina (Camacho, Dal Bianco, and Martínez-Martín 2015), Brazil (Bragoli, Metelli, and Modugno 2015), India (Bragoli and Fosten 2018), and Russia (Porshakov, Ponomarenko, and Sinyakov 2016), and small European economies such as the Czech Republic (Rusnák 2016). To date, no multi-country analysis of the European countries has been conducted.

The recent literature uses a dynamic factor model approach, but much of the work is focused on a single country or a few countries, employs only a handful of models, and covers relatively short periods, in which variations in GDP were not a significant issue. However, multiple countries are analyzed by Jansen, Jin, and de Winter (2016), and five euro-area countries are analyzed by Angelini et al. (2011).

Our paper accounts for the limitations mentioned above and extends the literature on the nowcasting and short-term forecasting methodology in the following ways. We assess nowcasting and short-term forecasting models when the variation in each country's GDP increases over time; our sample includes the period of volatility in the 2008 financial crisis. Further, we assess whether nowcasting outperforms all other models with statistical significance under higher-amplitude volatility in GDP growth. For that, we use data on actual GDP growth for a broad range of old and new EU countries with different levels of variation in GDP growth. As a result, we offer a comprehensive and comparative analysis targeting European economies.

Methodology

One of the major advantages of nowcasting is its ability to use high-frequency data to estimate quarterly macroeconomic variables, particularly growth in real GDP in real time. We now present the methodology for extraction of the dynamic factors via an algorithm proposed by Giannone, Reichlin, and Small (2008). This methodology relies on the two-step estimator proposed by Doz, Giannone, and Reichlin (2011). According to Doz, Giannone, and Reichlin (2011), the dynamic factor model in the state-space form can be presented as:

$$x_{t_m} = \Lambda f_{t_m} + \xi_{t_m}, \xi_{t_m} \sim N(0, \Sigma_\xi) \tag{1}$$

$$f_{t_m} = \sum_{i=1}^{p} A_i f_{t_{m-i}} + B\eta_{t_m}, \eta_{t_m} \sim N(0, I_q) \tag{2}$$

Equation 1 depicts the N monthly series x_{t_m} to an $(r \times 1)$ vector of latent factors f_{t_m}, through the matrix of factor loadings Λ, plus an idiosyncratic component ξ_{t_m}, assumed to be a multivariate with noise in diagonal covariance matrix Σ_ξ. Equation 2 describes the law of motion in the latent factors, which are driven by a q-dimensional standardized white noise η_{t_m}, where B is an $(r \times q)$ matrix.

To deal with missing observations at the end of the sample, the authors use a two-step estimator. In the first step, the parameters of the model are estimated consistently through principal components on a balanced panel, created by truncating the dataset by the date of the least timely release. In the second step, Kalman smoothing is applied to update the estimates of the factor and the forecast on the basis of the entire unbalanced dataset.[5]

We use calculated common factors in a bridge Equation (3) as explanatory variables in simple OLS to produce a nowcast for GDP.

$$\hat{y}_{t_q} = \alpha + \beta \hat{f}_{t_q} \tag{3}$$

where $\hat{f}_{t_q}$ is the quarterly aggregated correspondent of f_{t_m}.

As an alternative to nowcasting, we also assess the performance of nine widely used short-term forecasting models. In both nowcasting and forecasting models, the target time is the same: the current quarter. To estimate real GDP growth in real time, additional high-frequency data are used that are available in time to be employed by a nowcasting algorithm. Forecasting uses only past information and does not take into account data that are available in the current quarter (Jansen, Jin, and de Winter 2016). Finally, a nowcasting model exploits data in a mixed-frequency domain (including monthly data), while alternative models use only quarterly variables.

In the current paper, we use both univariate and multivariate models. As a univariate model, we use the AR(p) model. Adding unobservable factors to the AR(p) process, we obtain a so-called factor augmented AR(p) model. In a multivariate setting, we use a traditional unrestricted VAR(p) model as well as untraditional and more advanced models such as Bayesian VAR, factor augmented VAR, and Bayesian factor augmented VAR. We now briefly present the main idea and computational characteristics of those models.

We begin with an unrestricted VAR model (4), which can be presented as follows:

$$y_t = A_0 + A_1 y_{t-1} + A_2 y_{t-2} + \ldots + A_p y_{t-p} + v_t,\ \ t = 1, \ldots, T \tag{4}$$

in which y_t is an $(n \times 1)$ vector of variables to be forecasted, A_0 is a $(n \times 1)$ vector of constant terms, $A_1, A_2, \cdots, A_p$ is an $(n \times n)$ matrix of estimated parameters for different lag lengths $(l = 1, 2, \cdots, p)$, and v_t is an $(n \times 1)$ vector of error terms. We assume that $v \sim N\left(0, \sigma^2 I_{(n\times n)}\right)$, where $I_{(n\times n)}$ is an $(n \times n)$ identity matrix.[6] The parameters of the unrestricted VAR models can be consistently estimated by using an OLS algorithm. But, because in the VAR model we often need to estimate many parameters, this over-parametrization could cause inefficient estimates and large out-of-sample forecast errors. The Bayesian estimation approach is a viable alternative to overcome this over-parametrization problem (Banbura, Giannone, and Reichlin 2010).

In this paper, we use a standard Bayesian VAR model with well-known Minnesota-style priors. According to these priors, the restrictions are imposed by specifying normal prior distributions with zero mean and small standard deviations that decrease as the number of lags increases. The exception to this is that the coefficient on the first lag of a variable has a mean of 1.[7] Thus, according to the Minnesota-type priors, the prior mean and standard deviation can be set as follows: the parameters of the first lag of the dependent variables follow an AR(1) process while parameters for other lags equal zero, and the variance in the priors can be specified as follows:

$$\left(\frac{\lambda_1}{l^{\lambda_3}}\right)^2 \text{if } i = j,\ \left(\frac{\sigma_i \lambda_1 \lambda_2}{\sigma_j l^{\lambda_3}}\right)^2 \text{if } i \neq j,\ (\sigma_1 \lambda_4)^2 \text{ for the constant term.}$$

in which i is the dependent variable in the i-th equation and j is the independent variable in that equation, σ_i and σ_j are standard errors from the AR regressions estimated via OLS. The ratio of σ_i and σ_j controls for the possibility that variables i and j may have different scales (i is the lag length). The parameters λ_S are set by a researcher to control for the tightness of the prior. After setting the values of the priors, we can calculate the posterior parameters using a Bayesian approach.

It is well known that a traditional VAR model cannot accommodate a large number of variables, as they can cause serious problems in the forecasting accuracy of the model. Thus, in addition to small-scale unrestricted VAR and Bayesian VAR models, we also use factor augmented VAR (FAVAR) and Bayesian factor augmented VAR (BFAVAR) models. Following Bernanke, Boivin, and Eliasz (2005), we present the FAVAR and BFAVAR models (5) as follows:

$$\begin{vmatrix} Y_t \\ F_t \end{vmatrix} = A_1 \begin{vmatrix} Y_{t-1} \\ F_{t-1} \end{vmatrix} + A_2 \begin{vmatrix} Y_{t-2} \\ F_{t-2} \end{vmatrix} + \cdots + A_p \begin{vmatrix} Y_{t-p} \\ F_{t-p} \end{vmatrix} + \begin{vmatrix} v_t \\ u_t \end{vmatrix} \tag{5}$$

in which Y_t is the vector of observable variables, F_t is the vector of unobservable variables estimated via a static principal component or two-step Kalman filter algorithm, $A_1, A_2 \cdots A_p$ are $(r \times r)$ matrices of estimated parameters, and v_t and u_t are the error terms with zero mean and diagonal variance-covariance matrices, Q and V. In the model presented, the parameters can be estimated by either OLS or the Bayesian estimation approach.

As a rule, FAVAR and BFAVAR models can be estimated in two steps: in the first step, we estimate the dynamics of principal components, and in the second step, we estimate the model parameters and conduct forecasts. The unobservable factors can be estimated via three popular approaches: the Stock-Watson static principal component approach (Stock and Watson 2002), time domain (Doz, Giannone, and Reichlin 2011, 2012), and frequency domain (Forni et al. 2005). In this paper, as mentioned, above we follow the approach suggested by Doz, Giannone, and Reichlin (2011).[8] After estimating the factor dynamics, the FAVAR and BFAVAR models can be estimated in a traditional manner using OLS. In other words, we use a small-scale VAR model with variables of interest augmented by extracted factors.

Data and Selection of Indicators

We employ 25 types of data series to nowcast GDP growth. We sort the variables used into two groups: hard indicators (based on production and sales data) and soft indicators (based on survey data). When forming an appropriate dataset, we closely follow approach of Camacho and Perez-Quiros (2010), who divide the variables into three groups: GDP revisions (flash, first, and second revisions), hard indicators (based on economic activity data), and soft indicators (based on survey data).[9] In our analysis, we follow these classifications with minor exceptions. We use only the second GDP revision (because we do not have information on flash and first GDP revisions), and for hard and soft indicators, we use almost the same variables but with more disaggregation, shown in Table 1. Our selected dataset includes ten hard and fifteen soft variables. Hard indicators have a direct link with real economic activity, but they have one important disadvantage of being published with at least 45 days delay.[10] For this reason, we also include soft indicators, as they are available on a timely basis. Specifically, the soft indicator, which refers to a given month, is available before the end of that month. Further, Banbura and Rünstler (2011) have shown that the inclusion of soft indicators in a nowcasting model can substantially improve nowcasting performance, and Karasoy Can and Yüncüler (2018) show that soft indicators such as consumer confidence help to better predict future private consumption growth. The monthly data are available on a seasonally adjusted basis from the source.

The second set of data we use are the yearly real GDP growth rates for eighteen European countries listed in Table 2.[11] Further, we calculated the coefficient of percentage variability C_v defined as: $C_v = \sigma_y \times 100/\bar{y}$, where $\bar{y}$ is average value of GDP growth. The variation in GDP growth (σ_y) is defined as: $\sigma_y = \sqrt{\sum_{t=1}^{T} \left(y_t - \hat{y}_t\right)^2/(n-p)}$, where $y_t$is GDP growth at time t, $\hat{y}_t$ is a fitted value of GDP growth at time t calculated by a trend-line equation, n is the number of observations in the sample, and p is the number of parameters in the trend-line equation (p = 2). In essence, the variation formula is an unbiased standard deviation adjusted for the number of parameters in a model.

To clarify how variation in real GDP growth changes, we divide the full period into two subperiods: before and after the global financial crisis. Then, we compute the actual values of the coefficient of variation for two subperiods: 2000Q1–2007Q4 and 2008Q1–2018Q4 (see Table 2). Table 2 shows that the coefficient of variation increased during the post-crisis period in all the countries in our dataset. We assume that an increase in the coefficient of variation might be caused by the global financial crisis at the end of 2008. Further, we can compare the minimum and maximum values of the coefficient of variation in various countries and two subperiods separately. For example, the minimum

Table 2. Real GDP growth rate and coefficient of variability (in %).

	2000	2005	2010	2015	2016	2017	2018	Coefficient of variation, 2000–2007%	Coefficient of variation, 2008–2018, %
Austria	3.40	2.20	1.80	1.10	2.00	2.60	2.70	1.25	1.61
Belgium	3.63	2.09	2.74	1.74	1.45	1.73	1.44	1.00	1.14
Czech Republic	4.27	6.53	2.27	5.31	2.45	4.35	2.96	2.33	3.68
Denmark	3.75	2.34	1.87	2.34	2.40	2.26	1.49	1.36	2.27
Estonia	10.57	9.37	2.26	1.90	3.49	4.86	3.87	2.55	5.44
Finland	5.63	2.78	2.99	0.50	2.77	3.04	1.66	1.45	3.00
France	3.92	1.66	1.95	1.11	1.10	2.26	1.72	0.68	1.19
Germany	2.90	0.73	4.18	1.74	2.23	2.47	1.52	1.66	1.86
Hungary	4.21	4.39	0.66	3.54	2.28	4.14	4.94	0.98	4.26
Italy	3.71	0.95	1.69	0.92	1.12	1.68	0.86	0.79	2.32
Netherlands	4.20	2.05	1.34	1.96	2.19	2.91	2.60	1.53	2.38
Poland	4.56	3.49	3.61	3.84	3.06	4.94	5.15	1.98	2.00
Portugal	3.79	0.77	1.90	1.82	1.93	2.80	2.14	0.97	3.21
Slovakia	1.21	6.75	5.04	4.17	3.13	3.19	4.11	2.95	3.18
Slovenia	4.16	4.00	1.24	2.30	3.07	4.88	4.49	1.95	5.00
Spain	5.29	3.72	0.01	3.64	3.17	2.98	2.58	0.57	4.14
Sweden	4.75	2.82	5.99	4.46	2.68	2.10	2.36	1.22	2.31
United Kingdom	3.45	3.15	1.71	2.35	1.79	1.82	1.40	0.42	2.02

Source: OECD (http://stats.oecd.org).

The last two columns list the values of the coefficient of variation. The coefficients of variation are calculated for two subperiods, particularly before and after the global crisis. The coefficient of variation is always positive; if the value is larger, then, we conclude that during the post-crisis subperiod the variation in real GDP growth was higher than before the crisis. It is evidenced that in all countries the variation in real GDP growth is higher in the post-crisis period than in the pre-crisis period.

value of the coefficient of variation in the pre-crisis period is observed in the United Kingdom (0.42%), while the maximum value is observed in Slovakia (2.95%). Hence, the amplitude of variation during the pre-crisis period is 2.53 percentage points (2.95– 0.42). During the post-crisis period, the minimum value of the coefficient of variation is observed in Belgium (1.14%), while the maximum value is observed in Estonia (5.44%). Hence the amplitude of variation during the post-crisis period is 4.30 percentage points (5.44– 1.14). Thus, during the post-crisis period, the amplitude of the coefficient of variation among countries nearly doubled. Our observations about differences in output volatility are consistent with those by Benczúr and Rátfai (2014) for the period prior to the global financial crisis in 2008.

Experimental Design

We employ a recursive regression scheme to analyze the relative performance of nowcasting versus short-term forecasting models (AR, VAR, BVAR, FAAR_SW, FAAR_TS, FAVAR_SW, FAVAR_TS, BFAVAR_SW, and BFAVAR_TS) when the coefficient of variation in real GDP is changing (increasing and decreasing) both over time and across countries. Based on the out-of-sample RMSE criterion, we assess the performance of the nowcasting versus alternative short-term forecasting models.

Under the optimal scenario, we work with real-time data, as suggested by Croushore and Stark (2001). However, we do not have a real-time dataset. Therefore, we perform a simulated real-time analysis using revised data instead. For all countries in our dataset, we simulate a nowcasting exercise at the end of the third month of the current quarter. At this time, for most countries we have hard data for the first month of the current quarter, and for all countries we have soft data for all three months. Then we conduct out-of-sample nowcasting experiments with this available set of hard and soft data and the methodology mentioned earlier.[12] To conduct these experiments, we divide the full data sample into in-sample and out-of-sample groups. Considering the length of the real GDP quarterly data for each country, we use 70% of the observations as our in-sample group and the remaining 30% as our out-of-sample group (see Table A1).[13] Given the data length, our forecast horizon for each country is different, and the horizon consists of more than one period.

For a out-of-sample forecast comparison, we use a recursive regression scheme. A forecasting model with a recursive window assumes that the initial estimation period is fixed, and additional observations are added to the estimation period one at a time. In the nowcasting model, we perform a recursive simulation experiment. The main difference in the experiment design is that in the nowcasting model, we take into account all information available in the current quarter, whereas in the short-term forecasting model, we ignore the information available in the current quarter, as it would not be available in the real world. The main task is to describe whether the information available in the current quarter significantly helps to improve the accuracy of the forecast for the target variable. A detailed description of the out-of-sample design for both nowcasting and forecasting is in Appendix 1.

After obtaining all the forecast points for all available models, we compare nowcasting and different short-term forecasting models to determine which is better. To do that, we use out-of-sample nowcasts and short-term forecasts from the recursive regression scheme to verify the forecast accuracy produced by different models for different countries. Like Banbura et al. (2013), Barhoumi, Darne, and Ferrara (2010), Jansen, Jin, and de Winter

(2016), and Pirschel and Wolters (2014), we assess the forecast accuracy with the standard root mean squared forecast error (RMSE) measure (6), defined as:

$$RMSE^{i}_{gdp} = \sqrt{\frac{1}{T_i^* - 1}\sum\nolimits_{t=1}^{T_i^*-1}\left(\hat{y}_{i,t} - y_{i,t}\right)^2} \tag{6}$$

where $RMSE^{i}_{gdp}$is the calculated root mean squared forecast error for the i-th country, $y_{i,t}$ is the actual value of the GDP growth rate for the i-th country, $\hat{y}_{i,t}$ is the estimated value of the GDP growth rate for the i-th country, and T_i^* denotes the out-of-sample period for the i-th country.

Empirical Results

In this section, we present estimation results for ten models: namely, one nowcasting model and nine alternative models for short-term forecasting. We test a nowcasting model so as to use available information in the current quarter. In contrast, the short-term forecasting models generate forecasts based only on the past information set. At the same time, we also want to check the behavior of the nowcasting model versus short-term forecasting models when the coefficient of variation in real GDP is changing (increasing or decreasing) both in time and across countries – this is an important issue with respect to old and new EU members in our sample.

For our analysis to be robust, we use different lag lengths to estimate the parameters in short-term forecasting models, particularly from one lag to four lags, as in Pirschel and Wolters (2014). In addition to selecting the lag length, we also determine the optimal combination of static and dynamic factors, in a way that is similar to Poghosyan (2016); the technical presentation of factor selection is in Appendix 2. Finally, we compare a variety of different specifications for each model and choose the one that yields the best ex-post forecasting performance.

Our key empirical out-of-sample evaluation results (in absolute terms) are in Table 3 for new EU markets and in Table 4 for the old EU economies. Table 3 shows that in five out of six new EU markets, the nowcasting model outperforms all the short-term forecasting models considered; the exception is Slovakia. Recall that in Table 2 the new EU economies are characterized by relatively high variation in output, especially after the global financial crisis. Hence, even when variation is relatively high, the nowcasting model helps to reduce the errors and produces more accurate one-step-ahead forecasts than the short-term forecasting models.

In Slovakia, the majority of short-term forecasting models considered are still beaten by the nowcasting model. At the same time, Bayesian VAR and FAVAR models perform better than the nowcasting model. That could be due to the increased volatility in 2008–2009 but relative stability of GDP growth rates in the rest of the evaluation period, because the two Bayesian models take the persistence of GDP growth rates into account by imposing a prior in the estimation procedure (for more details, see Adam and Novotny 2018).

Further, in Table 4, the out-of-sample evaluation results for old EU countries show that, for eight of the twelve countries, the nowcasting model outperforms all the short-term forecasting models considered. The exceptions are Belgium, the Netherlands, Portugal, and Spain, but even for these countries, many of the short-term forecasting models considered are still beaten by the

Table 3. Out-of-sample RMSE indices for real GDP growth (New EU markets).

	Nowcasting	AR	VAR	BVAR	FAAR_SW	FAAR_TS	FAVAR_SW	FAVAR_TS	BFAVAR_SW	BFAVAR_TS
Czech Republic	**0.575**	0.591	0.627	0.626	0.583	0.588	0.585	0.597	0.625	0.625
Estonia	**0.562**	0.647	1.032	0.885	0.718	0.724	1.040	0.984	0.994	0.967
Hungary	**0.371**	0.581	0.626	0.645	0.557	0.555	0.604	0.602	0.656	0.653
Poland	**0.456**	0.578	0.653	0.639	0.578	0.576	0.632	0.632	0.637	0.635
Slovak Republic	0.294	0.343	0.514	0.223	0.382	0.461	0.474	0.546	0.205	**0.204**
Slovenia	**0.347**	0.530	0.621	0.5960	0.566	0.567	0.607	0.635	0.593	0.588

The RMSE results are presented for six new EU economies. The best-performing models are in boldface.

Table 4. Out-of-sample RMSE indices for real GDP growth (Old EU economies).

	Nowcasting	AR	VAR	BVAR	FAAR_SW	FAAR_TS	FAVAR_SW	FAVAR_TS	BFAVAR_SW	BFAVAR_TS
Austria	**0.289**	0.328	0.336	0.384	0.311	0.312	0.315	0.317	0.384	0.384
Belgium	0.275	0.282	0.278	0.310	0.278	0.277	**0.255**	0.267	0.321	0.317
Denmark	**0.724**	0.754	0.814	0.895	0.776	0.756	0.816	0.821	0.894	0.895
Finland	**0.532**	0.556	0.673	0.634	0.645	0.650	0.692	0.692	0.637	0.624
France	**0.250**	0.333	0.351	0.393	0.294	0.298	0.318	0.323	0.393	0.393
Germany	**0.302**	0.419	0.534	0.499	0.369	0.381	0.475	0.478	0.477	0.475
Italy	**0.268**	0.278	0.304	0.299	0.273	0.277	0.330	0.341	0.299	0.299
Netherlands	0.386	0.396	0.402	0.481	**0.366**	0.367	0.374	0.387	0.480	0.484
Portugal	0.614	**0.514**	0.589	0.585	0.555	0.554	0.671	0.677	0.582	0.578
Spain	0.345	0.204	0.237	0.250	**0.204**	0.205	0.224	0.226	0.242	0.241
Sweden	**0.434**	0.532	0.606	0.670	0.556	0.553	0.581	0.591	0.684	0.685
United Kingdom	**0.246**	0.352	0.366	0.406	0.343	0.344	0.360	0.362	0.406	0.406

The RMSE results are presented for the twelve old EU economies. The best-performing models are in boldface. From table we see that the nowcasting model outperforms the short-term forecasting models for 8 countries out of twelve. Only for four countries (Belgium, Netherlands, Portugal and Spain) the short-term forecasting models outperform nowcasting results.

nowcasting model, as with Slovakia. The coefficient of variation, in particular, after the post-crisis period is lower in the old EU economies than in the new EU economies (see Table 2). Hence, and less surprisingly, with a low level of variation, nowcasting has the power to reduce errors and produce more accurate one-step-ahead forecasts than the short-term forecasting models.

To show our key out-of-sample evaluation results from a directly comparative perspective, we present normalized RMSE values in Tables 5 and 6. The normalized RMSE values for each country are calculated as the absolute RMSE values divided by the corresponding standard deviation of the GDP growth rate during the out-of-sample evaluation period. In the new EU markets (Table 5), nowcasting again outperforms the short-term forecasting models for five out of six countries. Similarly, for the old EU economies (Table 6), nowcasting has better performance than the short-term forecasting models in eight of the twelve countries.

These results enable us to quantify the economic significance (or effect) of the nowcasting model in terms of the reduction in forecast variation compared to the naïve forecast (average decrease/increase over a testing period) and to provide a comparison among countries. Table 5 shows that nowcasting has the highest reduction in variation GDP growth forecasts for Slovenia (by 32.8%; [0.672*100 – 100]) and Estonia (by 26.6%; [0.734*100–100]); both countries also have the highest variation in GDP growth. During the post-crisis period, the average reduction in the forecast variation is about 25.0% for five new EU economies. Further, Table 6 enables us to deduce that nowcasting brings the highest reduction in variation in GDP growth forecasts for France and Sweden (19% and 18%, respectively). Unlike in the new EU markets, for Finland, which has the most-volatile GDP growth among the old EU economies, the reduction in forecast variation is less than 11%. For eight old EU countries, where nowcasting outperforms the short-term forecasting models, the average reduction in forecast accuracy is about 15%.

The results in Tables 5 and 6 can be summarized as follows. The average lower values in the normalized RMSE indicate a somewhat greater reduction in nowcasting and forecast variation in the new EU markets than in old EU economies. The difference in reduction is about 10%, though this should not be overplayed. Somewhat less reduction in forecast variation in the old EU economies might be intuitively due to their lower variation in GDP growth. Hence, the dynamics of GDP growth can be forecasted more accurately in the old EU economies with the trivial models, and the inclusion of other explanatory variables could actually lead to an increase in out-of-sample forecast error. In addition, Table 2 shows that after the global financial crisis, the ratio between the maximum and minimum values of variation in GDP was almost 4.8 (5.44/1.14), whereas the ratio between average normalized RMSE values calculated for nowcasting was almost 1.14 (0.853/0.750). This means that the nowcasting algorithm is a useful tool for the new EU economies in terms of achieving results that are comparable to those for the old EU economies, despite significant differences in GDP variation between the two groups.

We also compare the normalized RMSE values in Tables 5 and 6 across models, showing that when nowcasting outperforms other competing models, the outcomes differ between country groups only slightly. In the new EU markets, the minimum improvement between nowcasting and the second-best alternative forecast model ranges from less than 2% between nowcasting and FAAR_SW model in the Czech Republic to about 35% in Hungary (various FAVAR models). In the old EU economies, the situation is similar: the

Table 5. Out-of-sample normalized RMSE indices for real GDP growth (New EU markets).

	Nowcasting	AR	VAR	BVAR	FAAR_SW	FAAR_TS	FAVAR_SW	FAVAR_TS	BFAVAR_SW	BFAVAR_TS
Czech Republic	**0.809**	0.832	0.882	0.881	0.821	0.827	0.823	0.840	0.879	0.880
Estonia	**0.734**	0.844	1.347	1.156	0.937	0.945	1.358	1.285	1.297	1.263
Hungary	**0.749**	1.175	1.266	1.303	1.126	1.121	1.221	1.216	1.326	1.321
Poland	**0.788**	0.999	1.128	1.104	0.998	0.995	1.092	1.091	1.101	1.097
Slovak Republic	0.999	1.167	1.750	0.759	1.301	1.568	1.612	1.857	0.696	**0.692**
Slovenia	**0.672**	1.025	1.202	1.153	1.096	1.098	1.175	1.228	1.146	1.138

The normalized RMSE results are presented for the six new EU economies. The normalized RMSE is calculated by dividing RMSE in Table 3 by the standard deviation of GDP growth in the out-of-sample period. The best-performing models are in boldface.

Table 6. Out-of-sample normalized RMSE indices for real GDP growth (Old EU economies).

	Nowcasting	AR	VAR	BVAR	FAAR_SW	FAAR_TS	FAVAR_SW	FAVAR_TS	BFAVAR_SW	BFAVAR_TS
Austria	**0.841**	0.955	0.977	1.118	0.906	0.907	0.918	0.922	1.117	1.117
Belgium	1.082	1.107	1.091	1.217	1.094	1.089	**1.003**	1.050	1.261	1.246
Denmark	**1.011**	1.053	1.138	1.250	1.083	1.056	1.140	1.146	1.249	1.250
Finland	**0.892**	0.931	1.127	1.061	1.081	1.089	1.160	1.160	1.067	1.045
France	**0.810**	1.077	1.135	1.271	0.950	0.962	1.028	1.043	1.271	1.271
Germany	**0.752**	1.043	1.329	1.241	0.919	0.948	1.181	1.191	1.188	1.182
Italy	**0.726**	0.754	0.824	0.811	0.740	0.750	0.896	0.926	0.811	0.811
Netherlands	0.890	0.911	0.926	1.108	**0.844**	0.844	0.861	0.892	1.106	1.115
Portugal	1.030	**0.862**	0.988	0.982	0.931	0.929	1.125	1.136	0.976	0.970
Spain	0.657	0.389	0.451	0.476	**0.388**	0.390	0.427	0.430	0.460	0.459
Sweden	**0.816**	0.999	1.139	1.259	1.044	1.040	1.091	1.111	1.285	1.286
United Kingdom	**0.979**	1.400	1.455	1.617	1.366	1.371	1.435	1.443	1.617	1.617

The RMSE results are presented for the twelve old EU economies. The normalized RMSE is calculated by dividing RMSE in Table 4 by the standard deviation of GDP growth in the out-of-sample period. The best performing models are in boldface.

minimum improvement between nowcasting and the second-best alternative is from a mere 2% (nowcasting vs. FAAR_SW for Italy) to about 28% (nowcasting vs. FAAR_SW for the UK). Hence, the differences in economic significance between the two EU groups in terms of improvement from using nowcasts versus forecasts are negligible.

To verify whether the results obtained for RMSE are significantly different statistically among models, we perform further cross-model tests to differentiate between nowcasting and nine short-term forecasting models. The cross-model test is based on a statistic proposed by Diebold and Mariano (1995). We calculate the Diebold-Mariano (hereafter DM) statistic by regressing the loss differential on an intercept, using heteroscedasticity autocorrelation (HAC) robust standard errors in the following way. ε_t^{nc} denotes the forecast errors in the nowcasting model, andε_t^i denotes the forecast errors in the alternative short-term forecasting models. Then, the loss differential l_t can be calculated as $l_t = \left(\varepsilon_t^{nc}\right)^2 - \left(\varepsilon_t^i\right)^2$. The null hypothesis is that the loss differential equals zero ($H_0 : l_t = 0$). The results of t-statistics obtained from regressing the loss differential on the intercept are presented separately for the new and old EU economies in panels A and B of Table 7.

First, recall that nowcasting outperforms the short-term competing models for thirteen out of eighteen countries (based on the results in Tables 3 and 4). In the remaining five countries (Slovakia, Belgium, the Netherlands, Portugal, and Spain), short-term forecasting models outperform the nowcasting model; these five countries are listed in boldface in Table 7. The information in Table 7 indicates whether the performance results of the nowcasting and forecasting models in Tables 3 and 4 are significantly different statistically. The main importance of Table 7 is that the results related to better performance in the short-term forecasting models (BFAVAR_TS and FAAR_SW) than in nowcasting are

Table 7. Diebold-Mariano test.

	NOWCASTING VS.								
	AR	VAR	BVAR	FAAR_SW	FAAR_TS	FAVAR_SW	FAVAR_TS	BFAVAR_SW	BFAVAR_TS
Panel A: New EU markets									
Czech Republic	−0.21	−0.58	−0.49	−0.13	−0.20	−0.15	−0.32	−0.48	−0.48
Estonia	−0.82	3.23***	−2.53**	−1.35	−1.37	−2.62**	−1.97*	−2.86***	−2.89***
Hungary	−1.79*	−2.52**	−1.72*	−1.56	−1.52	−1.98*	−1.95*	−1.77*	−1.76*
Poland	−1.43	−1.61	−2.11**	−1.51	−1.48	−1.72*	−1.70*	−1.97*	−1.94*
Slovak Republic	−0.67	−2.08*	1.28	−2.87***	−2.73***	−2.79***	−3.37***	1.58	1.69*
Slovenia	−1.62	−1.90*	−2.01*	−3.20***	−3.45***	−1.85*	−2.83***	−2.04*	−2.01*
Panel B: Old EU markets									
Austria	−1.11	−1.12	−1.78*	−0.71	−0.53	−0.69	−0.71	−1.78*	−1.78*
Belgium	−0.18	−0.06	−0.80	−0.08	−0.05	0.42	0.18	−1.10	−1.01
Denmark	−0.69	−1.65	−1.58	−1.46	−0.85	−1.82*	−1.91*	−1.57	−1.57
Finland	−0.29	−1.27	−1.03	−1.12	−1.09	−1.40	−1.45	−1.12	−1.00
France	−2.76***	−3.26***	−3.14***	−1.95*	−1.84*	−2.17**	−2.19**	−3.14***	−3.14***
Germany	−2.73**	−4.31***	−3.41***	−1.92*	−2.29**	−3.75***	−3.22***	−3.13***	−3.17***
Italy	−0.35	−0.68	−0.57	−0.16	−0.24	−1.19	−1.37	−0.57	−0.57
Netherlands	−0.24	−0.38	−1.64	0.78	0.81	0.25	−0.02	−1.70*	−1.70*
Portugal	1.01	0.26	0.29	0.99	0.97	−0.62	−0.64	0.32	0.36
Spain	2.68**	2.17**	2.30**	2.38**	2.37**	1.97*	1.96*	2.20**	2.20**
Sweden	−1.27	−1.33	−1.78*	−1.49	−1.32	−1.45	−1.52	−1.84*	−1.85*
United Kingdom	−1.46	−1.39	−1.50	−1.62	−1.65	−1.61	−1.62	−1.50	−1.50

***, **, and * indicate 1%, 5%, and 10% level of significance. The null hypothesis is whether the loss differential is zero ($H_0 : l_t = 0$).

statistically supported in two countries (Slovakia and Spain). In the other countries marked in boldface (Belgium, the Netherlands, and Portugal), the differences are not statistically significant.

We now consider the reverse view: the results in which nowcasting (statistically significantly) outperforms the short-term statistical models; again, the results in Table 7 are based on those in Tables 3 and 4. When we compare nowcasting results with those for the short-term forecasting models (BVAR, FAVAR, and BFAVAR), for most countries the nowcasting significantly outperforms the short-term forecasting models. For example, BVAR is outperformed by nowcasting for eight out of thirteen countries, the FAVAR model for seven out of thirteen countries, and the BFAVAR model for eight out of thirteen countries. However, when we compare nowcasting to the AR, VAR, and FAAR models, for most countries, the differences are not statistically significant (e.g., the AR model is outperformed for only three out of thirteen countries, VAR is outperformed for only five, and FAAR is outperformed for three). The overall results are further detailed in the notes to Table 5 separately for the new and old EU economies.

In general, the results in Table 7 show that nowcasting significantly outperforms large-scale models such as FAVAR and BFAVAR for both EU groups. But when we compare nowcasting with the benchmark models, such as AR or VAR, then the differences between procedures are not statistically significant for most countries. This means that the results obtained by the benchmark models might be just as good as the results obtained with the nowcasting model. In other words, there is no sufficient evidence for preferring nowcasting over small-scale AR or VAR benchmark models, but there is strong evidence for preferring nowcasting over large-scale models such as FAVAR and BFAVAR. Finally, despite the statistical evidence, we should not overlook the economic significance of the nowcasting results. In particular, in the new EU markets with greater variation in output, nowcasting should be considered an efficient tool with clear predictive power.

Conclusions

We analyze the performance of a broad range of nowcasting and short-term forecasting models for a representative set of six new and twelve old EU members that are characterized by substantial differences in aggregate output variation. In our analysis, we generate ex-post out-of-sample nowcasts and forecasts based on hard and soft indicators from a comparable set of identical data.

Based on our results, we conclude that for most countries, the nowcasting algorithm outperforms the short-term forecasting models in terms of root mean squared errors. Our result show that nowcasting reduces forecasting errors and increases the accuracy of a one-step ahead forecast compared to the short-term forecasting models, even when the variation in GDP growth is relatively large. When we apply the DM statistic, we see that the nowcasting model significantly outperforms statistical models such as BVAR, FAVAR, and BFAVAR. However, when we compare nowcasting with the AR, VAR, and FAAR models, we conclude that the differences are not statistically significant for most countries.

Thus, using actual data, we observe that the nowcasting algorithm outperforms large-scale short-term forecasting models for most European countries in our sample. This is true even when the coefficient of variation in real GDP substantially changes over time and across countries. Further, nowcasting works well for new EU countries because, even though that variation in GDP growth data is larger than that of the old EU economies, the economic significance of nowcasting on average is somewhat greater than in old EU economies.

These results offer straightforward implications for policy makers, financial analysts, and economic actors: considering timely signals on the state of the economy improves the assessment of its current state even in times of economic fluctuation or distress. Thus, the nowcasting algorithm based on a dynamic factor model is a suitable candidate for generating an accurate one-period-ahead forecast of real GDP growth under uncertainty.

Notes

1. Nowcasting is the prediction of the present, the very near future and the very recent past in economics to monitor the state of the economy in real time.
2. Our motivation is further supported by findings of Benczúr and Rátfai (2014), who cover a large sample of countries prior to 2008, that includes also those analyzed by us, and show existence of strong degree of heterogeneity in macroeconomic behavior both across and within groups of emerging and developed countries. Further, they document differences in the volatility of the cyclical component of output. In general, they show that output is about twice as volatile in emerging market countries as in industrial countries.
3. Alternative short-term forecasting algorithms include the traditional autoregression (AR) model, factor-augmented autoregression (FAAR) model, unrestricted vector autoregressive (VAR) model, small-scale Bayesian VAR (BVAR), unrestricted factor augmented VAR (FAVAR), and Bayesian factor augmented VAR (BFAVAR).
4. In our analysis, we use a nowcasting algorithm proposed by Giannone, Reichlin, and Small (2008) because a large number of central banks have adopted it successfully. Further, our aim is to compare the most popular nowcasting method with other short-term forecasting algorithms, not with alternative nowcasting algorithms, such as MIDAS or mixed-frequency VAR. A comparison of various nowcasting algorithms is a topic for future research.
5. The MATLAB codes for nowcasting are available at https://www.newyorkfed.org/research/economists/giannone/pub/, which presents the computational steps of the nowcasting algorithms proposed by Giannone, Reichlin, and Small (2008) in full detail.
6. We assume that no structural relationship exists among endogenous variables and the existence of a left-hand side unity matrix (1 on the diagonal and 0 otherwise).
7. Following Blake and Mumtaz (2012), we use the Minnesota-type prior representing the belief that endogenous variables in a VAR model follow an AR(1) process or a random walk.
8. The MATLAB codes for two-step dynamic factor model is available at https://www.newyorkfed.org/research/economists/giannone/pub/, which presents the computational steps of the time domain algorithm proposed by Doz, Giannone, and Reichlin (2011) in full detail.
9. Camacho and Perez-Quiros (2010) use five hard indicators (Euro area industrial production index, excluding construction; Euro area total retail sales volume; industrial new orders index; total manufacturing orders; extra-Euro area exports) and five soft indicators (Belgium overall business indicator, Euro-zone economic sentiment indicator, Germany IFO business climate index, Euro area manufacturing purchasing managers index, Euro area services purchasing managers index).
10. Specifically, for our set of countries, the industrial production and retail trade indices are available at least four to seven weeks after the end of the current month (1–1.5 months). Similar lags are also present for imports and exports indices (3–5 weeks).
11. The countries are members of the OECD. We do not use the full OECD membership (36 countries) because of inconsistencies in data availability for the rest of the countries and because we are concentrating here on European countries.

12. We do not conduct a backcast because we do not have flash and first-revision GDP data. That is why we can generate a nowcast with revised data. Further, we do not perform one- or two-quarter-ahead forecasts because our target is to compare nowcasting and short-term forecasting algorithms for the current quarter.
13. This strategy is a good compromise among the standard in-sample and out-of-sample proportions of 50/50, 70/30, and 90/10 employed in modern machine learning modeling (see, e.g., https://machinelearningmastery.com/backtest-machine-learning-models-time-series-forecasting/).

Acknowledgments

We are thankful for valuable comments we received from Tomáš Adam, Joe Brada (Editor), Ali Kutan, Ashot Mkrtchyan, Armen Nurbekyan, Nerses Yeritsyan, and participants at several presentations. Kočenda acknowledges support from the Grant Agency of the Czech Republic grant No. 19-15650S. The views presented in this paper do not represent the official views of the Central Bank of Armenia. The usual disclaimer applies.

Disclosure statement

No potential conflict of interest was reported by the authors.

Funding

This work was supported by the Grantová Agentura České Republiky [19-15650S].

References

Aastveit, K., and T. Trovik. 2012. “Nowcasting Norwegian GDP: The Role of Asset Prices in a Small Open Economy.” *Empirical Economics* 42 (1):95–119. doi:10.1007/s00181-010-0429-9.

Adam, T., and F. Novotny (2018). Assessing the External Demand of the Czech Economy: Nowcasting Foreign GDP Using Bridge Equations. Working Papers 2018/18, Czech National Bank.

Alessi, L., M. Barigozzi, and M. Capasso. 2010. “Improved Penalization for Determining the Number of Factors in Approximate Factor Models.” *Statistics and Probability Letters* 80 (23--24):1806–13. doi:10.1016/j.spl.2010.08.005.

Alvares, R., C. Maximo, and G. Perez-Quiros. 2016. “Aggregate versus Disaggregate Information in Dynamic Factor Models.” *International Journal of Forecasting* 32 (3):680–94. doi:10.1016/j.ijforecast.2015.10.006.

Angelini, E., G. Camba-Mendez, D. Giannone, L. Reichlin, and G. Rünstler. 2011. “Short-Term Forecasts of Euro Area GDP Growth.” *Econometrics Journal* 14:25–44. doi: 10.1111/j.1368-423X.2010.00328.x.

Baffigi, A., R. Golinelli, and Parigi G. 2004. “Bridge Models to Forecast The Euro Area Gdp.” *International Journal Of Forecasting* 20 (3):447–460.

Bai, J., and S. Ng. 2002. “Determining the Number of Factors in Approximate Factor Models.” *Econometrica* 70 (1):191–221. doi:10.1111/ecta.2002.70.issue-1.

Banbura, M., D. Giannone, and L. Reichlin. 2010. “Large Bayesian Vector Auto Regressions.” *Journal of Applied Econometrics* 25 (1):71–92. doi:10.1002/jae.v25:1.

Banbura, M., D. Giannone, M. Modugno, and L. Reichlin (2013). Nowcasting and the Real Time Data. Working Paper Series 1564, European Central Bank.

Banbura, M., and G. Rünstler. 2011. “A Look into the Factor Model Black Box: Publication Lags and the Role of Hard and Soft Data in Forecasting GDP.” *International Journal of Forecasting* 27 (2):333–46. doi:10.1016/j.ijforecast.2010.01.011.

Barhoumi, K., O. Darne, and L. Ferrara. 2010. “Are Disaggregated Data Useful for Factor Analysis in Forecasting French GDP?” *Journal of Forecasting* 29:132–44. doi: 10.1002/for.1162.

Benczúr, P., and A. Rátfai. 2014. “Business Cycles around the Globe: Some Key Facts.” *Emerging Markets Finance and Trade* 50 (2):102–09. doi:10.2753/REE1540-496X500206.

Bernanke, B., J. Boivin, and P. Eliasz. 2005. “Measuring the Effects of Monetary Policy: A Factor-augmented Vector Autoregressive (FAVAR) Approach.” *Quarterly Journal of Economics* 120:387–422.

Blake, A.P., and H. Mumtaz. 2012. *Applied Bayesian Econometrics for Central Bankers*. Technical Books, Centre for Central Banking Studies, Bank of England, London, number 4.

Bragoli, D. 2017. “Nowcasting the Japanese Economy.” *International Journal of Forecasting* 33 (2):390–402. doi:10.1016/j.ijforecast.2016.11.004.

Bragoli, D., and J. Fosten. 2018. “Nowcasting Indian GDP.” *Oxford Bulletin of Economics and Statistics* 80 (2):259–82. doi:10.1111/obes.2018.80.issue-2.

Bragoli, D., L. Metelli, and M. Modugno. 2015. “The Importance of Updating: Evidence from a Brazilian Nowcasting Model.” *OECD Journal: Journal of Business Cycle Measurement and Analysis* 2015 (1):5–22.

Camacho, M., and G. Perez-Quiros. 2010. “Introducing the Euro-sting: Short-term Indicator of Euro Area Growth.” *Journal of Applied Econometrics* 25 (4):663–94. doi:10.1002/jae.1174.

Camacho, M., G. Perez-Quiros, and P. Poncela. 2013. “Short-term Forecasting for Empirical Economists: A Survey of the Recently Proposed Algorithms.” *Foundations and Trends in Econometrics* 6 (2):101–61. doi:10.1561/0800000018.

Camacho, M., M. Dal Bianco, and J. Martínez-Martín. 2015. “Short-run Forecasting of Argentine Gross Domestic Product Growth.” *Emerging Markets Finance and Trade* 51 (3):473–85. doi:10.1080/1540496X.2015.1025668.

Clements, M. P., and A.B. Galvão. 2008. “Macroeconomic Forecasting with Mixed Frequency Data.” *Journal Of Business & Economic Statistics* 26: 546-554. doi:10.1198/073500108000000015

Croushore, D., and T. Stark. 2001. “A Real Time Data Set for Macroeconomists.” *Journal of Econometrics* 105:111–30. doi: 10.1016/S0304-4076(01)00072-0.

D’Agostino, A., K. McQuinn, and D. O’Brien. 2012. “Nowcasting Irish GDP.” *OECD Journal: Journal of Business Cycle Measurement and Analysis* 2012 (2):21–31.

Diebold, F.X., and R.S. Mariano. 1995. “Comparing Predictive Accuracy.” *Journal of Business and Economic Statistics* 13:253–63.

Doz, C., D. Giannone, and L. Reichlin. 2011. “A Two-step Estimator for Large Approximate Dynamic Factor Models Based on Kalman Filtering.” *Journal of Econometrics* 164 (1):188–205. doi:10.1016/j.jeconom.2011.02.012.

Doz, C., D. Giannone, and L. Reichlin. 2012. “A Quasi-maximum Likelihood Approach for Large, Approximate Dynamic Factor Models.” *Review of Economics and Statistics* 94 (4):1014–24. doi:10.1162/REST_a_00225.

Forni, M., M. Hallin, M. Lippi, and L. Reichlin. 2005. “The Generalized Dynamic Factor Model: One-sided Estimation and Forecasting.” *Journal of American Statistical Association* 100:830–40. doi: 10.1198/016214504000002050.

Ghysels, E.P., P. Santa-Clara, and R. Valkanov. 2007. “MIDAS Regressions: Further Results and New Directions.” *Econometric Reviews* 26:53–90. doi: 10.1080/07474930600972467.

Giannone, D., L. Reichlin, and D. Small. 2008. “Nowcasting: The Real-time Information Content of Macroeconomic Data Releases.” *Journal of Monetary Economics* 55 (4):665–76. doi:10.1016/j.jmoneco.2008.05.010.

Giannone, D., S.M. Agrippino, and M. Modugno. 2013. *Nowcasting China Real GDP*. New York, NY: Mimeo.

Jansen, W. Jos., X.W. Jin, and Jasper M. de Winter. 2016. “Forecasting and Nowcasting Real GDP: Comparing Statistical Models and Subjective Forecasts.” *International Journal of Forecasting* 32 (2):411–36. doi:10.1016/j.ijforecast.2015.05.008.

Karasoy Can, H. G., and Ç. Yüncüler. 2018. “The Explanatory Power and the Forecast Performance of Consumer Confidence Indices for Private Consumption Growth in Turkey.” *Emerging Markets Finance and Trade* 54 (9):2136–52. doi:10.1080/1540496X.2017.1358608.

Kuzin, V., M. Marcellino, and C. Schumacher. 2011. "MIDAS Vs. Mixed-frequency Var: Nowcasting Gdp in The Euro Area." *International Journal Of Forecasting* 27 (2):529-542. doi:10.1016/j.ijforecast.2010.02.006.
Kuzin, V., M. Marcellino, and C. Schumacher. 2013. "Pooling versus Model Selection for Nowcasting with Many Predictors: An Application to German GDP." *Journal of Applied Econometrics* 28 (3):392–411. doi:10.1002/jae.2279.
Liebermann, J. 2012. "Short-term Forecasting of Quarterly Gross Domestic Product Growth." *Quarterly Bulletin Articles: Central Bank of Ireland* 3:74–84.
Luciani, M., M. P. Madhavi, A. Ramayandi, and G. Veronese. 2018. "Nowcasting Indonesia." *Empirical Economics* 55 (2):597–619. doi:10.1007/s00181-017-1288-4.
Marcellino, M., and C. Schumacher. 2010. "Factor-MIDAS for Now- and Forecasting with Ragged-edge Data: A Model Comparison for German GDP." *Oxford Bulletin of Economics and Statistics* 72:518–50. doi: 10.1111/j.1468-0084.2010.00591.x.
Pirschel, I., and Wolters, M. H. (2014). Forecasting German Key Macroeconomic Variables Using Large Dataset Methods. Kiel Working Papers 1925, Kiel Institute for the World Economy.
Poghosyan, K. 2016. "A Comparison of Different Short-term Macroeconomic Forecasting Models: Evidence from Armenia." *Journal of Central Banking Theory and Practice* 5 (2):81–99. doi:10.1515/jcbtp-2016-0012.
Porshakov, A., A. Ponomarenko, and A. Sinyakov. 2016. "Nowcasting and Short-term Forecasting of Russian GDP with a Dynamic Factor Model." *Journal of the New Economic Association* 30 (2):60–76. doi:10.31737/2221-2264-2016-30-2-3.
Rusnák, M. 2016. "Nowcasting Czech GDP in Real Time." *Economic Modelling* 54:26–39. doi: 10.1016/j.econmod.2015.12.010.
Stock, J., and M. Watson. 2002. "Macroeconomic Forecasting Using Diffusion Indexes." *Journal of Business and Economic Statistics* 20:147–62. doi: 10.1198/073500102317351921.
Yiu, M.S., and K.K. Chow. 2010. "Nowcasting Chinese GDP: Information Content of Economic and Financial Data." *China Economic Journal* 3 (3):223–40. doi:10.1080/17538963.2010.562028.

Appendix 1

In section 5, we discuss that, for out-of-sample forecast comparison, we use a recursive regression scheme. A forecasting model with a recursive window assumes that the initial estimation period is fixed, and additional observations are added to the estimation period one at a time. For the nowcasting model, we perform a recursive simulation experiment. Let us consider Germany as an example to explain the steps of experiments.

Table A1 shows that the available time period for Germany is from 1995Q1 to 2018Q4 (96 observations). Hence, according to the 70/30 rule, we have 67 observations for the in-sample period and 29 observations for the out-of-sample period. In monthly terms, it means that we have 201 (67*3) observations for the in-sample period and 87 (288–201) observations for the out-of-sample period. Therefore, the out-of-sample nowcast should start in 2011Q4 (or, in monthly terms, the starting date is 2011M10). As mentioned above, we do not have the actual data, so we use only revised data. We also mentioned that, at the end of current quarter, the hard data are available only for the first month (i.e., for 2011M10). Thus, for two remaining months, 2011M11 and 2011M12, we have missing values and deal with a ragged ends problem. However, the soft data are available for all three months, and their values are subject to minor or no revisions. Thus, using the method proposed by Giannone, Reichlin, and Small (2008), we have to extract the unobservable factors. As mentioned above, this method allows us to deal with missing data and ragged ends – therefore, we can compute the dynamics of unobservable factors until 2011M12. According to Giannone, Reichlin, and Small (2008), in the first step the principal components are estimated with using a balanced panel. Before applying the Kalman smoothing filter, the missing values are assigned arbitrary values. Then, the Kalman filter is applied to update the estimates of the factors. After estimating the monthly factors until 2011M12, we select only quarterly values, because real GDP growth is reported on a quarterly basis. Our goal is to forecast real GDP growth for 2011Q4, because we assume that for this quarter, we do not yet have its final value. To do so, we skip the quarter 2011Q4 and estimate the OLS

Table A1. The appropriate number of static factors.

Country	Sample period	No. observations for in-sample period	No. observations for out-of-sample period	No. additional explanatory variables	No. extracted factors with eigenvalues more than 1	Total variance explained, %
1	2	3	4	5	6	7
Austria	1997Q1 – 2018Q4	62	26	21	5	73.07
Belgium	1995Q1 – 2018Q4	67	29	21	6	76.44
Czech Republic	1996Q1 – 2018Q4	65	27	18	6	74.18
Denmark	1998Q1 – 2018Q4	59	25	17	6	77.20
Estonia	2000Q1 – 2018Q4	53	23	17	5	78.13
Finland	1997Q1 – 2018Q4	62	26	22	6	68.30
France	1991Q1 – 2018Q4	79	33	24	6	71.89
Germany	1995Q1 – 2018Q4	67	29	23	6	75.06
Hungary	1998Q1 – 2018Q4	59	25	20	5	73.30
Italy	1998Q1 – 2018Q4	59	25	22	6	75.58
Netherlands	1996Q1 – 2018Q4	65	27	20	6	75.14
Poland	1998Q1 – 2018Q4	59	25	17	5	79.89
Portugal	1997Q1 – 2018Q4	62	26	20	6	72.34
Slovakia	1996Q1 – 2018Q4	65	27	18	6	74.56
Slovenia	1999Q1 – 2018Q4	56	24	16	5	76.22
Spain	1997Q1 – 2018Q4	62	26	23	6	73.73
Sweden	1997Q1 – 2018Q4	62	26	21	6	80.49
United Kingdom	1998Q1 – 2018Q4	59	25	22	6	75.99

Characteristics for selecting the appropriate number of factors are indicated in parentheses – time spans used for unobservable factor extraction (2); number of observations for in-sample (3) and out-of-sample (4) periods; number of available additional explanatory variables (5); number of extracted factors with eigenvalues greater than 1 (6); total variance explained (7).

regression for the period 1995Q1-2011Q3 (which coincides with the in-sample period with 67 observations), where the dependent variable is real GDP growth, and the independent variables are extracted factors. Then, with estimated coefficients for the regression model along with the actual value of the extracted factors for 2011Q4, we can compute the value of real GDP for the fourth quarter of 2011. Afterward, we increase our sample by one observation (i.e., 1997Q1–2012Q1 or the months 2012M1, 2012M2, and 2012M3) and then we re-estimate the dynamic factors in the same way. Then again, we skip the most recent quarter (2012Q1) and estimate the regression model for the period 1997Q1-2011Q4 (which coincides with the in-sample period with 68 observations, because we have added one additional observation). After obtaining actual values of the dynamic factors for 2012Q1 we calculate the value of real GDP growth for 2012Q1. Continuing in this manner, we obtain 29 points one-step-ahead nowcasts for the German real GDP growth rate. In the same way, we conduct out-of-sample nowcasting experiments for the other countries in our analysis.

A slightly different experiment design is used for the short-term forecasting models (VAR, BVAR, FAVAR, and BFAVAR). For these models, the out-of-sample recursive experiments proceed as follows. Let us again consider the case of Germany. First, we estimate the factors for 1995Q1-2011Q3 (or, in monthly terms, 1995M1–2011M9), because we assume in this case that the actual values of the additional variables are unknown. Using a broad range of statistical models, we estimate the unknown parameters and generate a one-step-ahead forecast. Then, we increase the sample size by one (68 observations) and regenerate the one-step-ahead forecast. Continuing in this manner, we obtain 29 points for one-step-ahead forecasts.

Appendix 2

In addition to selecting the lag length, we also determine the optimal combination of static and dynamic factors. First, we determine the appropriate number of static factors. In doing so, we retain factors with eigenvalues of more than 1.[a] For example, Table A1 presents the number of additional explanatory variables (column 5), the number of extracted static factors (column 6), and the total variance explained by the extracted factors (column 7). Column 5 in Table A1 shows that the minimum number of additional variables used is 16 (Slovenia) and the maximum is 24 (France). The number of extracted static factors with eigenvalues of more than 1 is 5 or 6. The variance explained by the extracted static factors fluctuates between 68.30% and 80.49%. Thus, we see that most of the variance in the initial variables can be explained by only a few static factors.

Second, we select the number of dynamic factors. The number of dynamic factors cannot exceed the number of static factors (Forni et al. 2005). Therefore, we can restrict the number of dynamic factors to the maximum number of static factors. For example, we have 6 static factors for Germany (Table A1). Therefore, the maximum number of dynamic factors can be less or equal to 6. Then we chose different combinations of dynamic and static factors to obtain the maximum number of all possible combinations.[b] In our sample country (Germany), the maximum equals 21. The best of all possible combinations of static and dynamic factors is chosen based on the RMSE criterion.[c]

In Tables 3 and 4 for the AR model, we report the RMSE indices. To select the appropriate lag length, we run the AR model separately for 1, 2, 3, and 4 lags and select those that have lower RMSE values. We proceed in the same way for the unrestricted VAR, but, in contrast to the AR model, we run the model for four key macroeconomic variables, namely, the GDP growth rate, inflation, the nominal short-term interest rate, and the harmonized unemployment rate, as in Pirschel and Wolters (2014). Again, as with AR, we run VAR models separately for 1, 2, 3, and 4 lags and choose those lags based on the lowest RMSE value.

To run a small-scale Bayesian VAR, we go through the same steps as in the case of VAR. The difference is that, with Bayesian VAR, we must use two additional parameters: overall tightness and lag decay. Overall tightness is set at 0.1–0.3, with increments of 0.1. The decay factor takes values of 1 and 2. Thus, we run a grid search over all possible combinations of hyper-parameters and lag lengths. In our case, the lag length equals 1, 2, 3, and 4, overall tightness is 0.1, 0.2, and 0.3, and lag decay takes values of 1 and 2. Thus all possible combinations of hyper-parameters (overall tightness and lag decay) and lag length yield 24 BVAR models (for nineteen countries, it yields $24 \times 19 = 456$ models). As in the case of AR and VAR, the out-of-sample forecast accuracy is measured in terms of RMSE. We select the hyper-parameters and lag length by inspecting the pseudo out-of-sample forecast performance; the model with the minimum RMSE is selected as the best model, and its results are reported in Tables 3 and 4.

To estimate the FAAR model, we repeat the same steps as for the AR model. The main difference is that here we use additional factors. In Tables 3 and 4 we report the results for two FAAR models: the FAAR model with static factors and the FAAR model with dynamic factors. To select the appropriate combination of dynamic and static factors as well as the lag lengths, we go through all the possible combinations of dynamic and static factors.[d] Afterward, we select the appropriate number of static and dynamic factors and lag length by assessing the out-of-sample forecast performance: the FAAR model with the minimum RMSE is selected as the best model. In a similar manner, we select an appropriate model for the FAVAR and BFVAR models. The only difference is that here we use four target variables: GDP growth rate, inflation, nominal short-term interest rate, and unemployment rate.

Appendix Endnotes

a. An alternative is to use formal statistical tests. For example, the Bai and Ng (2002) information criterion can be used to determine the number of common factors. In addition, it is possible to implement the recently proposed criterion by Alessi, Barigozzi, and Capasso (2010). An alternative approach yields results that are not materially different.

b. Thus, we can conduct experiments for different combinations of dynamic and static factors: for example, one dynamic and one static factor. Then if we select two static factors, then we can have one dynamic and two static or two dynamic and two static factors. Then if we select three static factors, then the possible combinations can be one dynamic and three static, two dynamic and three static, and three dynamic and three static factors. Finally, if we select six static factors, then the possible combinations are one dynamic and six static, two dynamic and six static, three dynamic and six static, four dynamic and six static, and six dynamic and six static factors.

c. We should mention that all the necessary procedures for nowcasting and short-term forecasting were performed with software specially created for this purpose. This software is written in C#. NET and VBA (Visual Basic for Applications), which are powerful object-oriented programming languages. This software works directly in MS Excel 2010, 2013, and 2016 spreadsheets and can be downloaded from https://github.com/KarenPoghos/ForecastXL/.

d. We illustrate the issue with the following example. For Germany, we extracted 6 static factors and therefore have a total of 21 combinations. Considering that we run models for 4 different lag lengths, we have 84 scenarios in total (21 x 4). Then we recursively estimate each model and construct a one-step-ahead forecast.