Прикладная эконометрика, 2020, т. 61, с. 28–46.
Applied Econometrics, 2020, v. 61, pp. 28–46.
DOI: 10.22394/1993-7601-2021-61-28-46

K. Poghosyan, R. Poghosyan [1]

On the applicability of dynamic factor models for forecasting real GDP growth in Armenia

In this paper, we are trying to find out whether large-scale factor-augmented models can be successfully used for forecasting real GDP growth rate in Armenia. We compare the forecasting performance of factor-augmented models such as FAAR, FAVAR and Bayesian FAVAR with their small-scale benchmark counterpart models like AR, VAR and Bayesian VAR. Based on the ex-post out-of-sample recursive and rolling forecast evaluations and using RMSFE's, we conclude that large-scale factor-augmented models outperform small-scale benchmark models. However, the differences in forecasts among the models are not statistically significant when we apply statistical test.

Keywords: factor-augmented models; static and dynamic factors; recursive and rolling regression; out-of-sample forecast; RMSFE; Armenia.

JEL classification: E17; E37; C11; C15; C32; C53; C55.

1. Introduction

One of the main activities of the central banks is the use of modern forecasting methodology to conduct effective monetary policy. In the forecasting framework of the Central Bank of Armenia (hereafter CBA), the medium and long term forecasts of the key macroeconomic variables are based on the information obtained from the short-term forecasts (mainly one or two quarter ahead). Therefore, it is essential to the CBA to make the short-term forecasts as accurate as possible. For that, the CBA must constantly improve forecasting methodology. From this point of view, models with large datasets (or factor models) have become a popular tool for central banks for producing short-term forecasts. One of the important advantages of factor models is that potentially significant information is not neglected. There are many applications of dynamic factor models to forecasting macroeconomic and financial variables (Stock, Watson, 2002; Schumacher, 2007; Artis et al., 2005; Angelini et al., 2011; Matheson, 2006). The main finding of these applications is that the forecasts generated from the models with large datasets are superior to traditional small-scale benchmark models, like AR and VAR. In this paper, we want to consider the applicability of the large dataset models to real GDP growth rate forecasting in Armenia. We specifically assess real GDP growth forecast, because GDP is one of the most important indicator of economic activity and it is the main variable of interest that providing information about effectiveness in economic policy-making process.

[1] **Poghosyan, Karen** — Central Bank of Armenia, Yerevan, Armenia; karen.poghosyan@cba.am.
Poghosyan, Ruben — Yerevan State University, Yerevan, Armenia; poghosyan26@gmail.com.

In this paper we estimate a series of models that are frequently employed in the forecasting studies of most central banks. We employ three traditional small-scale benchmark models, namely univariate autoregression, unrestricted vector autoregression and Bayesian vector autoregression (hereafter AR, VAR and BVAR) and their factor-augmented counterparts, particularly, factor-augmented autoregression, factor-augmented vector autoregression and Bayesian factor-augmented vector autoregression (hereafter FAAR, FAVAR and BFAVAR). The factor-augmented models can be constructed in two steps: factor extraction, followed by model estimation and forecasting. Following Barhoumi et al. (2014), there are three main algorithms for extracting factors, namely static principal component, as in (Stock, Watson, 2002), dynamic principal components estimated in the time domain, as in (Doz et al., 2011, 2012) and dynamic principal components in the frequency domain, as in (Forni et al., 2000, 2005). All of these methods for factor extraction have the same purpose, namely, given a large number of initial variables, to extract only a small number of factors, which summarize most of the information contained in the whole dataset. In this paper, we use all of the aforementioned methods to extract the dynamics of unobservable factors. After the unobservable factors are extracted in the usual manner, they are added into standard small-scale forecasting models such as AR, VAR and BVAR, and then the factor augmented models are estimated and used to forecast the key macroeconomic variables.

To extract the dynamics of the factors we use Armenian actual quarterly macroeconomic time series from 1996Q1 to 2019Q4. The additional dataset includes 42 macroeconomic variables, comprising information on national accounts and consumer price indices, labor force and unemployment variables, monetary and financial variables and international macroeconomic variables. The main sources for our dataset are the Central Bank of Armenia (https://www.cba.am/) databases as well as external source databases, like World Bank (https://www.worldbank.org/), OECD (https://www.data.oecd.org/) and IndexMundi (https://www.indexmundi.com/). Using additional macroeconomic time series, we calculate the dynamics of unobservable factors with the help of one static and two dynamic algorithms (time and frequency domain). After extracting the dynamics of factors, we estimate the unknown parameters for all competing models included in our analysis.

A forecasting model with a good in-sample fit does not necessarily imply that it will have a good out-of-sample performance. For that, we also design out-of-sample forecast evaluation experiments based on the recursive and rolling regression scheme. Then using the results of the out-of-sample forecast evaluation we calculate root mean squared forecast error (RMSFE) indices. To keep robustness of our conclusions we conduct out-of-sample forecast experiments for different lag lengths and various combinations of dynamic and static factors. Based on the out-of-sample forecast evaluations and the calculated RMSFE indices, we conclude that models with large dataset are outperform small-scale benchmark models. However, the forecasts generated by the large-scale models are not statistically different from the forecasts generated by the small-scale models when we apply Diebold–Mariano statistical test.

The remaining paper is organized as follows. In section 2 we briefly present the main forecasting models. In section 3 we present the dynamics of actual macroeconomic variables and give some explanations for their fluctuations. In this section, we also consider the preliminary treatment of the additional explanatory variables, which we use for extraction of unobservable factors. In section 4 we explain in detail the experimental design that we use for recursive and rolling out-of-sample forecast evaluation. In section 5 we present the forecast evaluation results. The last section concludes the paper.

2. Literature review

As it was mentioned above the central banks are constantly involved in improving forecasting models to improve forecast accuracy and conduct more effective monetary policy. In this section we cover relevant literature on short-term forecasting models with large data set as well as on empirical contribution. The literature offers several approaches in terms of short-term forecasting models with large dataset, particularly static principal component, dynamic principal component in the frequency and time domain approaches.

In the paper by Stock and Watson (2002) they proposed univariate dynamic forecasting model augmented with static factors obtained by static principal component analysis. This method they have used to construct 6-, 12-, and 24-month-ahead forecasts for eight monthly US macroeconomic time series using 215 variables from 1970 to 1998. Based on the out-of-sample forecasts evaluation they concluded that new forecasts outperform univariate AR, small-scale VAR and leading indicator models. There are various other applications where the authors provided favorable evidence for the forecasting accuracy of the Stock and Watson (2002) static factor model. For example, Brisson et al. (2003) for Canadian data, Camacho and Sancho (2003) for Spanish data, Artis et al. (2005) for forecasting UK time series. From the other side there are a number of papers where the authors cast doubts on the empirical accuracy of large factor model based on static principal components. For example, Banerjee et al. (2005) compare static factor and single indicator forecasts for euro area variables and do not find improvements in the static factor models over single indicator approach. Other example is the paper by Schumacher and Dreger (2004), where the authors do not find significant advantages of factor models according to statistical tests of forecasting accuracy in similar experiments using German data. The Stock and Watson (2002) approach does not allow for use of the different dynamics that may exist between the variables used. To take account of this dynamic structure in factor models, several alternatives to the static factor model have been developed and suggested in the literature. Specifically, there are two main types of dynamic factor models. The first approach, proposed by Forni et al. (2000, 2005), is based on the frequency domain, while the second approach, developed by Doz et al. (2011, 2012), is based on the time domain.

In a series of articles, Forni et al. (2000, 2005) propose a dynamic principal component analysis in the frequency domain, also called a generalized dynamic factor model, to estimate dynamic factors. The method proposed by Forni et al. (2000, 2005), makes it possible to estimate dynamic factors in a first step and, then, obtain the static factors from the estimated dynamic factors in a second step. They discuss the theoretical advantages of dynamic over static model and show that the key advantage is that the dynamic model link variables at different point in times, while only contemporaneous variables enter in static models. However, despite its theoretical advantages, the empirical success of the dynamic approach does not seem to have been reached. For example, in (Forni et al., 2003) using dynamic factor model in the frequency domain, have found that the financial variables help to forecast inflation but not industrial production.

Another dynamic factor model approach proposed by Doz et al. (2011, 2012), is based on a state-space representation of the models in the time-domain. Specifically, the authors estimate their dynamic factor models using two different approaches, particularly two-step and quasi-maximum likelihood approaches. The two-step approach consists in first estimating the parameters by the standard static principal component. Then, in the second step, the factor dynamics are estimated via Kalman filtering and smoothing algorithm. A second approach is based on quasi-maximum likelihood estimations of an approximate dynamic factor model. The main idea is to treat the exact factor model as a misspecified approximating model and analyze the properties of maximum likelihood estimator under the different misspecifications. The authors have shown that the effect of misspecifications on the estimation of the factors is negligible for large sample size and cross-sectional dimensions.

In our paper we have used all above mentioned three approaches, particularly we discuss whether the more sophisticated dynamic factor models can significantly outperform more traditional small-scale benchmark models like AR, VAR and BVAR.

In this paper we contribute to the existing literature in two ways. First we provide a comprehensive comparison of the above mentioned three factor models using out-of-sample simulation experiments. From our knowledge, there are not yet any out-of-sample comparison incorporating both the Forni et al. (2000, 2005) and factor models proposed by Doz et al. (2011, 2012). In a number of empirical papers either Forni et al., or Doz et al., approach was used[2]. In our paper we also employ a recursive and rolling simulation scheme for out-of-sample forecasting. In most part of papers, the authors use mainly recursive regression scheme. Additionally, to keep the robustness of our results we apply various lag length and different combinations of dynamic and static factors. Among of all possible combinations we select the model which has a smallest value of RMSFE index. Second, we compare the three factor models and their accuracy using Armenian actual macroeconomic variables. We should mention that Armenia as a developing country has experienced more uncertainty. Developing countries tend to have the most volatile GDP growth rates. In the panel of 60 countries with available growth and financial data that Bloom (2014) examined, those with low incomes (less than $10000 GDP per capita) had 50 percent higher volatility of growth rates. He concludes, overall, developing countries experience about one-third higher macro uncertainty. Taking into account this fact it is worth to apply above mentioned approaches to a country like Armenia to check the applicability of the advanced forecasting methods to economies with relatively higher volatility of real GDP growth rates. For the Armenian economy a comparison of different large scale factor models has not been carried out yet, though there were some attempts to model real GDP growth and inflation using traditional approaches (AR, ARDL, VAR, BVAR, small-scale static factor models). For example, in the paper by Ghazaryan (2015) a system for short-term forecasting of private consumption and private investments is described. The main methods used are BVAR and FAVAR with only static factors. In sample quasi-real time recursive forecast evaluation shows that pooled forecasts outperform traditional univariate model forecasts. Another attempt to model inflation and monetary policy in Armenia by Dabla-Norris and Floerkemeier (2006) based on unrestricted VAR model. The authors have found that that the capability of monetary policy to influence on economic activity and inflation is still limited, as important channels of monetary transmission are not fully functional. In the paper by Dabla-Norris and Floerkermeier (2007) by employing traditional VAR approach the authors have shown while the CBA's focus on the repo rate is appropriate, the exchange rate would still need to be carefully monitored to the extent that it affects inflation. In the paper by Bordon and Weber (2010) by employing a Markov-switching VAR framework the authors analyzing whether the transmission mechanism in Armenia has been subject to a structural break. The authors have shown that reduced levels of dollarization are an important

---

[2] See (Angelini et al., 2010, 2011; Jovanovic, Petrovska, 2010; Godbout, Lombardi, 2012; Dias et al., 2015; Buss, 2010; Siliverstovs, Kholodilin, 2012; Ajevskis, Davidsons, 2008).

determinant of the effectiveness of monetary policy. Thus, we see that there has not been research on short-term forecasting the real GDP growth by employing dynamic factor model approaches. This paper fills the gap in the empirical literature.

3. A brief review of existing models

In this section, we present the basic forecasting models, particularly AR, VAR and BVAR and their factor-augmented counterpart models, FAAR, FAVAR and BFAVAR. We use the three small-scale models in order to evaluate the out-of-sample forecast performances of the three factor-augmented models. Below we briefly introduce each of them and discuss the main empirical aspects.

Univariate AR models are commonly used as benchmarks in the forecasting literature. Quite often AR models are perceived as advantageous compared to large multiple-equation models such as vector autoregression and traditional structural macroeconomic models (Hoffman, 2008; Arratibel et al., 2009). It is well known that the univariate AR model can be estimated by using the following regression model: $y_t = c + \sum_{j=1}^{p} \rho_j y_{t-j} + \varepsilon_t$. The unknown parameters of the model can be consistently estimated by using traditional OLS algorithm (Hamilton, 1994).

Another small-scale model that is used to forecast real GDP growth is the unrestricted VAR model. A standard VAR with p lags is expressed as $y_t = A_0 + A(L)y_t + \varepsilon_t$, where y_t is a $(n \times 1)$ vector of variables to be forecasted, A_0 is a $(n \times 1)$ vector of constant terms, $A(L)$ is a $(n \times n)$ polynomial matrix in the backshift operator L with lag length p, ε_t is a $(n \times 1)$ vector of error terms. In our case we assume that $\varepsilon_t \sim N(0, \sigma^2 I_n)$, where I_n is an $(n \times n)$ identity matrix. The unknown parameters of the VAR model can be consistently estimated by using traditional OLS algorithm (Hamilton, 1994). However, from the other side in the VAR model we very often need to estimate many parameters. This over-parametrization could cause inefficient estimates and hence a large out-of-sample forecast error. Thus, to overcome this over-parametrization we also implement the BVAR model. In order to use BVAR we first need to identify the priors. In this paper, we use the "Minnesota" type priors according to which the prior mean and standard deviation of the BVAR model can be set as follows:

1. The parameters of the first lag of the dependent variables follow an AR(1) process while parameters for other lags are equal to zero.
2. The variances of the priors can be specified as follows:

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

where i refers to the dependent variable in the j-th equation and j to independent variables in that equation, σ_i and σ_j are standard errors from AR(1) regressions estimated via OLS. The ratio of σ_i and σ_j controls for the possibility that variable i and j may have different scale (l is the lag length). The $\lambda's$ are set by the researcher and control the tightness of the priors. Thus, having "Minnesota" type priors it is possible to calculate the posterior parameters using the Bayesian approach to estimation.

Unlike small-scale benchmark models (AR, VAR and BVAR), the large-scale factor-augmented models include static or dynamic factors. As a rule, the factor-augmented models are estimated in two steps. First, we estimate the dynamics of unobservable factors using static and dynamic approaches and then we employ these extracted factors to forecast quarterly real GDP growth. In the modern time series econometrics literature there are three main algorithms for extracting factors, namely the static principal components as in (Stock, Watson, 2002), the dynamic principal component (frequency domain) approach as in (Forni et al., 2000, 2005) and the dynamic principal component approach (time domain) as in (Doz et al., 2011, 2012). There are a number of papers that present the computational steps of these factor models in great detail (Forni et al., 2000, 2005; Doz et al., 2011, 2012; Barhoumi et al., 2014).

Second, we add the extracted factors into the small-scale benchmark models as additional explanatory variables. Following Bernanke and Boivin (2003) we can present the factor augmented model as follows:

$$\begin{vmatrix} X_t \\ F_t \end{vmatrix} = A_0 + A_1 \begin{vmatrix} X_{t-1} \\ F_{t-1} \end{vmatrix} + A_2 \begin{vmatrix} X_{t-2} \\ F_{t-2} \end{vmatrix} + \ldots + A_p \begin{vmatrix} X_{t-p} \\ F_{t-p} \end{vmatrix} + \begin{vmatrix} \nu_t \\ u_t \end{vmatrix},$$

where X_t is the vector of observable variables, F_t is the vector of unobservable variables estimated using the three previously mentioned methods, $A_1, A_2, \ldots, A_p$ are ($r \times r$) matrices of estimated parameters, v_t and u_t are the error terms with zero mean and diagonal variance-covariance matrices. The above presented model can be consistently estimated by OLS and Bayesian approach (Hamilton, 1994).

4. Description of the dataset

For estimating small-scale benchmark models, namely AR, VAR and BVAR we use four key macroeconomic variables, particularly real GDP growth rates, CPI inflation, short-term nominal interest rates and unemployment rates. We closely follow approach by Pirschel and Wolters (2014) to select the macroeconomic time series to be included in the small-scale benchmark models. Thus, our dataset includes four key macroeconomic variables, which we mainly use in the small-scale benchmark models, and 42 additional macroeconomic variables, which we use to extract the dynamics of unobservable factors. According to (Barhoumi et al., 2014) our additional set of variables is medium-sized. Some studies have shown that the usage of smaller datasets which include about 10–30 series outperform the usage of larger datasets with disaggregated data with more than 100 series (Alvarez et al., 2016). Our dataset is balanced and in quarterly terms from 1996Q2 to 2019Q4. The dataset has 95 observations for each variable. We choose quarterly time series because we want to discuss the empirical properties of the dynamic factor models with respect to real GDP growth, which is available in the quarterly frequency.

The first important macroeconomic variable is the real GDP growth rate with respect to the previous quarter. In order to calculate the values of the growth rate we made the following calculations. First the absolute values of the real GDP in terms of average 2005 prices were transformed to natural logarithms. Then transformed values were seasonally adjusted (X12 ARIMA method) and first differenced. As a result, we obtain the real GDP growth rates (Figure 1).

From Figure 1 we can clearly see that real GDP growth rate series have shown little persistence and can thus be expected to be hard to predict. Also we can indicate that after years of sluggish growth that followed the 2008–2009 global financial crisis, the Armenian real GDP growth rate expanded by 7.5% in 2017. Its growth slowed to 5.2% in 2018, due to a dramatic regime change in the country. In 2019 the real GDP growth for Armenia was 7.6%, which is the largest recorded growth since 2008. This growth was caused by the increase in consumption of households and supported by stronger export growth. The increase in consumption was led by household credit, up by 30% in 2019 and by a 10% increase in money transfers from abroad. On the production side, growth was led by the service sector growth following an acceleration in tourism output and domestic trade. Industry output has also expanded strongly, driven by a rebound in mining production.

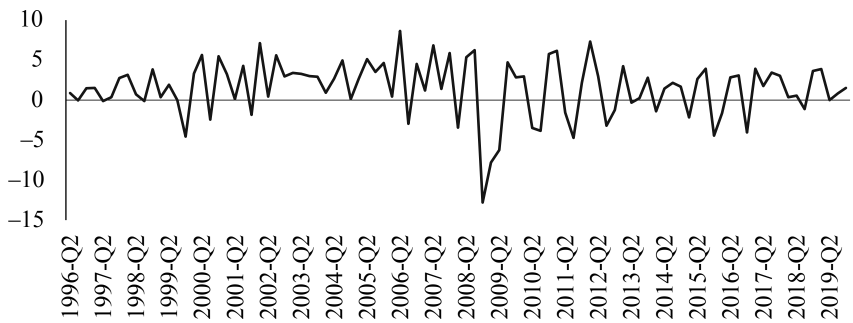

Fig. 1. Real GDP growth (%-th change to the previous quarter)

The second macroeconomic variable is the CPI inflation with respect to the previous quarter. In order to calculate the values of CPI inflation we made the following preliminary calculations. We first transformed the CPI month to month indices to the base month indices. Then we transformed these values to natural logarithms, seasonally adjusted the results and calculated the first differences. Figure 2 presents the seasonally adjusted CPI inflation rates with respect to previous quarter.

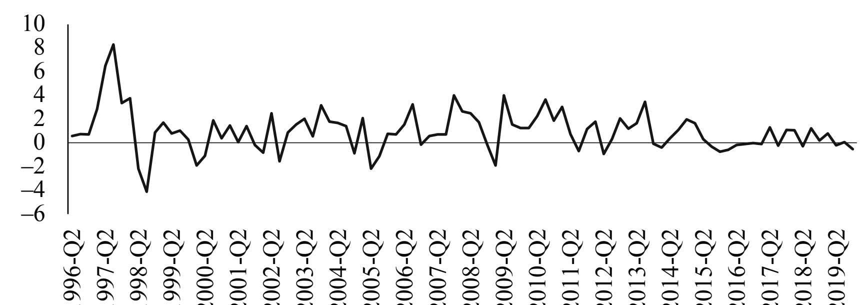

Fig. 2. CPI inflation (%-th change to the previous quarter)

From Figure 2 we can see that CPI inflation is more persistent than real GDP growth, but it shows many spikes which will be hard to predict. In Armenia the calculations of monthly CPI start from 1993 and it is the only indicator characterizing inflation dynamics. The Central Bank of Armenia targeted monetary aggregates prior to 2006, but after 2006 it switched to inflation targeting through interest rates, as managing the monetary aggregates proved ineffective due to the large inflow of remittances from abroad. The inflation target was initially 3.0% for 2006, and changed only once in 2007; from 2007 onward it is maintained at 4.0% with a confidence band of ±1.5%. Even after the global financial crisis the inflationary pressure remained low with an average annual inflation rate of 1.4% in 2019 (down from 2.5% in 2018), well below the lower band of the Central Bank of Armenia's inflation target range.

The third macroeconomic variable is the short-term nominal interest rate for time deposits in national currency. This time series is much more persistent than real GDP growth or CPI inflation. The preliminary treatment for this variable includes only first differences (in percentage points). The short-term nominal interest rate shows an overall downward trend. For example, as we see in Figure 3 the nominal interest rate is characterized by relatively large fluctuations before 2005, but since 2006 the fluctuations of interest rates have become smaller. Such behavior could be explained by fact that before 2006 the CBA policies targeted monetary aggregates, while after 2006 the CBA adopted the inflation-targeting regime.

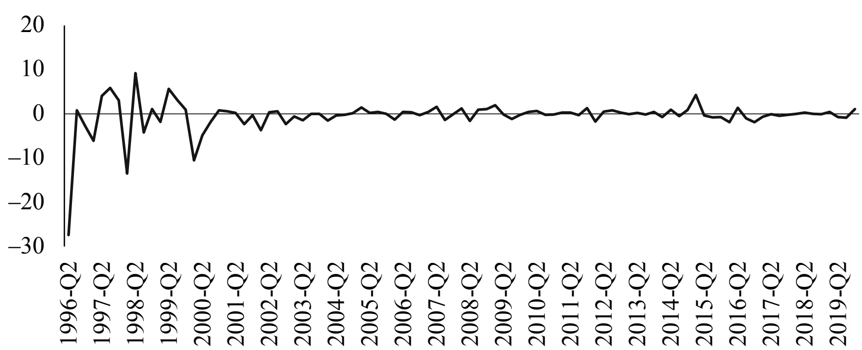

Fig. 3. Short-term nominal interest rate (percentage points change to the previous quarter)

The next macroeconomic variable is the total unemployment rate. The labor market in Armenia has improved, but the unemployment rate remains exceptionally high at about 18%. A large segment of population remains employed in agriculture and the informal sector. According to the International Labor Organization (ILO) estimates, the ratio of self-employed in total employment is still large at about 40%. The unemployment rate dynamics with respect to the previous quarter were calculated by the authors of this paper. The calculations have been done as follows. First, we obtained the official values for unemployment (in persons) in yearly terms from the World Bank development indicators. Then, using temporal decomposition method, particularly Boot et al. (1967) mechanical projection algorithm, we decomposed the yearly unemployment data into the quarterly frequencies. Finally, we transformed the unemployment data into natural logarithms and calculated the first differences (Figure 4).

After all transformations we have to check whether the key macroeconomic variables have become stationary, since our models can be fitted only to the stationary series. To check whether the key macroeconomic variables appears to be stationary, below we present the results of the formal statistical tests for unit roots.

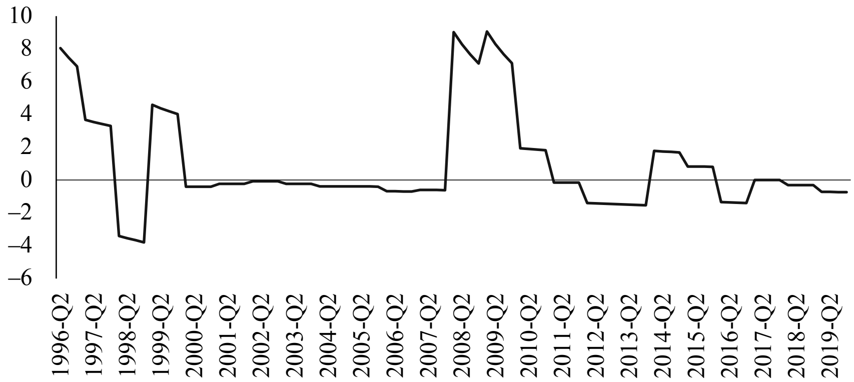

Fig. 4. Total unemployment rate (%-th change to the previous quarter)

Table 1. Results of unit root tests for transformed values of the key macroeconomic variables

	Augmented Dickey–Fuller test			Phillips–Perron test		
	Without C,T	C	T	Without C,T	C	T
Real GDP growth	–8.294 (0)	–9.597 (0)	–9.631 (0)	–8.309 (0)	–9.627 (0)	–9.675 (0)
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Inflation	–4.323 (1)	–5.395 (1)	–5.549 (1)	–5.325 (1)	–6.432 (1)	–6.559 (1)
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Interest rate	–8.129 (3)	–8.169 (3)	–8.175 (3)	–17.770 (3)	–17.853 (3)	–17.897 (3)
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Unemployment rate	–3.628 (1)	–3.633 (1)	–3.573 (1)	–3.588 (1)	–3.586 (1)	–3.526 (1)
	(0.00)	(0.01)	(0.03)	(0.00)	(0.01)	(0.04)

Notes. Lags in unit root tests were determined automatically using the SIC criteria. The number in parentheses behind the test statistic is the number of lags. *P*-values are in parentheses below the test statistic. C — constant, T — trend.

The results of the unit root tests in Table 1 confirm that the seasonally adjusted and first differentiated series follow a stationary pattern and that at the risk of error of 5%, the null hypothesis of unit root existence is rejected.

Besides these four key macroeconomic variables, our dataset also includes additional 42 variables, as mentioned above. This set of additional variables can be grouped into the following categories: national accounts, consumer and producer prices, labor force and unemployment, monetary and financial variables and international variables on growth rates and price indexes. Most of the time series are obtained from CBA internal databases. Some of the other time series were obtained from different sources, particularly from Index Mundi (IM), OECD and World Bank (WB). A detailed data description is provided in Appendix. The vector of time series presented in Appendix was preliminary treated. First, the time series were corrected for outliers and then seasonally adjusted as explained in Appendix. All nonstationary time series were transformed to be stationary, by taking first differences. We formally check the stationarity of the transformed additional series by using ADF (augmented Dickey–Fuller test) unit root tests. Based on the results of the tests we have concluded that all transformed series are stationary. From Appendix we also see, that time series included in national accounts group comparing with other time series are relatively volatile. This is because the standard deviations are relatively high and the first lag autocorrelations are small and negative. Positive autocorrelation is an indication of a specific form of persistence, the tendency of a time series to remain in the same state from one observation to the next. Hence, we conclude that real GDP growth rates is a little persistence (because of its components are volatile) and it may be difficult to predict. Finally, the series were normalized to have mean zero and unit variance. Most of the calculations were done using MATLAB (r2018a) code. The MATLAB code for extracting unobservable components in time and frequency domain was obtained from the internet sources[3]. Some other code was written by the authors of the paper. For example, we have C# codes for time domain factor model, as well as recursive and rolling regressions, which can be conducted directly from MS Excel spreadsheet (download for free from https://github.com/KarenPoghos/ForecastXL).

5. Experimental design

To conduct out-of-sample forecast evaluation we use both recursive and rolling regression schemes. For out-of-sample forecast evaluation we divide the whole sample into two subsamples, namely in-sample and out-of-sample periods. The first period is the training sample (in-sample), and the second period is the forecasting sample (out-of-sample). In our experiments the in-sample period includes 70% of observations, while out-of-sample period 30% of observations. The 70/30 proportion is a good compromise among the standard in-sample and out-of-sample proportions of 50/50, 70/30 and 90/10 broadly employed in modern machine learning algorithms[4]. After choosing the proportion between in-sample and out-of-sample periods the recursive simulation scheme proceeds as follows.

First, we estimate the models using subsample 1996Q2–2012Q4 (67 observations). Using estimated model, we generate and then to store 1 to 4 steps-ahead forecasts results. Then we increase the sample size by one (68 observations, 1996Q2–2013Q1) and generate again 1 to 4 steps-ahead forecasts and then we store the forecast results. We continue increasing the sample size by one and generating 1 to 4 steps-ahead forecasts until the sample size 91 (1996Q2–2018Q4). Then we increase the sample size by one but only generate 1 to 3 steps-ahead forecasts (since we only have 92 observations in total). We continue increasing the sample size until we have 94 observations in the sample, in which case we can only compute the 1-step-ahead forecast. In this way, we obtain 28 1-step-ahead forecasts, 27 forecasts for 2-steps-ahead, 26 for 3-steps-ahead and finally 25 forecasts for 4-steps-ahead.

For the rolling forecast scheme, the initial sample is the same as in recursive scheme, but when the additional observation is added after the first forecast, the first values of the initial estimation sample are also deleted. Hence, while in the recursive scheme the sample size increases by one

---

[3] The corresponding MATLAB code for factor model proposed by Doz et al. (2011, 2012) can be found at https://www.newyorkfed.org/research/economists/giannone/pub, MATLAB code for factor model proposed by Forni et al. (2005) can be found at http://www.barigozzi.eu/Codes.html.

[4] https://machinelearningmastery.com/backtest-machine-learning-models-time-series-forecasting/.

quarter at each step, in the rolling scheme the estimation size remains constant. The rolling regression scheme proceeds as follows.

First we fix the sample size at 67 observations. As in the recursive regression scheme the forecasts are computed with a forecast horizon from 1 up to 4 and the results are stored. Then we add one observation to the sample and delete the first observations (in total we have 67 observations). Then we generate again 1 to 4 steps-ahead forecasts and the results are stored. Continuing in this manner we obtain the same number of forecasts as in the case of recursive regression. The recursive simulation scheme has an advantage of using all available data at a certain point of time, while the rolling scheme skips available information. But considering that our time series includes the episodes of global financial crisis we have assumed that rolling regression scheme may be useful to use.

Next, we use the out-of-sample forecasts from recursive and rolling regression to compute the corresponding root mean squared forecast error (RMSFE) indices for each of the fourth forecasting horizons. More formally, we denote the out-of-sample period by T^* and forecast horizons $h = 1,2,3,4$. Then the RMSFE index is calculated by the following formula:

$$RMSFE_h = \sqrt{\frac{1}{T^* - (h-1)} \sum_{t=1}^{T^*-(h-1)} (\hat{y}_t - y_t)^2},$$

where RMSFE is the root mean squared forecast error for the h-th forecast horizon, $\hat{y}_t$ is the forecasted value of the real GDP growth, y_t is the actual value of the real GDP growth.

6. Empirical results

In this section we present the out-of-sample forecast results for 15 competing models. To keep robustness of our results we have estimated models with different lags length and different combinations of dynamic and static factors. Following (Pirshel, Wolters, 2014; Jos Jansen et al., 2016), we vary the number of lags from 1 up to 4 lags. In addition, we vary the number of static and dynamic factors, enabling us to consider all possible combinations. Thus, varying both the number of lags and the number of static and dynamic factors we compare the estimated models to each other. Finally, we select the lags length and number of static and dynamic factors by looking at the pseudo out-of-sample forecast performances, particularly we select the combination that minimizes the RMSFE, evaluated over the entire out-of-sample period. The model with the smallest RMSFE index is selected as a model for forecasting at all horizons. Now, we will explain how we determine the number of static and dynamic factors in more detail for each model separately.

For the standard AR(p) model we only vary the number of lags from 1 up to 4. We conduct an out-of-sample forecast evaluation for four models, particularly models with $p = 1,...,4$ lags and choose the model with the smallest RMSFE. We use the same approach for the VAR model. The only difference is that the VAR model includes four explanatory variables (real GDP growth, inflation, short-term nominal interest rate and unemployment rate). Again we conduct an out of sample forecast evaluations for four models with $p = 1,...,4$. We chose the model with the smallest RMSFE. We use approximately the same approach for small-scale BVAR model. The only difference is that in addition to varying lags, we also vary overall tightness and lag decay. Following the papers (Gupta, Kabundi, 2011; Kočenda, Poghosyan, 2020) we set the overall tightness 0.1 to 0.3 with increment equal to 0.1. The decay factor takes values of 1 and 2. Thus all possible combination for lags length and hyperparameters yield 24 BVAR models. Then we select the optimal combinations of lags and parameters by looking at the pseudo out-of-sample forecast performances and we select the model with the smallest value of RMSFE. Selected models with optimal lags and number of dynamic and static factors are presented in the Tables 3 and 4.

For FAAR_SW[5] model, in contrast above models, we vary both the number of lags and the number of static factors. Now we will explain how we estimate the number of static factors. Taking into account that additional set of variables is medium-sized, we estimate the number of static factors as follows. To estimate the number of static factors we retain in the analysis only the factors with eigenvalues more than 1[6]. Using this simple rule, we have extracted 12 static factors. In the Table 2 is presented the total variance explained by the extracted 12 static factors. Also we present the contributions of each variables group in total variance of each factor separately.

Table 2. Total variance explained by factors and variable groups

	Factors											
	1	2	3	4	5	6	7	8	9	10	11	12
National accounts	0.71	0.15	2.66	**2.79**	0.85	**2.41**	**2.73**	**2.47**	0.86	**1.41**	0.88	**0.83**
Consumer and producer prices	0.57	0.92	0.31	0.32	**3.33**	1.09	0.35	0.51	0.09	0.57	**0.99**	0.73
Labor force and unemployment	0.06	**9.73**	0.75	1.92	0.07	0.13	0.55	0.42	0.24	0.22	0.10	0.04
Monetary and financial variables	0.66	1.61	**7.16**	1.85	0.18	1.18	0.41	0.41	**1.25**	0.20	0.64	0.26
International variables	**13.22**	0.20	1.12	0.71	1.83	0.12	0.42	0.20	0.80	0.65	0.05	0.58
% in total variance	15.21	12.61	12.01	7.57	6.24	4.94	4.46	4.02	3.24	3.06	2.67	2.45
Cumulative % in total variance	15.21	27.83	39.84	47.41	53.66	58.60	63.06	67.08	70.32	73.38	76.04	78.49

From Table 2 we see that the first extracted factor explains the 15.21% of total variance of 42 initial variables. The most part of the "Factor 1" total variance is explained by the "International variables" group (13.22%). The second factor explains 12.61% of total variance of 42 initial variables, but the most contribution in the factor variance has "Labor force and unemployment" group. In a such manner we can observe the contributions of all factors and variable groups. Also based on the values of contributions we can give a name for particular factor. For example, the first factor we can title as "Oil and energy price index", the second — "Labor variable's index", the third — "Monetary index", and so on. The name of factors we also can formulate based on a separate variable included in a group. For example, in the fourth factor the largest contribution has "National accounts" group. The same group has the largest contribution in the factors 6, 7, 8 and 10. Thus, in such situation we can formulate the factor name based on a particular variable which is included in a specific group of variables. For example, the "Factor 4" can be titled as "Foreign trade index", which is included in a "National accounts" group.

After extracting the static factors we examine all possible combination of lags ($p = 1,\ldots,4$) and static factors ($r = 12$), which yield 48 FAAR_SW models. Based on the out-of-sample forecast evaluation results we choose the model with the smallest RMSFE. For FAAR_FHLR, FAAR_2S

[5] FAAR_SW is a FAAR model with static factors (Stock, Watson, 2002).

[6] But for some relatively large additional datasets (with more than 100 variables) it is more appropriate to use some developed statistical tests, for example (Bai, Ng, 2002; Alessi et al., 2010).

and FAAR_QML[7] we vary the number of lags as well as the number of dynamic and static factors. To select the appropriate number of dynamic factors we follow the principal according to which the number of dynamic factors cannot exceeds the number of static factors (Forni et al., 2005; Jos Jansen et al., 2016). For example, if we have 3 static factors then the number of dynamic factors can range from 1 to 3. In other words, we can construct the following combinations for dynamic and static factors: 1 dynamic and 3 static, 2 dynamic and 3 static and finally 3 dynamic and 3 static factors. We use this idea to construct of all possible combinations of dynamic and static factors. Thus, if we estimate 12 static factors then according to our approach, we can have 78 combinations for dynamic and static factors in total. Taking into account that we also vary the number of lags ($p=1,...,4$) then all possible combinations yield 312 FAAR_FHLR, 312 FAAR_2S and 312 FAAR_QML models. We should note that the QML algorithm is an iterative procedure and for each model we run 100 iterations. Finally, we choose the model with the smallest RMSFE (Tables 3 and 4).

For the FAVAR_SW, FAVAR_FHLR, FAVAR_2S and FAVAR_QML the selection procedures is the same as in the case of FAAR, with the only difference being that in the FAVAR models there are 4 target variables (real GDP growth rate, inflation rate, short-term interest rate and unemployment rate). For BFAVAR_SW model we also vary additional hyperparameters (overall tightness and lag decay). As mentioned above, we vary overall tightness from 0.1 to 0.3 with increments equal to 0.1. The decay takes a value of 1 or 2. Thus, varying all inputs, that is number of lags, the number of static factors and hyperparameters, all possible combinations yield 288 BFAVAR_SW models. Then as in the case of previous models we choose the model with smallest RMSFE. For BFAVAR_FHLR, BFAVAR_2S and BFAVAR_QML the number of possible combinations is much higher, because for these models we also vary the number of dynamic factors. Thus varying all possible inputs parameters, we get 1872 combinations in total for each model separately. Again we choose the model with smallest RMSFE and store the results in the Tables 3 and 4.

Tables 3 and 4 present the results for various forecast horizons. We see that models with large data set always outperforms small-scale benchmark models at all forecast horizons. For example, as we see from Table 3 for the recursive regression the FAAR_2S model outperforms all small scale benchmark models producing the minimum value of RMSFE's. For two step forecast horizon the FAAR_QML outperforms all small-scale benchmark models producing the minimum RMSFE's. For three and four steps ahead forecast horizons the best model is the BFAVAR_2S. We can reach the same conclusions using the results presented in Table 4, that is in the case of rolling regression the large data set models outperform small-scale benchmark models.

The next question that arises from Tables 3 and 4 is whether the differences between forecasts generated by the large- and small-scale models are significantly different. To give an answer on this question we have to conduct the equal forecast accuracy test, particularly Diebold–Mariano (1995) test. Before presenting the results of the tests we present some explanations related to this test.

In this paper we calculate the Diebold–Mariano statistic by regressing the loss differential on an intercept, using heteroscedasticity autocorrelation robust (HAC) standard errors (Diebold, 2015). Let ε_t^{AR} denote the forecast errors in the benchmark AR(*p*) model and ε_t^{i} denote the forecast errors in the competing *i*-th short term forecasting models (*i* = FAAR_SW, FAAR_FHLR,

[7] FAAR_FHLR is a FAAR model with dynamic factors estimated in the frequency domain (Forni et al., 2005), FAAR_2S and FAAR_QML is a FAAR model estimated in the time domain (Doz et al., 2011, 2012). In the same way it is possible to explain the abbreviations for FAVAR and BFAVAR models.

Table 3. RMSFE indices for the real GDP growth (recursive regression scheme)

Model	Forecast horizon			
	$h = 1$	$h = 2$	$h = 3$	$h = 4$
AR ($p = 2$)[8]	2.379	2.367	2.435	2.473
VAR ($p = 1$)	2.443	2.502	2.489	2.512
BVAR ($p = 1, w = 0.3, d = 1$)[9]	2.473	2.479	2.479	2.505
FAAR_SW ($p = 2, r = 5$)	2.078	2.200	2.390	2.475
FAAR_FHLR ($p = 2, q = 4, r = 5$)[10]	2.066	2.237	2.388	2.438
FAAR_2S[11] ($p = 2, q = 3, r = 5$)	***1.939***	2.282	2.576	2.500
FAAR_QML ($p = 2, q = 5, r = 5$)	2.208	***2.118***	2.401	2.490
FAVAR_SW ($p = 1, r = 1$)	2.365	2.601	2.402	2.415
FAVAR_FHLR ($p = 1, q = 1, r = 1$)	2.317	2.536	2.437	2.431
FAVAR_2S ($p = 1, q = 1, r = 1$)	2.444	2.685	2.398	2.408
FAVAR_QML ($p = 1, q = 1, r = 1$)	2.319	2.427	2.503	2.434
BFAVAR_SW ($p = 1, w = 0.3, d = 1$)	2.309	2.569	2.379	2.405
BFAVAR_FHLR ($p = 1, q = 1, r = 1, w = 0.3, d = 1$)	2.295	2.511	2.390	2.459
BFAVAR_2S ($p = 1, q = 1, r = 1, w = 0.3, d = 1$)	2.344	2.634	***2.364***	***2.391***
BFAVAR_QML ($p = 1, q = 1, r = 1, w = 0.3, d = 1$)	2.231	2.418	2.423	2.416

FAAR_QML, FAAR_2S). Then the loss differential l_t can be calculated as $l_t = (\varepsilon_t^{AR})^2 - (\varepsilon_t^i)^2$. Thus, we regress the loss differential on an intercept using HAC standard errors. The null hypothesis is that the loss differentials equal to zero ($\mathbf{H_0}: \ l_t = 0$). The results of *t*-statistics obtained from regressing the loss differentials on the intercept both for recursive and rolling regressions are presented in the Table 5.

The statistics presented in Table 5 indicates whether the performance results of large-scale and small-scale benchmark models are significantly different. From Table 5 we see that when we compare the predictive accuracy of the large-scale models with small-scale models then for both recursive and rolling regressions the differences are not statistically significant. In other words, there is not sufficient evidence to favor large-scale models over small-scale benchmark models. This means that the forecasting results for real GDP growth rate obtained by the small-scale benchmark models could be just as good as the results obtained from models based on large data set.

---

[8] *p* is the number of lags in the model. The number in the brackets says that the smallest RMSFE has been achieved in the case of 2 lags. For all other models in the Tables 3 and 4 *p* has the same meaning.

[9] *w* and *d* are hyperparameters that we use in the BVAR and BFAVAR models. The first coefficient (overall tightness) we have implemented to the diagonal matrix of the variances, while the second coefficient (decay) is implemented to the lags. In the Tables 3 and 4 we have presented the values of *w* and *d* for which the model has the smallest RMSFE.

[10] *r* is the number of static factors; *q* is the number of dynamic factors. In the Tables 3 and 4 we have presented the values of *r* and *q* for which the model has the smallest RMSFE.

[11] FAAR_QML is more accurate method than FAAR_2S. In a FAAR_QML the FAAR_2S serve as an initial step for iterations. The main difference between FAAR_QML and FAAR_2S is that in the FAAR_QML is iterative procedure and the parameters are updated at each iteration, while desired correctness will be achieved. In the case of FAAR_2S the Kalman filtering and smoothing is using only in one iteration and parameters are not updated.

Table 4. RMSFE indices for the real GDP growth (rolling regression scheme)

Model	Forecast horizon			
	$h = 1$	$h = 2$	$h = 3$	$h = 4$
AR ($p = 2$)	2.383	2.376	2.435	2.471
VAR ($p = 1$)	2.482	2.550	2.517	2.537
BVAR ($p = 2, w = 0.1, d = 1$)	2.549	2.447	2.458	2.494
FAAR_SW ($p = 2, r = 2$)	2.243	2.468	2.504	2.486
FAAR_FHLR ($p = 2, q = 4, r = 5$)	***2.223***	2.508	2.492	2.443
FAAR_2S ($p = 2, q = 3, r = 5$)	2.504	2.581	***2.327***	2.402
FAAR_QML ($p = 2, q = 1, r = 2$)	2.343	2.371	2.531	2.512
FAVAR_SW ($p = 1, r = 1$)	2.537	2.626	2.413	2.407
FAVAR_FHLR ($p = 1, q = 1, r = 1$)	2.486	2.544	2.459	2.435
FAVAR_2S ($p = 1, q = 1, r = 1$)	2.629	2.736	2.416	2.394
FAVAR_QML ($p = 1, q = 1, r = 1$)	2.489	2.605	2.563	2.432
BFAVAR_SW ($p = 1, r = 1, w = 0.3, d = 1$)	2.372	2.558	2.366	2.401
BFAVAR_FHLR ($p = 1, q = 1, r = 1, w = 0.3, d = 1$)	2.352	2.488	2.381	2.446
BFAVAR_2S ($p = 1, q = 1, r = 1, w = 0.3, d = 1$)	2.406	2.626	2.351	***2.384***
BFAVAR_QML ($p = 1, q = 1, r = 1, w = 0.3, d = 1$)	2.282	***2.368***	2.436	2.426

Table 5. Diebold–Mariano statistics

Compared Model	Forecast horizon			
	$h = 1$	$h = 2$	$h = 3$	$h = 4$
Recursive regression scheme				
FAAR_SW versus AR	0.71	0.82	0.52	–0.03
FAAR_FHLR versus AR	0.09	–0.90	–1.36	–0.21
FAAR_2S versus AR	1.08	0.65	–1.04	–0.31
FAAR_QML versus AR	0.60	0.29	–0.20	–0.18
FAVAR_SW versus VAR	0.24	–0.73	0.82	1.35
FAVAR_FHLR versus VAR	0.09	–1.15	0.30	1.07
FAVAR_2S versus VAR	0.00	–1.05	0.76	1.27
FAVAR_QML versus VAR	–0.28	–1.21	0.56	1.35
BFAVAR_SW versus BVAR	0.78	–0.97	1.09	1.55
BFAVAR_FHLR versus BVAR	0.55	–1.23	0.79	1.43
BFAVAR_2S versus BVAR	0.58	–1.23	1.04	1.50
BFAVAR_QML versus BVAR	0.51	–1.37	0.94	1.59
Rolling regression scheme				
FAAR_SW versus AR	0.33	–0.67	–0.72	–0.16
FAAR_FHLR versus AR	–0.01	— 1.79*	–0.61	0.65
FAAR_2S versus AR	0.14	–1.49	–1.21	–0.15
FAAR_QML versus AR	0.04	–1.50	–1.52	–1.42
FAVAR_SW versus VAR	–0.16	–0.43	0.88	1.19
FAVAR_FHLR versus VAR	–0.36	–1.40	–1.03	–0.54
FAVAR_2S versus VAR	–0.42	–0.83	0.75	1.18
FAVAR_QML versus VAR	–0.78	–1.15	0.64	1.25
BFAVAR_SW versus BVAR	0.71	–1.00	0.92	1.08
BFAVAR_FHLR versus BVAR	0.44	–1.50	–0.67	–1.22
BFAVAR_2S versus BVAR	0.53	–1.27	0.89	1.09
BFAVAR_QML versus BVAR	0.39	–1.56	0.79	1.16

Note. * indicates 10% level of significance.

7. Conclusion

We analyze the forecast performances of the 15 competing short-term forecasting models. In our analysis we generate ex-post out-of-sample forecasts based on the quarterly actual Armenian time series. For the ex-post out-of-sample simulations we use both recursive and rolling regression schemes. Based on the recursive and rolling forecast simulation results we conclude that out-of-sample forecasts obtained by the large-scale factor augmented models outperform forecasts obtained by the small-scale benchmark models at all forecast horizons. Based on these results we conclude that the forecasts of the real GDP growth rate obtained by large-scale models are more appropriate from the practical point of view. Then, in order to check whether the differences in forecasts obtained by the different models are statistically significant we apply Diebold–Mariano test. We conduct this test both for recursive and rolling regression schemes. Based on the results of this test we conclude that there is not sufficient evidence to favor large-scale over small-scale models. This means that the forecast results obtained for real GDP growth rate by using the small scale models would not be statistically different from the results obtained by the large scale models.

References

Alessi L., Barigozzi M., Capasso M. (2010). Improved penalization for determining the number of factors in approximate factor models. *Statistics and Probability Letters*, 80 (23–24), 1806–1813.

Alvarez R., Camacho M., Perez-Quiros G. (2016). Aggregate versus disaggregate information in dynamic factor models. *International Journal of Forecasting*, 32 (2), 680–694.

Angelini E., Banbura M., Rünstler G. (2010). Estimating and forecasting the euro area monthly national accounts from a dynamic factor model. *OECD Journal: Journal of Business Cycle Measurement and Analysis*, 2010 (1), 1–22.

Angelini E., Camba-Méndez G., Giannone D., Rünstler G., Reichlin L. (2011). Short-term forecasts of Euro area GDP growth. *Econometrics Journal*, 14 (1), 25–44.

Ajevskis V., Davidsons G. (2008). Dynamic factor models in forecasting Latvia's gross domestic product. *Working Papers* 2008/02, Latvijas Banka.

Arratibel O., Kamps C., Leiner-Killinger N. (2009). Inflation forecasting in the new EU member states. *ECB Working Paper* No. 1015.

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

Bai J., Ng S. (2002). Determining the number of factors in approximate factor models. *Econometrica*, 70 (1), 191–221.

Banerjee A., Marcellino M., Masten I. (2005). Leading indicators for Euro-area inflation and GDP growth. *Oxford Bulletin of Economics and Statistics*, 67 (s1), 785–813.

Barhoumi K., Darné O., Ferrara L. (2014). Dynamic factor models: A review of the literature. *OECD Journal: Journal of Business Cycle Measurement and Analysis*, 2013 (2), 73–107.

Bernanke B., Boivin J. (2003). Monetary policy in a data-rich environment. *Journal of Monetary Economics*, 50 (3), 525–546.

Bloom N. (2014). Fluctuations in uncertainty. *Journal of Economic Perspectives*, 28 (2), 153–176.

Bordon A. R., Weber A. (2010). The transmission mechanism in Armenia: New evidence from a regime switching VAR analysis. *IMF Working Paper* No. 10/270.

Boot J. C. G., Feibes W., Lisman J. H. C. (1967). Further methods of derivation of quarterly figures from annual data. *Applied Statistics*, 16 (1), 65–75.

Brisson M., Campbell B., Galbraith J. W. (2003). Forecasting some low-predictability time series using diffusion indices. *Journal of Forecasting*, 22, 515–531.

Buss G. (2010). A note on GDP now-/forecasting with dynamic versus static factor models along a business cycle. *MPRA Paper* 22147. Munich, Germany.

Camacho M., Sancho I. (2003). Spanish diffusion indexes. *Spanish Economic Review*, 5, 173–203.

Dabla-Norris E., Floerkermeier H. (2006). Transmission mechanisms of monetary policy in Armenia: Evidence from VAR Analysis. *IMF Working Paper* No. 06/248.

Dabla-Norris E., Floerkermeier H. (2007). Bank efficiency and market structure: What determines banking spreads in Armenia? *IMF Working Paper* 07/134.

Dias F., Maximiano P., Rua A. (2015). Forecasting Portuguese GDP with factor models: Pre- and post-crisis evidence. *Economic Modelling*, 44 (C), 266–272.

Diebold F., Mariano R. (1995). Comparing predictive accuracy. *Journal of Business and Economic Statistics*, 13 (3), 253–263.

Diebold F. (2015). Comparing predictive accuracy, twenty years later: A personal perspective on the use and abuse of Diebold–Mariano tests. *Journal of Business and Economic Statistics*, 33 (1), 1.

Godbout C., Lombardi M. J. (2012). Short-term forecasting of the Japanese economy using factor models. *Staff Working Papers* 12–7. Bank of Canada.

Doz C., Giannone D., Reichlin L. (2011). A two-step estimator for large approximate dynamic factor models based on Kalman filtering. *Journal of Econometrics*, 164 (1), 188–205.

Doz C., Giannone D., Reichlin L. (2012). A quasi-maximum likelihood approach for large, approximate dynamic factor models. *The Review of Economics and Statistics*, 94 (4), 1014–1024.

Gupta R., Kabundi A. (2011). A large factor model for forecasting macroeconomic variables in South Africa. *International Journal of Forecasting*, 27 (4), 1076–1088.

Jos Jansen W., Xiaowen J., de Winter J. M. (2016). Forecasting and nowcasting real GDP: Comparing statistical models and subjective forecasts. *International Journal of Forecasting*, 32 (2), 411–436.

Jovanovic B., Petrovska M. (2010). Forecasting Macedonian GDP: Evaluation of different models for short-term forecasting. *Working paper* 2010–02. National Bank of the Republic of North Macedonia.

Forni M., Hallin M., Lippi M., Reichlin L. (2000). The generalized dynamic-factor model: Identification and estimation. *The Review of Economics and Statistics*, 82 (4), 540–554.

Forni M., Hallin M., Lippi M., Reichlin, L. (2003). Do financial variables help forecasting inflation and real activity in the Euro area? *Journal of Monetary Economics*, 50 (6), 1243–1255.

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

Ghazaryan N. (2015). Short term forecasting system of private demand components in Armenia. *Armenian Journal of Economics*, 1, 21–52.

Hamilton J. D. (1994). *Time series analysis*. Princeton University Press.

Hoffman B. (2008). Do monetary indicators lead euro area inflation? *ECB Working Paper* No. 867.

Kočenda E., Poghosyan K. (2020). Nowcasting real GDP growth: Comparison between old and new EU countries. *Eastern European Economics*, 58 (3), 197–220.

Matheson T. (2006). Factor model forecasts for New Zealand. *International Journal of Central Banking*, 2 (2), May.

Pirschel I., Wolters M. (2014). Forecasting German key macroeconomic variables using large dataset methods. *Kiel Working Papers* 1925. Kiel Institute for the World Economy.

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

Schumacher C., Dreger C. (2004). Estimating large-scale factor models for economic activity in Germany: Do they outperform simpler models? *Journal of Economics and Statistics*, 224 (6), 731–750.

Siliverstovs B., Kholodilin K. (2012). Assessing the real-time informational content of macroeconomic data releases for now-/forecasting GDP: Evidence for Switzerland. *Journal of Economics and Statistics*. 232 (4), 429–444.

Stock J., Watson M. (2002). Macroeconomic forecasting using diffusion indexes. *Journal of Business and Economic Statistics*, 20 (2), 147–162.

Received 28.07.2020; accepted 20.09.2020.

Appendix

Dataset description

Series description	Source	SA	Transf.	Mean	Stand. deviation	$\rho(1)$	ADF	Corr. with real GDP growth
			National accounts					
Value added in industry, q/q, %	CBA	Yes	ln and Δ	1.15	4.77	–0.14	–11.01	0.437
Value added in agriculture, q/q, %	CBA	Yes	ln and Δ	0.84	7.11	–0.16	–11.27	0.479
Value added in construction, q/q, %	CBA	Yes	ln and Δ	1.56	11.92	–0.06	–10.46	0.646
Value added in services, q/q, %	CBA	Yes	ln and Δ	1.84	3.46	–0.29	–12.99	0.359
Final consumption, q/q, %	CBA	Yes	ln and Δ	1.26	3.52	–0.14	–10.99	0.443
Private consumption, q/q, %	CBA	Yes	ln and Δ	1.27	3.75	–0.13	–10.94	0.425
Government consumption, q/q, %	CBA	Yes	ln and Δ	1.26	9.42	–0.35	–13.71	0.071
Gross fixed capital formation, q/q, %	CBA	Yes	ln and Δ	1.35	12.24	–0.20	–11.76	0.561
Exports of goods and services, q/q, %	CBA	Yes	ln and Δ	1.74	10.90	–0.27	–12.63	0.164
Imports of goods and services, q/q, %	CBA	Yes	ln and Δ	0.75	7.06	0.03	–9.27	0.407
			Consumer and producer prices					
CPI-food prices, q/q, %	CBA	Yes	ln and Δ	0.90	2.39	0.25	–7.44	0.074
CPI-nonfood, q/q, %	CBA	No	ln and Δ	0.65	1.46	0.13	–8.44	0.029
CPI-services, q/q, %	CBA	No	ln and Δ	1.54	2.90	0.49	–6.32	–0.133
Ind. production price index, q/q, %	CBA	No	ln and Δ	1.50	3.68	0.26	–7.48	0.342
Construction price index, q/q, %	CBA	No	ln and Δ	1.09	3.15	0.25	–7.66	0.304
Tariffs for transportation, q/q, %	CBA	No	ln and Δ	3.08	10.17	–0.11	–10.73	0.018
			Labor force and unemployment					
Labor force, q/q, %	WB	No	ln and Δ	0.03	0.31	0.90	–4.32	–0.027
Employment, q/q, %	WB	No	ln and Δ	–0.08	0.55	0.80	–4.28	0.203

End of the table

Series description	Source	SA	Transf.	Mean	Stand. deviation	$\rho(1)$	ADF	Corr. with real GDP growth
Employment in industry q/q, %	WB	No	ln and Δ	–0.04	1.14	0.78	–4.33	0.303
Employment in agriculture q/q, %	WB	No	ln and Δ	–0.53	0.61	0.81	–4.28	–0.07
Employment in services q/q, %	WB	No	ln and Δ	0.27	0.70	0.80	–4.27	0.219
Self-employed, total q/q, %	WB	No	ln and Δ	–0.31	0.65	0.83	–3.98	0.074
Employment vulnerable, q/q, %	WB	No	ln and Δ	–0.33	0.64	0.84	–3.85	0.069
Monetary and financial variables								
Monetary base, q/q, %	CBA	Yes	ln and Δ	3.93	5.68	0.04	–9.41	0.285
Cash money in circulation, q/q, %	CBA	Yes	ln and Δ	3.19	6.00	0.31	–7.26	0.416
Broad money, q/q, %	CBA	Yes	ln and Δ	4.54	3.78	0.31	–7.53	0.462
Total deposits, q/q, %	CBA	Yes	ln and Δ	5.27	4.86	0.16	–9.07	0.243
Firms time deposits, q/q, %	CBA	Yes	ln and Δ	4.72	20.40	0.00	–9.73	–0.03
Households time deposits, q/q, %	CBA	Yes	ln and Δ	6.07	6.02	0.38	–7.06	0.134
Total time deposits, q/q, %	CBA	Yes	ln and Δ	5.96	7.39	0.27	–7.74	0.075
Total loans, q/q, %	CBA	Yes	ln and Δ	4.78	5.81	0.41	–6.99	0.237
Interest rates for loans, pp	CBA	No	Δ	–0.83	4.86	–0.32	–15.20	–0.056
International variables								
EA GDP growth rate, q/q, %	OECD	Yes	ln and Δ	0.39	0.58	0.59	–4.92	0.322
EA industrial production, q/q, %	OECD	Yes	ln and Δ	0.25	1.60	0.54	–5.22	0.376
Russia industrial production, q/q, %	OECD	Yes	ln and Δ	0.69	2.14	0.33	–6.74	0.346
Russian natural gas, q/q, %	OECD	No	ln and Δ	1.38	12.17	0.44	–5.56	0.142
Crude oil; Dated Brent, q/q, %	IM	No	ln and Δ	2.38	14.38	0.23	–7.55	0.289
Crude oil; Dubai Fateh, q/q, %	IM	No	ln and Δ	2.52	14.52	0.18	–7.96	0.305
Crude oil; West Texas, q/q, %	IM	No	ln and Δ	2.22	14.36	0.18	–8.03	0.296
Fuel price index, q/q, %	IM	No	ln and Δ	1.90	12.66	0.28	–7.17	0.338
Non-fuel price index, q/q, %	IM	No	ln and Δ	0.51	5.71	0.41	–6.18	0.421
All commodity price index, q/q, %	IM	No	ln and Δ	1.01	7.53	0.41	–6.22	0.434

Notes. SA — seasonal adjustment, Transf. — transformation (ln — natural logarithm, Δ — first difference), $\rho(1)$ — first lag autocorrelation, ADF — augmented Dickey–Fuller unit root tests. According to ADF unit root tests for all 42 additional time series the null hypothesis is rejected. In the last column is presented the correlation coefficients between real GDP growth rates and additional variable.