JEL Classification: C11, C13, C52, C53
Keywords: factor-augmented models, static and dynamic factors, recursive and rolling regression, out-of-sample forecast, RMSFE, Armenia

On the Applicability of Dynamic Factor Models for Forecasting Real GDP Growth in Armenia

Karen POGHOSYAN - Central Bank of Armenia, Macroeconomic Department, Yerevan, Armenia (karen.poghosyan@cba.am) corresponding author

Ruben POGHOSYAN - Yerevan State University, Faculty of Physics, Yerevan, Armenia

Abstract

In this paper, we are trying to find out whether large-scale factor-augmented models can be successfully employed for forecasting real GDP growth rate in Armenia. We use Armenian data because as a developing country Armenia has experienced a relatively higher volatility of GDP growth rate in comparison to other countries. Based on our calculation using growth rate data from 40 countries, we argue that low-income countries have about 57% higher volatility of growth rates than high-income countries. Taking this into account, it is worth testing the forecasting performance of factor models on a country like Armenia to check the applicability of the advanced forecasting methods to economies with highly volatile growth rates. For this, 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, Bayesian VAR and mixed-frequency 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 when we apply these methods to forecasting real GDP growth. However, the differences in forecasts among the models are not statistically significant when we apply statistical test.

1. Introduction

One of the main activities of 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 quarters 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 and Watson, 2002; Schumacher, 2007; Artis et al., 2005; Angelini et

https://doi.org/10.32065/CJEF.2021.01.03

We would like to thank the two anonymous referee and participants at several presentations for their suggestions and useful comments.

al., 2011; Matheson, 2006). The main finding of these applications is that the forecasts generated from the models with large datasets perform, on average, slightly better than traditional small-scale benchmark models, like AR and VAR (Eickmeier and Zigler, 2008). In this paper, we want to consider the applicability of the large dataset models to real GDP growth forecasting in Armenia. We specifically assess real GDP growth forecast, because GDP is one of the most important indicators of economic activity, and it is the main variable of interest that provides information about effectiveness in the economic policy-making process.

We use an Armenian macroeconomic dataset, because as a developing country Armenia has experienced relatively higher volatility of GDP growth rate in comparison to other countries. From Appendix Table A1 we see that for Armenia the coefficient of variation of GDP growth rate is 5.62 %, which is the third highest among the developing countries included in our analysis. We also see from Appendix Table A1 that of the 40 countries, 16 are classified as low-income and 24 as high-income. If we calculate the average values for the coefficient of variation, we find that in low-income countries it is about 3.81 %, while in high-income countries it is 2.42 %. Using these values, we can calculate that low-income countries have 57 % higher volatility of growth rates. As will be shown in the next section, the bulk of working papers devoted to dynamic factor models are applied to data from developed and emerging economies (Eickmeier and Zigler, 2008). Therefore, it is worth testing the forecasting performance of dynamic factor models on a country like Armenia to check the applicability of the advanced forecasting methods to other economies with relatively low income and high volatility of growth rates.

Also we use Armenian as a model economy, because it is attractive choice from both theoretical and practical point of view. We can point out the following reasons. First, Armenia's economic freedom score is 71.9, making its economy the 32nd (up from 34th last year) freest in the world, according to the 2021 Index of Economic Freedom published by the Heritage Foundation. Second, inflation rate in Armenia relatively lower than in other transition economies. Investment (% of GDP) ratio is lower than in the Baltic states, but it is close to CIS (Commonwealth Independent States) group. Third, Armenia's financial intermediation is lower than the Baltic states, but close to the CIS group and Armenia is one of the few transition countries that never operated under a fixed exchange rate regime after gaining independence (Poghosyan et al., 2008).

In this paper we estimate a series of models that are frequently employed in the forecasting studies of most central banks. We employ the following small-scale benchmark models: univariate autoregression, unrestricted vector autoregression, Bayesian vector autoregression and mixed-frequency vector autoregression (hereafter AR, VAR, BVAR, MF-VAR, respectively). We also employ their factor-augmented counterparts, particularly factor-augmented autoregression, factor-augmented vector autoregression and Bayesian factor-augmented vector autoregression (hereafter FAAR, FAVAR and BFAVAR, respectively). 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 and 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. We select quarterly data, because there is some evidence that quarterly data lend themselves better to factor forecasts than monthly data (Eickmeier and Zigler, 2008). The additional dataset includes 42 quarterly 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 algorithm 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) values. 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 datasets 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 the Diebold-Mariano (1995) statistical test.

The remaining paper is organized as follows. Section 2 reviews the existing literature. In section 3 we present the main forecasting models and give some intuitive descriptions of the models. Section 4 presents the dynamics of actual macroeconomic variables and gives 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 5 we explain in detail the experimental design that we use for recursive and rolling out-of-sample forecast evaluation. In section 6 we present the forecast evaluation results. The last section concludes the paper.

2. Literature Review

As mentioned above, 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 datasets as well as on empirical contributions. The literature offers several approaches in terms of short-term forecasting models with large datasets, particularly static principal component and dynamic principal component in the frequency and time domains.

In the paper by Stock and Watson (2002) they proposed univariate dynamic forecasting model augmented with static factors obtained by static principal component analysis. They used this method 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 and Artis et al. (2005) for forecasting UK time series. From the other side there are a number of papers where the authors cast doubt on the empirical accuracy of large factor models 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 the single indicator approach. Another example is the paper by Schumacher and Dreger (2003), where the authors do not find significant advantages of factor models according to statistical tests of forecasting accuracy in a similar experiment 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 account for 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 reach. 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 first step of the two-step approach involves estimating the parameters by the standard static principal component. Then, in the second step, the factor dynamics are estimated via Kalman filtering and a 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 estimators under the different misspecifications. The authors have shown that the effect of misspecifications on the estimation of the factors is negligible for large sample sizes and cross-sectional dimensions. These factor models were further developed in the following papers, among others: Li et al. (2017), Forni et al. (2017), Forni et al. (2018), Fiorentini et al. (2018), Ma and Su (2018), Baltagi et al. (2017), Su and Wang (2017).

There are many applications of the three above-mentioned factor extraction methods to forecast real GDP growth and inflation. The main finding of these applications is that, on average, factor forecasts are slightly better than other models forecasts. In particular, factor models tend to outperform small-scale models, whereas they perform slightly worse than alternative methods which are also able to exploit large datasets (Eickmeier and Zigler, 2008). In this paper, we have used all three above-mentioned approaches. In particular, we discuss whether the more sophisticated dynamic factor models can significantly outperform more traditional small-scale benchmark models like AR, VAR, BVAR and MF-VAR. At the same time, we also compare the forecasting performances of three factor models, using actual Armenian macroeconomic variables. We find that the dynamic factors estimated by the approach of Doz et al. (2011, 2012) perform better for real GDP growth forecasting.

Why did we choose data on Armenia, and can the lessons learned from Armenia be useful for practitioners from the other countries?

We perform our analysis on data from Armenia, because as a developing country Armenia has experienced relatively higher volatility of GDP growth rate. In general, developing countries tend to have the most volatile GDP growth rates (Appendix Table A1). In Appendix Table A1 we present the descriptive statistics for real GDP growth rate of 40 countries. These countries are classified as either high- or low-income countries based on the GNI per capita values. The countries with GNI per capita of less than 10000 US dollars are classified as low-income, while all others are classified as high-income. Using coefficient of variations (Appendix Table A1, column 8) it is possible to calculate the average volatility of GDP growth for both high- and low-income countries. From our calculations, we have obtained that the average volatility of GDP growth rate in low-income countries is about 3.81%, while in high-income countries it is 2.42%. From these values, we calculate that low-income countries have 57% higher volatility of growth rates. Our finding closely resembles that of Bloom (2014), according to which in a panel of 60 countries, those with low incomes (less than 10000 GDP per capita) had 50 percent higher volatility of growth rates. Taking this into account, it is useful to apply the 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. Therefore, if we find that above-mentioned methods work for Armenia, then these techniques could be applied also to other developing countries that have similar economic characteristics. From this point of view, the results obtained for Armenia potentially could be of interest for practitioners from other developing countries to improve their forecast.

In this paper we contribute to the existing literature in two ways. First, we provide a comprehensive comparison of the broad range of factor models. From our knowledge, there are not yet any systematic out-of-sample comparisons incorporating all above-mentioned factor models with an application to developing countries. There are a number of papers, that apply factor models to data from developing countries, including Corona et al. (2017), Gunay (2018), Lopez-Buenache (2018), Abdić et al. (2020), Camacho et al. (2015), and Porshakov et al. (2016). In these papers, the authors use mainly one or two factor extraction methods and the number of models compared to factor models is relatively small. In contrast, based on our analysis, we are able to compare factor models not only with a broad range of forecasting models, but we can also compare three factor extraction methods to each other. This is important, because thus far the comparisons between different factor extraction models have been based on data from advanced and emerging economies (Eickmeier and Zigler, 2008). Furthermore, the choice between static and dynamic factors in forecasting GDP has gone unresolved until now. From this point of view our paper attempts to answer on this question by using macroeconomic variables from a developing country. Second, 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 with using traditional approaches (AR, ARDL, VAR, BVAR, small-scale static factor models)[1]. 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. We briefly introduce each of them below 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 algorithms (Hamilton, 1994).

[1] See Ghazaryan N. (2015), Dabla-Norris E. and Floerkemeier H. (2006), Bordon A. and Weber A. (2010).

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 and 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 algorithms (Hamilton, 1994).

However, 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. Following (Gupta and Kabundi, 2011), we use the Litterman/Minnesota (Litterman, 1980) type prior. The Litterman/Minnesota prior imposes the hypothesis that individual variables all follow random walk processes. This specification typically performs well in forecasts of macroeconomic time series (Kilian and Lutkepohl, 2017) and is often used as a benchmark to evaluate accuracy. According to this approach 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:

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

Where i refers to the dependent variable in the j-th equation and j to independent variables in that equation and 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, by having Litterman/Minnesota type priors it is possible to calculate the posterior parameters using the Bayesian approach to estimation.

$$\overline{\beta} = \left(\underline{H}^{-1} + \Sigma^{-1} \otimes X_t^{'} X_t\right)^{-1} \left(\underline{H}^{-1} \underline{\beta} + \Sigma^{-1} \otimes X_t^{'} X_t \hat{b}\right)$$
$$\text{var}(\overline{\beta}) = \left(\underline{H}^{-1} + \Sigma^{-1} \otimes X_t^{'} X_t\right)^{-1}$$

Where $\overline{\beta}$ is the vector of the posterior parameters, $\underline{\beta}$ is the vector of the prior parameters, $\underline{H}$ is the diagonal matrix with the prior variances on the diagonal, X is the $(T \times k)$ matrix of the initial time series, and Σ is assumed to be known (we replace Σ by the estimate $\hat{\Sigma}$ (variance-covariance matrix of the residuals)). As an alternative to the Litterman/Minnesota prior, we also use other priors, particularly normal-flat, independent normal-Wishat and Sims-Zha priors. We now discuss these alternatives in more detail.

The normal-flat prior relaxes the assumption of Σ being known but imposes no prior beliefs about Σ. Thus, Σ is taken from an "uninformative" prior, and as such has no prior distribution. For Σ, all of the prior information is contained in the degrees-of-freedom parameter, μ (Koop and Korobilis, 2010). Calculating the conditional posteriors only requires specifications of $\underline{\beta}$ (priors for parameters), $\underline{H}$ (priors variances for prior parameters) and $\underline{\mu}$ (prior for $\Sigma \sim IW(\underline{S}, \underline{\mu})$, where IW denotes the Inverse Wishart distribution (for more details see Koop and Korobilis, 2010). The priors for $\underline{\beta}$ is constructed identically to the Litterman/Minnesota prior. The covariance terms are constructed as: $\underline{H}^{-1} = c_2 I_k$, $\underline{\mu} = c_3$, where, c_2 and c_3 are hyper-parameters, I_k is identity matrix of size $k \times k$ ($k = Mp + d$, where M is the number of endogenous variables, p is the number of lags and d is the number of exogenous variables).

The independent normal-Wishart prior again relaxes the assumption of Σ being known, this time imposing a prior distribution (Koop, Korobilis, 2010). For the independent normal-Wishart model, the priors for Σ are given $\Sigma \sim IW(\underline{S}, \underline{\mu})$. $\underline{\beta}$ is constructed identically to the Litterman/Minnesota prior. The covariance terms are constructed as: $\underline{S}^{-1} = c_1 I_M$, $\underline{H}^{-1} = c_2 I_{Mk}$, $\underline{\mu} = c_3$, where, c_1, c_2 and c_3 are hyper-parameters, I_M is identity matrix of size $M \times M$ (M is the number of endogenous variables).

The construction of the Sims-Zha prior employs the structural form of the VAR (Sims and Zha, 1998). Sims-Zha place a conditional prior using either the normal-flat or normal-Wishart distributions. Sims-Zha suggest a form of $\underline{\alpha}$ $\underline{H}$ along similar lines to the Litterman/Minnesota prior. First, $\underline{\alpha}$ is set to a vector of nearly all zeros, with only the elements corresponding to the coefficient of a variable's own first lag, hyper-parameter μ_1, being non-zero. Second $\underline{H}$ is assumed to be diagonal:

- The diagonal elements corresponding to the *j*-th endogenous variables in the *i*-th equation at lag *l* are given by:

$$\left(\frac{\lambda_0 \lambda_1}{\sigma_j l^{\lambda_3}}\right)^2$$

- The diagonal elements corresponding to the constants and to the remaining exogenous are:

$$(\lambda_0 \lambda_4)^2, \ (\lambda_0 \lambda_5)^2$$

Where, $\lambda_0, \lambda_1, \lambda_3, \lambda_4$ and λ_5 are hyper-parameters, and σ_j is the square root of the corresponding diagonal element of Σ.

In addition to the small-scale VAR and Bayesian VAR we also using the Mixed-Frequency VAR approach. In this model we consider only two types of frequencies, particularly quarterly and monthly, which means that there are 3 high frequency periods per quarter. Then using the U-MIDAS approach we stack the high and low frequency variables into the corresponding matrices, and using classical VAR least squares we estimate the matrix of unknown parameters (Ghysels, 2016).

Unlike small-scale benchmark models (AR, VAR, BVAR and MF-VAR), 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 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 and 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:

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

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 and 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).

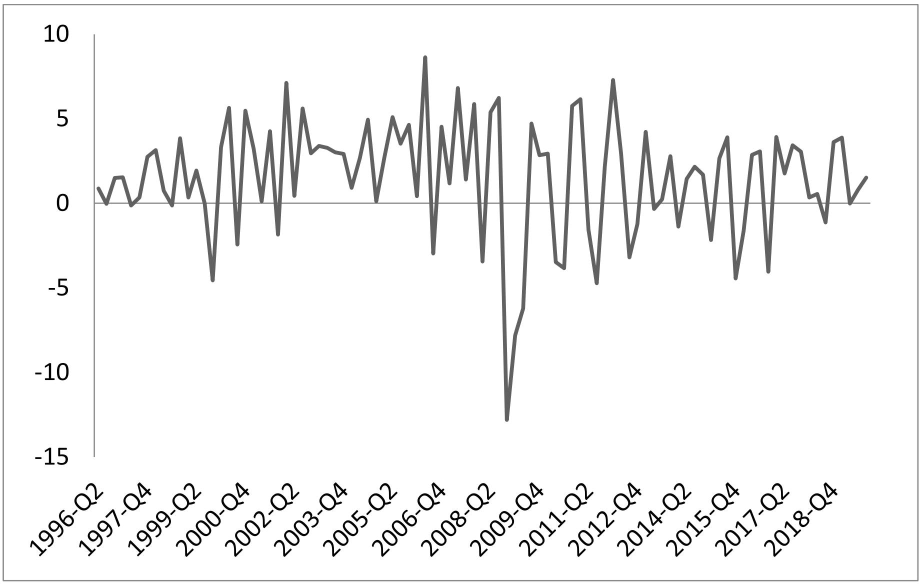
4. Description of the Dataset

For estimating small-scale benchmark models, namely AR, VAR and BVAR we use four key macroeconomic variables, particularly real GDP growth, CPI inflation, short-term nominal interest rates and unemployment rates. We closely follow the approach by Pirshel 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 following an acceleration in tourism output and domestic trade. Industry output has also expanded strongly, driven by a rebound in mining production.

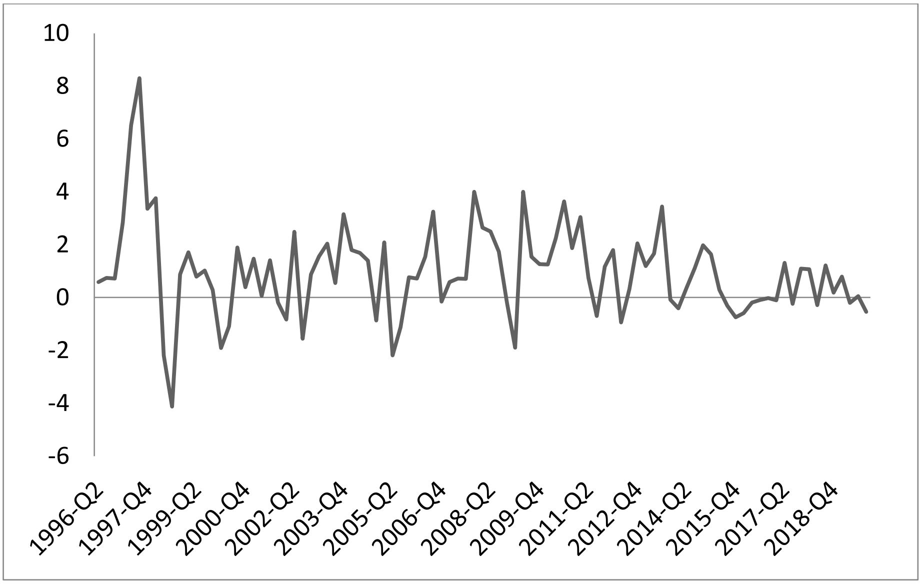
Figure 1 Real GDP Growth (%-th Change to the Previous Quarter)



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.

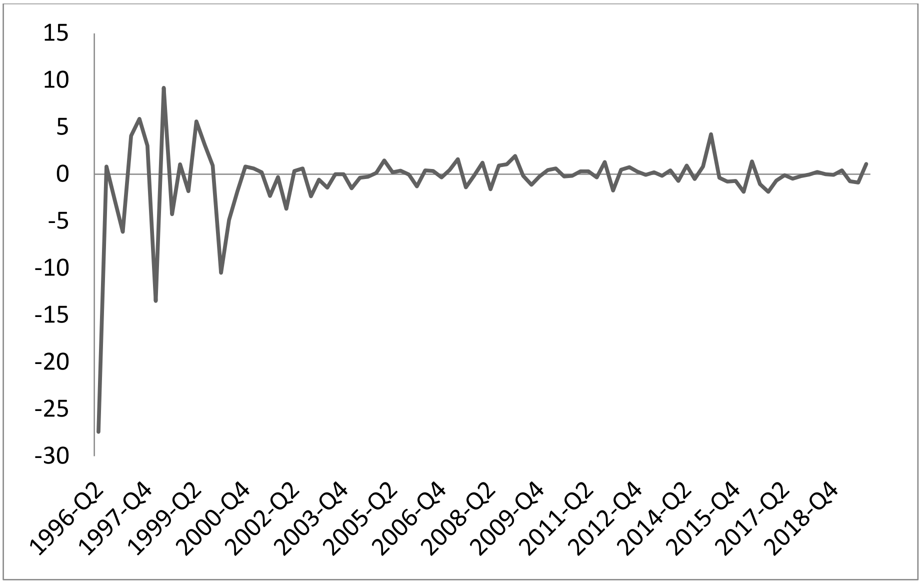
From Figure 2 we can see that CPI inflation is more persistent than real GDP growth, but it shows many spikes which will be hard to predict. In Armenia the calculations of monthly CPI start from 1993, and it is the only indicator characterizing inflation dynamics. The Central Bank of Armenia targeted 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. In this context Kočenda and Varga (2018) provide an analysis of the link between price stability-oriented monetary strategies and inflation persistence. They show that explicit inflation targeting has a stronger effect on taming inflation persistence than implicit inflation targeting and is effective even during and after the financial crisis. They also show that once a country hits the zero lower bound its inflation persistence mildly decreases and that there exists a mild pull to return to inflation persistence mean once a central bank moves away from its inflation target.

Figure 2 CPI Inflation (%-th Change to the Previous Quarter)



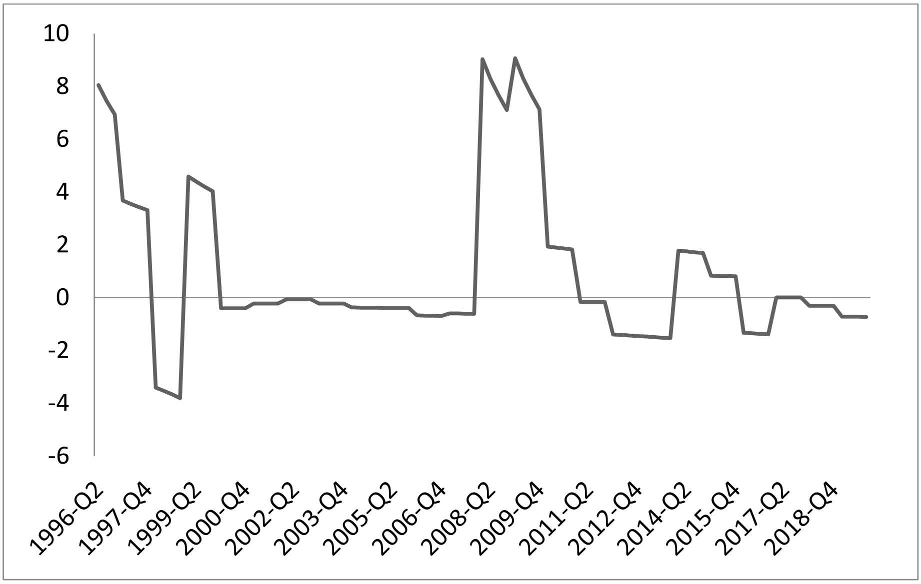
The third macroeconomic variable is the short-term nominal interest rate for time deposits in national currency. This time series is much more persistent than real GDP growth or CPI inflation. The preliminary treatment for this variable includes only first differences (in percentage points). The short-term 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.

Figure 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 the 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 Faibes and Lisman 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).

Figure 4 Total Unemployment Rate (%-th Change to the Previous Quarter)



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

Table 1 Results of Unit Root Tests for Transformed Values of the Key Macroeconomic Variables

	ADF			*PP*		
	Without C,T	***C***	***T***	***Without C,T***	***C***	***T***
Real GDP growth	-8.294 (0) (0.00)	-9.597 (0) (0.00)	-9.631 (0) (0.00)	-8.309 (0) (0.00)	-9.627 (0) (0.00)	-9.675 (0) (0.00)
Inflation	-4.323 (1) (0.00)	-5.395 (1) (0.00)	-5.549 (1) (0.00)	-5.325 (1) (0.00)	-6.432 (1) (0.00)	-6.559 (1) (0.00)
Interest rate	-8.129 (3) (0.00)	-8.169 (3) (0.00)	-8.175 (3) (0.00)	-17.770 (3) (0.00)	-17.853 (3) (0.00)	-17.897 (3) (0.00)
Unemployment rate	-3.628 (1) (0.00)	-3.633 (1) (0.01)	-3.573 (1) (0.03)	-3.588 (1) (0.00)	-3.586 (1) (0.01)	-3.526 (1) (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. C – constant, T – trend. ADF – augmented Dikey-Fuller test, PP – Philips-Perron test.

The results of the unit root tests in Table 1 confirm that usually 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 an 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 indices. Most of the time series are obtained from CBA internal databases. Some of the other time series were obtained from different sources, particularly from https://www.indexmundi.com/, https://www.oecd.org/ and https://www.worldbank.org/. A detailed data description is provided in Appendix Table A1. The vector of time series presented in Appendix Table A1 was preliminarily treated. First, the time series were corrected for outliers and then seasonally adjusted as explained in Appendix Table A1. 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 Table A1 we also see that time series included in the national accounts group are relatively volatile, compared to other time series. 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 rate has little persistence (because 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 with using MATLAB (r2018a) code. The MATLAB code for extracting unobservable components in time and frequency domain have been obtained from the internet sources[2]. 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[3]. 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 store 1- to 4-steps-ahead

[2] The corresponding MATLAB code for factor model proposed by Doz, Gianonne & Reichlin (2001, 2012) can be found here https://www.newyorkfed.org/research/economists/giannone/pub, MATLAB code for factor model proposed by Forni, Hallin, Lippi & Reichlin (2005) can be found here http://www.barigozzi.eu/Codes.html

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

forecasts results. Then we increase the sample size by one (68 observations, 1996Q2-2013Q1) and again generate 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 forecasts for 3-steps-ahead and finally 25 forecasts for 4-steps-ahead.

It is known that rolling forecasts are better able to account for structural breaks and therefore rolling forecasts are preferable to recursive forecasts (Eickmeier, Zigler, 2008). Hence, in our paper we also employ rolling simulation scheme. For the rolling forecast scheme, the initial sample is the same as in the 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 quarter at each step, in the rolling scheme the sample size remains constant. The rolling regression scheme proceed 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 of 1 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 again generate 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.

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 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 and 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 19 competing models. To keep robustness of our analysis we have estimated models with different lags length and different combinations of dynamic and static factors. Following (Pirshel and 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. In particular, 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. We now explain in more detail how we determine the number of factors for each model.

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 out-of-sample forecast evaluations for four models with $p = 1,...,4$. We choose 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 previous papers (Gupta, Kabundi 2011; Kocenda, Poghosyan, 2020), we set the overall tightness 0.1 to 0.3 with increments 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. In section 3 we mentioned that the Litterman/Minnesota prior performs well in forecasts of macroeconomic time series and is often used in practical applications. To keep robustness of our analysis in this paper, besides of the Litterman /Minnesota type prior, we also employ other frequently cited type of priors, such as normal-flat, independent normal-Wishart and Sims-Zha priors (Koop and Korobilis 2010; Sims and Zha 1998). The results of BVAR model with these types of priors are presented in the Tables 3 and 4.

In this paper we also employ mixed-frequency VAR approach to account for high frequency data. As mentioned above, our small-scale VAR model includes four traditional variables. Two of them, particularly inflation and short-term interest rate, have monthly observations. We hypothesize that including high frequency data in the model would improve forecast accuracy. Thus, to keep comparability with previous small-scale models, in this model we continue to keep the key four variables. But in contrast to previous small-scale models, here we take inflation and short-term interest rate in monthly terms and GDP growth and unemployment in quarterly terms. Then we produce forecasts for both recursive and rolling regression and the results are stored in Tables 3 and 4[4].

For FAAR_SW model, in contrast to the 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 the additional set of

[4] Additionally, as a high frequency variable we have selected industrial production index. Then we have conducted bivariate MF-VAR estimation and forecasting for a model that includes GDP growth (quarterly) and industrial production index (monthly). In the results we have concluded that factor augmented models are outperforming bivariate MF-VAR forecasts. The results of calculated RMSFE's can be provided by request.

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[5]. Using this simple rule, we have extracted 12 static factors. Table 2 presents the total variance explained by the extracted 12 static factors. We also present the contributions of each variable group in total variance of each factor separately.

Table 2 Total Variance Explained by Factors and Variable Groups

	Fac 1	*Fac 2*	*Fac 3*	*Fac 4*	*Fac 5*	*Fac 6*	*Fac 7*	*Fac 8*	*Fac 9*	*Fac 10*	*Fac 11*	*Fac 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
%-th 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 %-th 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 the 42 initial variables. The majority of the “Fac 1” total variance is explained by the “International variables” group (13.22%). The second factor explains 12.61% of total variance of the 42 initial variables, and the greatest contributor in the factor variance was the “Labor force and unemployment” group. In such a manner we can observe the contributions of all factors and variable groups. Also, based on the values of contributions we can give a name to each 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. We can also formulate the name of each factor based on a specific variable within the group contributing the most to the variance of said factor. For example, in the fourth factor the largest contributor is the “National accounts” group, but the same group has the largest contribution for “Fac 6”, “Fac 7”, “Fac 8”and “Fac 10”. Thus, in a 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 “Fac 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 combinations of lags ($p = 1,...,4$) 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 and FAAR_QML we vary the

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

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 that the number of dynamic factors cannot exceed 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 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 are 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 the 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 the number of lags, the number of static factors and the number of 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 datasets always outperform small-scale benchmark models for 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 dataset models outperform small-scale benchmark models.

Table 3 RMSFE Indices for the Real GDP Growth (Recursive Regression Scheme)[6]

Model	*Forecast horizon*			
	h = 1	***h = 2***	***h = 3***	***h = 4***
AR (p = 2)[7]	2.379	2.367	2.435	2.473
VAR (p = 1)	2.443	2.502	2.489	2.512
BVAR\Litterman (p = 1, w = 0.3, d = 1)[8]	2.473	2.479	2.479	2.505
BVAR\Normal-flat (p=1, w= 0.3, c2 = 0.1, c3 = 5)	2.427	2.510	2.477	2.473
BVAR\Indep. Normal-Wishart (p=1, c1 = 0.1, c2 = 0.1, c3 = 5)[9]	2.461	2.500	2.466	2.465
BVAR\Sims-Zha (p=1, w = 0.3, d = 1, c3 = 5)	2.509	2.504	2.468	2.469
Mixed-Frequency VAR (p = 1)	2.421	2.504	2.481	2.480
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, r = 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

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

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

[8] w and d 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 the w and d for which the model has the smallest RMSFE.

[9] c1, c2 and c3 are hyperpharameters for normal-flat, independent normal-Wishart and Sims-Zha priors. More concretely, c1 is S scale, c2 is V scale and c3 is degree of freedom. S and V scale parameters we need for the random generating covariance matrix from the inverse Wishart distribution.

[10] r is the number of static factors; q is the number of dynamic factors. In 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 a 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 an 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 used only in one iteration and parameters are not updated.

Table 4 RMSFE Indices for the Real GDP Growth (Rolling Regression Scheme)

Model	*Forecast horizon*			
	1	*2*	*3*	*4*
AR (p = 2)	2.383	2.376	2.435	2.471
VAR (p = 1)	2.482	2.550	2.517	2.537
BVAR \Litterman (p = 2, w = 0.1, d = 1)	2.549	2.447	2.458	2.494
BVAR\Normal-flat (p=1, w= 0.3, c2 = 0.1, c3 = 5)	2.470	2.515	2.503	2.453
BVAR\Indep. Normal-Wishart (p=1, c1 = 0.1, c2 = 0.1, c3 = 5)	2.459	2.501	2.490	2.443
BVAR\Sims-Zha (p=1, w = 0.3, d = 1, c3 = 5)	2.527	2.491	2.490	2.453
Mixed-Frequency VAR (p = 1)	2.500	2.706	2.508	2.465
FAAR_SW (p = 2, r = 2)	2.243	2.468	2.504	2.486
FAAR_FHLR (p = 2, q = 1, r = 2)	***2.223***	2.508	2.492	2.443
FAAR_2S (p = 2, q = 2, r = 2)	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

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 answer 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, 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 ($H_0 : l_t = 0$). The results of t-statistics obtained from regressing the loss differentials on the intercept both for recursive and rolling regressions are presented in Tables 5 and 6.

Table 5 Diebold-Mariano Statistics (Recursive Regression Scheme)

	Forecast horizon			
	h = 1	***h = 2***	***h = 3***	***h = 4***
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 \ Litterman	0.78	-0.97	1.09	1.55
BFAVAR_FHLR versus BVAR \Litterman	0.55	-1.23	0.79	1.43
BFAVAR_2S versus BVAR \Litterman	0.58	-1.23	1.04	1.50
BFAVAR_QML versus BVAR \Litterman	0.51	-1.37	0.94	1.59

Table 6 Diebold-Mariano Statistics (Rolling Regression Scheme)

	Forecast horizon			
	h = 1	***h = 2***	***h = 3***	***h = 4***
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 \ Litterman	0.71	-1.00	0.92	1.08
BFAVAR_FHLR versus BVAR \ Litterman	0.44	-1.50	-0.67	-1.22
BFAVAR_2S versus BVAR \ Litterman	0.53	-1.27	0.89	1.09
BFAVAR_QML versus BVAR \ Litterman	0.39	-1.56	0.79	1.16

Notes: * indicate 10% level of significance.

The statistics presented in Tables 5 and 6 indicate whether the performance results of large-scale and small-scale benchmark models are significantly different. From Table 5 and Table 6 we see that when we compare the predictive accuracy of the large-scale models with that of the small-scale models, the differences are not statistically significant for either recursive or rolling regressions. 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 dataset. Based on Tables 3 and 4 we are also able to compare different factor extraction methods to each other. From Tables 3 and 4 we see that factors extracted by the time domain approach (Doz et al., 2011, 2012) are more appropriate for GDP growth forecasting.

7. Conclusion

We analyze the forecast performances of the 19 competing short-term forecasting models. In our analysis we generate ex-post out-of-sample forecasts based on the actual quarterly 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 for 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. This finding is consistent with the finding for the advanced and emerging economies. When we compare the factor extraction methods, particularly SW, FHLR, 2S and QML, we conclude that factors extracted by 2S and QML methods are better. Hence for forecasting GDP growth rate the methods proposed by Doz et al. (2011, 2012) seem more appropriate. The main peculiar feature of this finding is that all previous findings were based on developed countries' datasets, while our finding is based on a developing country's dataset. To check whether the differences in forecasts obtained by the different models are statistically significant we apply the 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.

APPENDIX

Table A1 Descriptive Statistics for GDP Growth Rate (Percentage Change)

Countries	*Country classification*	*Average growth rate, per year, (1996 - 2019)*	*Median*	*Min.*	*Max.*	*Standard deviation*	*Coefficient of variation*
1	*2*	*3*	*4*	*5*	*6*	*7*	*8*
Armenia	lower income	6.13	6.40	-14.10	14.00	5.97	5.62
Austria	high-income	1.83	2.04	-3.76	3.73	1.58	1.55
Azerbaijan	lower income	7.91	6.62	-3.06	34.47	9.24	8.56
Belarus	lower income	4.78	4.89	-3.83	11.45	4.26	4.07
Belgium	high-income	1.85	1.73	-2.02	3.79	1.32	1.30
Brazil	lower income	2.23	2.06	-3.55	7.53	2.65	2.59
Bulgaria	lower income	2.37	3.82	-14.19	7.15	4.98	4.87
Canada	high-income	2.38	2.83	-3.85	6.87	2.31	2.26
Chile	high-income	3.87	4.07	-1.56	7.43	2.40	2.31
China	lower income	8.95	8.81	6.11	14.23	1.97	1.80
Czech Republic	high-income	2.59	2.61	-4.66	6.77	2.64	2.58
Finland	high-income	2.17	2.69	-8.07	6.33	3.01	2.94
France	high-income	1.62	1.73	-2.87	3.92	1.38	1.35
Georgia	lower income	5.59	4.93	-3.65	12.58	3.69	3.50
Germany	high-income	1.38	1.59	-5.69	4.18	1.97	1.94
Hungary	high-income	2.60	3.86	-6.70	5.41	2.66	2.60
India	lower income	6.49	7.23	3.09	8.85	1.77	1.66
Indonesia	lower income	4.31	5.03	-13.13	7.82	3.93	3.77
Ireland	high-income	5.57	5.62	-5.08	25.16	5.90	5.59
Italy	high-income	0.59	1.11	-5.28	3.79	1.85	1.84
Japan	high-income	0.85	1.15	-5.42	4.19	1.84	1.83
Kazakhstan	lower income	5.49	4.65	-1.90	13.50	3.99	3.78
Kyrgyz Republic	lower income	4.62	4.42	-0.47	10.92	3.13	2.99
Luxembourg	high-income	3.44	3.63	-4.36	8.48	2.97	2.87
Moldova	lower income	3.06	4.55	-6.54	9.04	4.67	4.53
Netherlands	high-income	2.00	2.11	-3.67	5.03	1.96	1.93
New Zealand	high-income	2.84	3.09	-1.03	5.45	1.46	1.42
Pakistan	lower income	4.03	4.33	0.99	7.55	1.78	1.71
Paraguay	lower income	2.95	3.71	-2.31	11.14	3.27	3.18
Poland	high-income	4.09	4.39	1.13	7.06	1.61	1.55
Portugal	high-income	1.41	1.79	-4.06	4.81	2.28	2.25
Romania	high-income	3.15	3.84	-5.52	10.43	4.17	4.04
Russian Federation	high-income	2.94	4.16	-7.80	10.00	4.47	4.34
Slovak Republic	high-income	3.80	3.92	-5.46	10.83	3.20	3.08
Slovenia	high-income	2.71	3.25	-7.55	6.98	3.01	2.93
Spain	high-income	2.15	2.98	-3.76	5.25	2.37	2.32
Ukraine	lower income	1.33	2.73	-14.76	12.11	6.60	6.51
United Kingdom	high-income	2.09	2.36	-4.25	3.85	1.63	1.60
United States	high-income	2.46	2.54	-2.54	4.75	1.59	1.55
Uzbekistan	lower income	6.09	6.52	1.70	9.47	1.94	1.83

Notes: 40 countries are classified as either high- and low-income countries based on the GNI per capita value. Average growth rate of real GDP is calculated using the geometric mean formula. The final column (column 8) presents the values for the coefficient of variation in terms of %. This coefficient is calculated dividing by standard deviation (column 7) on the average growth rate (column 3 value plus 100) multiplying by 100. We use this coefficient to compare the volatility of GDP growth between low- and high-income countries.

Table A2 Dataset Description (Quarter to the Previous Quarter, in Percentage Change)

Series description	*Source*	*SA*	*Transf.*	*Mean*	*Std. Dev*	*ρ(1)*	*ADF*	*Corr. with real GDP growth*
National accounts								
Value added in industry	CBA	Yes	Ln and Δ	1.15	4.77	-0.14	-11.01	0.437
Value added in agriculture	CBA	Yes	Ln and Δ	0.84	7.11	-0.16	-11.27	0.479
Value added in construction	CBA	Yes	Ln and Δ	1.56	11.92	-0.06	-10.46	0.646
Value added in services	CBA	Yes	Ln and Δ	1.84	3.46	-0.29	-12.99	0.359
Final consumption	CBA	Yes	Ln and Δ	1.26	3.52	-0.14	-10.99	0.443
Private consumption	CBA	Yes	Ln and Δ	1.27	3.75	-0.13	-10.94	0.425
Government consumption	CBA	Yes	Ln and Δ	1.26	9.42	-0.35	-13.71	0.071
Gross fixed capital formation	CBA	Yes	Ln and Δ	1.35	12.24	-0.20	-11.76	0.561
Exports of goods and services	CBA	Yes	Ln and Δ	1.74	10.90	-0.27	-12.63	0.164
Imports of goods and services	CBA	Yes	Ln and Δ	0.75	7.06	0.03	-9.27	0.407
Consumer and producer prices								
CPI-food prices	CBA	Yes	Ln and Δ	0.90	2.39	0.25	-7.44	0.074
CPI-nonfood	CBA	No	Ln and Δ	0.65	1.46	0.13	-8.44	0.029
CPI-services	CBA	No	Ln and Δ	1.54	2.90	0.49	-6.32	-0.133
Ind. production price index	CBA	No	Ln and Δ	1.50	3.68	0.26	-7.48	0.342
Construction price index	CBA	No	Ln and Δ	1.09	3.15	0.25	-7.66	0.304
Tariffs for transportation	CBA	No	Ln and Δ	3.08	10.17	-0.11	-10.73	0.018
Labor force and unemployment								
Labor force	World Bank	No	Ln and Δ	0.03	0.31	0.90	-4.32	-0.027
Employment	World Bank	No	Ln and Δ	-0.08	0.55	0.80	-4.28	0.203
Employment in industry	World Bank	No	Ln and Δ	-0.04	1.14	0.78	-4.33	0.303
Employment in agriculture	World Bank	No	Ln and Δ	-0.53	0.61	0.81	-4.28	-0.07
Employment in services	World Bank	No	Ln and Δ	0.27	0.70	0.80	-4.27	0.219
Self-employed, total	World Bank	No	Ln and Δ	-0.31	0.65	0.83	-3.98	0.074
Employment vulnerable	World Bank	No	Ln and Δ	-0.33	0.64	0.84	-3.85	0.069
Monetary and financial variables								
Monetary base	CBA	Yes	Ln and Δ	3.93	5.68	0.04	-9.41	0.285
Cash money in circulation	CBA	Yes	Ln and Δ	3.19	6.00	0.31	-7.26	0.416
Broad money	CBA	Yes	Ln and Δ	4.54	3.78	0.31	-7.53	0.462
Total deposits	CBA	Yes	Ln and Δ	5.27	4.86	0.16	-9.07	0.243
Firms time deposits	CBA	Yes	Ln and Δ	4.72	20.40	0.00	-9.73	-0.03
Households time deposits	CBA	Yes	Ln and Δ	6.07	6.02	0.38	-7.06	0.134
Total time deposits	CBA	Yes	Ln and Δ	5.96	7.39	0.27	-7.74	0.075
Total loans	CBA	Yes	Ln and Δ	4.78	5.81	0.41	-6.99	0.237
Interest rates for loans, pp	CBA	No	first diff.	-0.83	4.86	-0.32	-15.20	-0.056

Table A2 Dataset Description (Quarter to the Previous Quarter, in Percentage Change) Continued

Series description	*Source*	*SA*	*Transf.*	*Mean*	*Std. Dev*	*ρ(1)*	*ADF*	*Corr. with real GDP growth*
International variables								
EA GDP growth rate	OECD	Yes	Ln and Δ	0.39	0.58	0.59	-4.92	0.322
EA Industrial production	OECD	Yes	Ln and Δ	0.25	1.60	0.54	-5.22	0.376
Russia Industrial production	OECD	Yes	Ln and Δ	0.69	2.14	0.33	-6.74	0.346
Russian Natural Gas	OECD	No	Ln and Δ	1.38	12.17	0.44	-5.56	0.142
Crude Oil; Dated Brent	IndexMundi	No	Ln and Δ	2.38	14.38	0.23	-7.55	0.289
Crude Oil; Dubai Fateh	IndexMundi	No	Ln and Δ	2.52	14.52	0.18	-7.96	0.305
Crude Oil; West Texas	IndexMundi	No	Ln and Δ	2.22	14.36	0.18	-8.03	0.296
Fuel Price Index	IndexMundi	No	Ln and Δ	1.90	12.66	0.28	-7.17	0.338
Non-Fuel Price Index	IndexMundi	No	Ln and Δ	0.51	5.71	0.41	-6.18	0.421
All Commodity Price Index	IndexMundi	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), ρ(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. The last column presents the correlation coefficients between real GDP growth rates and additional variable.

REFERENCES

Abdić A, Resić E, Abdić A (2020): Modelling and Forecasting GDP Using Factor Model: An Empirical Study from Bosnia and Herzegovina. *Croatian Review of Economic, Business and Social Statistics,* 6(1):10-26.

Alessi L, Barigozzi M, Capasso M (2010): Improved Penalization for Determining the Number of Factors in Approximate Factor Models. *Statistics & 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, Camba-Méndez G, Giannone D, Rünstler G, Reichlin L (2011): Short-Term Forecasts of Euro Area GDP Growth. *Econometrics Journal,* 14(1):25-44.

Arratibel O, Kamps C, Leiner-Killinger N (2009): Inflation Forecasting in the New EU Member States. *ECB Working Paper* No. 1015, European Central Bank.

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.

Baltagi B, Kao C, Wang F (2017): Identification and Estimation of a Large Factor Model with Structural Instability. *Journal of Econometrics,* 197(1):87–100.

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, Weber A (2010): The Transmission Mechanism in Armenia: New Evidence from a Regime Switching VAR Analysis. *IMF Working Paper* No. 10/270.

Brisson M, Campbell B, Galbraith J (2003): Forecasting Some Low-Predictability Time Series Using Diffusion Indices. *Journal of Forecasting,* 22:515-531.

Camacho M, Sancho I (2003): Spanish Diffusion Indexes. *Spanish Economic Review,* 5(3):173-203.

Camacho M, Dal Bianco M, Martinez-Martin J (2015): Short-Run Forecasting of Argentine Gross Domestic Product Growth. *Emerging Markets Finance and Trade,* 51(3):473-485.

Corona F, González-Farías G, Orraca P (2017): A Dynamic Factor Model for the Mexican Economy: Are Common Trends Useful When Predicting Economic Activity? *Latin American Economic Review,* 26(1):1-35.

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

Diebold F, Mariano R (1995): Comparing Predictive Accuracy. *Journal of Business & 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 & Economic Statistics,* 33(1):1-1.

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.

Eickmeier S, Ziegler Z (2008): How Successful Are Dynamic Factor Models at Forecasting Output and Inflation? A Meta-Analytic Approach. *Journal of Forecasting,* 27(3):237-265.

Fiorentini G, Galesi A, Sentana E (2018): A Spectral EM Algorithm for Dynamic Factor Models. *Journal of Econometrics,* 205(1):249–279.

Ghysels E (2016): Macroeconomics and the Reality of Mixed Frequency Data. *Journal of Econometrics,* 193:294–314.

Gunay G (2018): Forecasting Industrial Production and Inflation in Turkey with Factor Models. *Central Bank Review,* 18(4):149-161.

Gupta R, Kabundi A (2011): A Large Factor Model for Forecasting Macroeconomic Variables in South Africa. *International Journal of Forecasting,* 27(4):1076-1088.

Jansen W. Jos, Xiaowen J, de Winter, Jasper M (2016): Forecasting and Nowcasting Real GDP: Comparing Statistical Models and Subjective Forecasts. *International Journal of Forecasting*, 32(2):411-436.

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.

Forni M, Hallin M, Lippi M, Zaffaroni P (2017): Dynamic Factor Models with Infinite-Dimensional Factor Space: Asymptotic Analysis. *Journal of Econometrics,* 199(1):74–92.

Forni M, Giovannelli A, Lippi M, Soccorsi S (2018): Dynamic Factor Model with Infinite-Dimensional Factor Space: Forecasting. *Journal of Applied Econometrics,* 33(5):625–642.

Ghazaryan N (2015): Short Term Forecasting System of Private Demand Components in Armenia. *Armenian Journal of Economics*, 1:21-52.

Hamilton J (1994): *Time Series Analysis*. Princeton University Press.

Hofmann B (2009): Do Monetary Indicators Lead Euro Area Inflation? *Journal of International Money and Finance,* 28(7):1165-1181.

Kilian,L, Lutkepohl H (2017): *Structural Vector Autoregressive Analysis*. Cambridge Books, Cambridge University Press, number 9781316647332, December.

Kočenda E, Poghosyan K (2020): Nowcasting Real GDP Growth: Comparison between Old and New EU Countries. *Eastern European Economics,* 58(3):197-220.

Kočenda E, Varga B (2018): The Impact of Monetary Strategies on Inflation Persistence. *International Journal of Central Banking,* 14(4):229-274.

Koop G, Korobilis D (2010): *Bayesian Multivariate Time Series Methods for Empirical Macroeconomics.* Foundations and Trends in Econometrics.

Li H, Li Q, Shi Y (2017): Determining the Number of Factors When the Number of Factors Can Increase With Sample Size. *Journal of Econometrics,* 197(1):76–86.

Litterman R (1980): A Bayesian Procedure for Forecasting with Vector Autoregressions. Manuscript, Massachusetts Institute of Technology, Department of Economics.

Lopez-Buenache G (2018): Forecast Accuracy of Small and Large Scale Dynamic Factor Models in Developing Economies. *Review of Development Economics,* 22(3):63-78.

Ma S, Su L (2018): Estimation of Large Dimensional Factor Models with an Unknown Number of Breaks. *Journal of Econometrics,* 207(1):1–29.

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.

Poghosyan T, Kočenda E, Zemčik P (2008): Modeling Foreign Exchange Risk Premium in Armenia. *Emerging Markets Finance and Trade,* 44(1):41-61.

Porshakov A, Ponomarenko A, Sinyakov A (2016): Nowcasting and Short-Term Forecasting of Russian GDP with a Dynamic Factor Model. *Journal of the New Economic Association,* 30(2):60-76.

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.

Sims C, Zha T (1998): Bayesian Methods for Dynamic Multivariate Models. *International Economic Review*, 39:949–968.

Stock J, Watson M (2002): Macroeconomic Forecasting Using Diffusion Indexes. Journal of Business & Economic Statistics 20(2):147-162.

Su L, Wang X (2017): On Time-Varying Factor Models: Estimation and Testing. *Journal of Econometrics*, 198(1):84–101.