Bank of Russia

Nowcasting Russian GDP using forecast combination approach

WORKING PAPER SERIES

No. 68 / February 2021

M. Zhemkov

Michael Zhemkov

Bank of Russia, Monetary policy Department

Email: zhemkovmi@mail.cbr.ru

Bank of Russia Working Paper Series is anonymously refereed by members of the Bank of Russia Research Advisory Board and external reviewers.

Author expresses deep gratitude to the specialists of the Bank of Russia's Monetary Policy Department and Research and Forecasting Department and to the anonymous reviewer for their valuable advice and assistance with this paper.

Cover image: Shutterstock.com

Address: 12 Neglinnaya street, Moscow, 107016
Tel.: +7 495 771-91-00, +7 495 621-64-65 (fax)
Website: www.cbr.ru

Abstract

This paper presents a forecast combination approach for short-term assessment of economic growth in Russia. Our method significantly expands the existing domestic academic literature and embraces most advanced nowcasting techniques. The key feature of our approach is forecasting the growth rates of components GDP by expenditure using scenario indicators and a huge set of high-frequency predictors. The total number of combined models is nearly 500. Comparing our approach with popular benchmark models over the period from January 2003 to March 2020, we conclude that our method has higher accuracy of short-term GDP forecasting. The findings of this research may be useful for the development of monetary and communication policy in Russia.

Key words: nowcast, forecast combination, GDP, inflation targeting, DFM, Bank of Russia

JEL: E27, E52, E58

1. Introduction

In 2015, the Bank of Russia switched to an inflation targeting, which based on a medium-term forecast of key macroeconomic indicators. Medium-term forecasting in most central banks largely depends on the estimates of the current macroeconomic situation. The reason for this is that large structural and semi-structural models are not always capable to accurately capture economic dynamics over a short-term forecast horizon, because they are rather aimed at medium-term forecasting. That is why finding an efficient technique for short-term forecasting is the most relevant task for many central banks, including the Bank of Russia.

Since statistics of most macroeconomic indicators have publication lags and a relatively low frequency, economists have to estimate not only current and future economic activity, but also the recent past. In academic literature, this process is called nowcasting. A major challenge in nowcasting is the selection of the better model due to the huge range of various forecast techniques. These models may differ in specifications, assumptions, and data used, while it is hard to rank most of them by forecasting quality because in practice they sometimes produce significantly varying results for different countries and samples. Since macroeconomists need to select the 'only true' model, the focus of studies has gradually shifted in recent years towards combining multiple forecasts.

Forecast combination is a relatively simple technique involving the averaging of independent forecasts produced by various models based on their historical accuracy. The most accurate models thus have higher weights in the aggregate forecast, compared to models with inaccurate forecasts. An apparent advantage of this approach is that it relies on a broad dataset, without compromising the quality of forecasting due to potential problems with the curse of dimensionality. This technique involves combining various indicators instead of incorporating them all at once in a model. As a result, macroeconomists may use the entire data pool available in their forecasting. The second advantage of this method is its adaptability since it implies that the structure of a model is flexible and may be re-evaluated based on incoming data. This makes it possible to automatically adjust weights in the aggregate forecast due to changes in their interconnections caused by structural shocks in the economy. The last but not the least benefit of forecast combination is the possibility to diversify accidental errors in the models.

The aim of this paper is to analyze the accuracy of forecast combination in estimating economic growth in Russia. For this purpose, we suggest a forecast combination technique that encompassing the most advanced nowcasting methods. We implement this technique in several stages. At the first stage, we collect a large array of economic indicators. The novelty of our research is that we combine several groups of predictors: key scenario indicators used in the Bank of Russia's forecasts (oil prices, exchange rate, etc.) and a vast array of monthly statistics, which allows us to make preliminary estimates of the economic situation. At the second stage, we build a dynamic factor model that identifies common unobservable factors from the array of collected indicators, divided into 3 groups – real sector indicators, the financial sector indicators and agent's expectations. Finally, we include these factors and scenario indicators in various model combinations, the forecasts from which are aggregated with weights. The weight of each model variation is calculated based on the forecast accuracy in the past. The total number of the models used in our method is approximately 500. Using out-of-sample forecasts in a pseudo real-time sample over the period from January 2003 to March 2020, we demonstrate the advantages of our forecasting technique compared to the standard set of benchmark models. As far as we know, forecast combination based on mixed-frequency factor models, with account of the ragged edge problem, has not been described in the Russian academic literature, which makes our research especially relevant.

The rest of the article is organized as follows. In Section 2, we discuss the development of research on short-term forecasting of economic activity and focus on the various forecasting methods, such as bridge equations, factor and mixed-frequency models, forecast combination approaches. In section 3 and 4, we provide the research methodology, specifying the models and data used in the article. In section 5, we report the results of our forecast combination accuracy comparing to popular benchmark models. The last section concludes.

2. Literature review

The use of bridge equations is one of the most popular methods implemented by many central banks for short-term forecasting. The key idea of these equations is to bridge a target indicator with one or several key variables released without significant time lags, with the frequencies of these predictors converted into a single one. According to the majority of studies, models of this type very often produce more accurate forecasts than simple models. For instance, Baffigi, Golinelli, and Parigi (2004) try to identify better models for GDP forecasting in the euro area over the period from 1980 to 2002, concluding that bridge equations generate better results compared to alternative benchmark models. The authors also prove that the use of disaggregate data enhances the accuracy of forecasting. Nonetheless, it is still an open question whether aggregate or disaggregate forecasting is more accurate. Specifically, some authors (Fair and Shiller, (1990)) state that disaggregation improves forecasting, while others (Hubrich, (2005)) believe that disaggregation does not necessarily enhance the quality of forecasting. Hendry and Hubrich (2011) cover this issue in detail. A number of papers rely on bridge equations to forecast both the demand side and the supply side of GDP. Thus, a relatively new research by Pinkwart (2018) estimates the output and expenditure components of GDP using a system of bridge equations. According to the findings, both approaches provide satisfactory results for short-term forecasting, while the forecast combination of demand and supply side components significantly increases the accuracy of GDP estimate.

However, the use of bridge equations involves a number of challenges. Firstly, analysts should have forecasts for all high-frequency exogenous indicators. These forecasts may rely on both simple (autoregressive, naïve, etc.) and complex models, but in any case, this may increase forecast error due to the inaccurate prediction of exogenous indicators. Secondly, as practice shows, it is difficult to justify the choice of several most important predictors for forecasting, what's why later studies increasingly use the 'bridging with factors' approach (bridging with the common factors extracted from a dataset). The usage of these factor models is currently a standard practice in most central banks and international organizations. One of the most well-known and earliest papers is the research by Bernanke, Boivin, and Eliasz (2005) suggesting a factor-augmented vector autoregressive (FAVAR) model to analyze the impact of monetary policy on a broad array of economic indicators. The majority of later papers show that analysts may considerably enhance the quality of forecasting by including factors in models. For instance, Giannone, Reichlin, and Small (2008) suggest a GDP forecast model incorporating factors from a large set of high-frequency predictors (approximately 200 macroeconomic indicators of the US economy for the period from 1982 to 2005). The authors prove that the dynamic factor model has a higher forecasting accuracy than alternative benchmark models and that the use of monthly indicators is critical to enhance the quality of forecasting. Furthermore, the authors identify key macroeconomic predictors for improving forecasting accuracy: these are mostly survey indicators and labor market indicators. These findings confirm the results of an earlier research by Schumacher and Breitung (2008) about using of monthly survey indicators for improving the forecast accuracy. A research by Marcellino and Schumacher (2010) continues the discussion of various ways to assess unobserved factors extracted from a set of high-frequency predictors and ways to employ these factors in forecasts. Using mixed-frequency factor models with the data on GDP growth in Germany over the period from 1992 to 2006, the authors conclude that the autoregressive component is crucial in forecasting, at the same time the method for assessing the polynomial function does not change forecasting quality.

In practice, the use of factor models may involve certain challenges. Firstly, the selection of the optimal number of extracted factors remains a matter of dispute, several article are devoted to this problem (e.g. by Boivin and Ng (2006)). Secondly, there are challenges associated with the delays in official statistics publication, i.e. the ragged edge (RE) problem (Wallis (1986)). For instance, the first available statistics on the past period are financial data and agent's expectations, while data on the real sector are released more than two weeks later. In this case, the panel data is unbalanced and standard methods to extract factors may cut off the most recent available data, which are crucial for high-quality nowcasting. The academic literature proposes various ways to deal with this problem. The most popular and convenient one is a two-step method to estimate factors. Specifically, this approach was described by Giannone, Reichlin, and Small (2008) and Doz et al. (2011).

Another restriction is that models should include indicators with the same frequency. However, this requirement may not always be met in practice. In particular, GDP is usually available at a lower frequency (generally, quarterly) than real economy indicators used for bridging (typically, monthly). For this reason, when higher-frequency predictors are included in models, analysts should aggregate them into a quarterly/annual frequency, assuming a potential loss of information. This aggregation may be avoided by using a special type of mixed-frequency (MF) or mixed data sampling (MIDAS) models. Models of this class have been applied increasingly more often in the academic literature, e.g. by Clements and Galvao (2008), Marcellino and Schumacher (2010), Kuzin, Marcellino, and Schumacher (2011). Thus, Clements and Galvao (2008) discuss whether the mixed-frequency data may help improve the quality of economic growth forecasts. For this purpose, the authors analyze the quality of out-of-sample pseudo real-time forecasts based on US data for 1959–2005, showing that the MIDAS method substantially enhances forecasting quality. A later paper by Kuzin, Marcellino, and Schumacher (2011) compares the predictive power of the MIDAS and MF-VAR models. According to the authors, the MF-VAR and MIDAS approaches do not give any significant differences in the results of forecasting: the MF-VAR are slightly better over longer forecast horizons, while the MIDAS – over shorter ones.

When an analyst uses different models with various sets of indicators, a question may arise: which model is optimal for short-term forecasting in selected period and in a particular country. It is rather difficult to rank most of the models by forecasting quality, because in practice they sometimes yield significantly varying results for different countries and samples. As long as it is hard to choose an optimal model, the focus of academic research in recent years has shifted towards combining multiple models and forecasts. Among these studies, a paper by Timmermann (2006) may be considered as a key one. The author carries out a theoretical analysis of the benefits of forecast combination, such as the aggregation of a broad dataset without compromising the quality of models, flexibility to structural changes, and a higher robustness to the problems of misspecification and forecast errors. Nevertheless, Timmermann points to certain problems of forecast combination associated with additional uncertainty due to the methods used to calculate model weights. Furthermore, if a researcher knows an ‘ideal’ model, forecast combination will be less accurate by default. Finally, the author proposes to include time-varying weights and exclude the worst models from combined approaches to improve estimates.

Currently forecast combination is a well-developed approach in the academic literature extensively applied to the majority of macroeconomic indicators. In addition, one of the conclusions by Kuzin, Marcellino, and Schumacher (2011) is that forecast combination based on the MIDAS and MF-VAR models enhances the accuracy of GDP forecasts. Pinkwart (2018) confirms that combining improves the quality of forecasts, but when the forecasts of the demand- and supply-side components of GDP are combined. Another example is the use of combinations in forecasting inflation, e.g. a study by Koop and Korobilis (2012), and the application of this technique in Russia described by Styrin (2019) and Andreev (2016).

Short-term forecasting of economic growth is a popular research topic in Russia as well. There are currently several studies suggesting the use of factor models to estimate GDP growth in Russia. For instance, Porshakov, Ponomarenko, and Sinyakov (2016) provide the results obtained through using dynamic factor model to forecast GDP growth in Russia. For this purpose, the authors use approximately 116 various indicators broken down into three groups – leading indicators, real economy indicators, and financial variables. Applying the principal component analysis and the Kalman filter, the authors extract unobserved common factors from these groups and predict GDP growth rates. Moreover, at the last stage, they convert high-frequency unobserved factors into low frequency ones to insert them into the bridge equation. Their model demonstrates rather accurate forecasts, as compared with alternative approaches to GDP estimates. Achkasov (2016) develop this approach in an article. In particular, he assesses factors separately for each group of predictors: economic agents' expectations, financial variables, and real economy and external sector indicators and then decomposes quarterly GDP growth into these factors.

The Russian academic community also employs mixed-frequency models. Thus, Mikosch and Solanko (2019) give an example of using mixed-frequency data to estimate and forecast economic growth rates. The main objective of this paper is to select predictors and models that would be most valuable in short-term GDP forecasting. The authors conclude that output indices, leading indicators, and individual financial and banking indicators improve the quality of forecasting. Comparing different types of models by forecasting accuracy, the MIDAS approach outperforms standard bridge equations.

Unfortunately, as far as we know, forecast combination based on mixed-frequency factor models, with account of the RE problem, has not been described in the Russian academic literature, which makes our research especially relevant. Nonetheless, there are papers on using this approach to forecast inflation. For instance, Styrin (2019) evaluates the accuracy of inflation forecasting based on dynamic model averaging (DMA). This approach was first suggested by Raftery et al.

(2010), and used by Koop and Korobilis (2012) to forecast inflation in the USA. DMA implies that the model structure and estimated parameters are time-varying, and an aggregate forecast is made of a combination of the forecasts produced by several models weighted based on their historical accuracy. Although this approach has a number of apparent advantages, the article Styrin (2019) based on Russian statistics for the period from January 2002 to September 2017 does not confirm that DMA consistently outperforms a set of benchmark models. In contrast, Andreev (2016) shows that forecast combination ensures higher forecasting accuracy compared with other benchmark models. We assume that forecast combination results may vary depending on predictors and samples used, which leaves room for further research.

3. Methodology

Our approach embraces most advanced techniques for short-term GDP forecasting described in the academic literature. The GDP forecasting scheme in its general form is presented in Figure 1. The headline GDP forecast is aggregated from the forecasts of every GDP by expenditure component. Since the dynamics of these components may be described through various processes, the model designed solely for aggregate GDP may be misspecified. However, the weighted sum of GDP by expenditure components may not always be the same as headline GDP. This is associated with both the technical specifics of seasonal adjustment (each series is adjusted separately, due to which differences in seasonal factors between the series may cause deviations from the headline GDP) and the statistical discrepancy. It is worth emphasizing that we use in our models only the demand-side components of GDP. This is primarily associated with the availability of time series (the supply-side components of GDP have a smaller sample, because the OKVED system (the Russian Classification of Economic Activities) was changed and certain series are incomparable). However, in the future, as the sample will expand gradually, it may be useful to incorporate the supply-side components of GDP into the forecast as well.

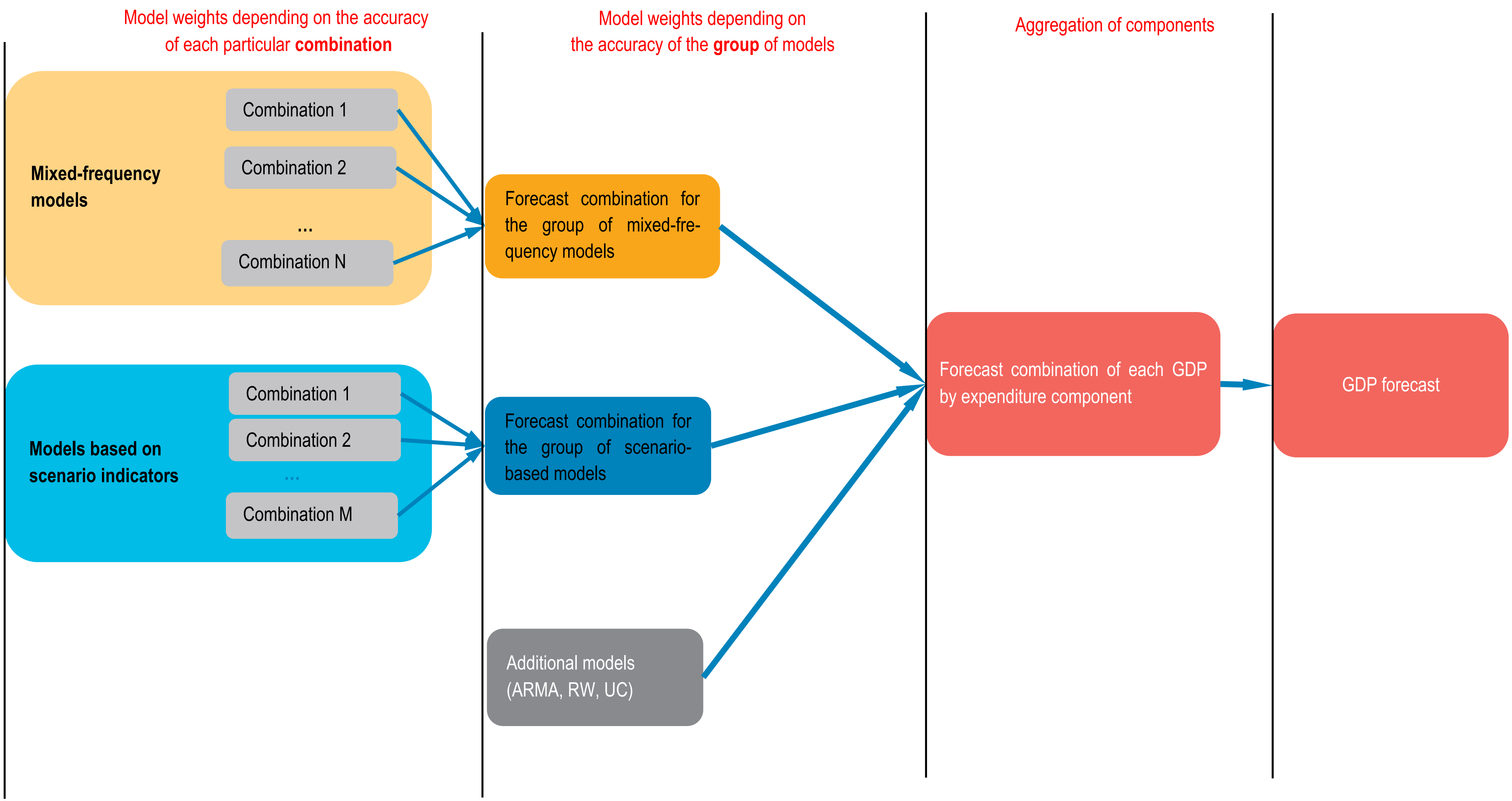

Figure 1 General scheme for forecast combination approach

We get a forecast of every GDP by expenditure components (such as final consumption by households and general government, gross fixed capital formation, changes in inventories and exports) from the forecast combination of three main groups:

1) Models based on scenario indicators
2) Mixed-frequency models
3) Additional models

The forecast combination technique is not applied for gross capital formation and import. Forecast for gross capital formation is aggregated from the forecasts of gross fixed capital formation and changes in inventories, and forecast for import –from final consumption by households and gross capital formation. The indicators of the first group are used to devise a scenario-based forecast enabling the Bank of Russia to make monetary policy decisions, and their usage in our model makes the GDP forecast consistent with the forecasts of other main macroeconomic indicators used by the Bank of Russia (inflation, the balance of payments, monetary conditions, etc.).

We use oil prices, the real exchange rate, the growth rate of the global economy, and short- and long-term loan rates as the main scenario indicators in first group of models. The movements of these predictors are exogenous, and they were selected based on the results of the preliminary studies about analysis of the relevance of variables for modelling GDP expenditure components. The forecasts of these indicators is set exogenously in accordance with the baseline forecast of the Bank of Russia in the course of monetary policy decision. Equation (1) presents a general scheme for forecast combination using scenario indicators.

$$y_t^f = \beta_0 + \sum_{i=1}^{M} \beta_i \ X_{it} + \varepsilon_t \qquad (1)$$

where $\mathrm{y}_\mathrm{t}^\mathrm{f}$ – the forecast of the GDP component growth from each individual combination, X_{it} – a scenario indicator

It should be noted that these scenario indicators do not always adequately capture the current situation and temporary shocks, due to which the accuracy of GDP estimates for the previous and current quarters may be insufficient. One of the ways to deal with this problem is to use the real-time indicators of economic activity in bridge equations. However, the application of such models raises quite a natural question: what high-frequency predictors should be included in a model, and what indicators are to be left out. The academic literature offers two answers to this question:

1) To select only few predictors. The selected predictors should be most informative for GDP forecasting. Moreover, the choice should be perfectly justified, and a researcher should be confident that no additional information would be lost because particular indicators are not included in a model.

2) To break down all available indicators into groups and extract unobserved common factors from them.

The idea of adding common factors in models was partially grounded by the theory that macroeconomic shocks should be economy-wide and affect the majority of economic indicators, due to which a wide range of macroeconomic indicators depend on unobserved common factors and they contain main data on the co-movements of the set of predictors. This type of models has become very popular among macroeconometrists since it allows the modelling of a broad range of indicators without the curse of dimensionality. Furthermore, the majority of studies demonstrate that unobserved common factors more accurately predict changes in economic activity in recent past and present.

Our approach relies on the second solution: our forecast combination technique based on the group of models with unobserved common factors extracted from high-frequency predictors. Common factors may be extracted from a large dataset in various ways, for example, using the standard principal component analysis. This method is rather simple to implement, but it imposes a rather serious restriction on the balance of a sample and use of the principal component technique may cause a partial loss of essential information. In order to avoid this problem, we use a rather popular two-step estimation method introduced by Doz et al. (2011). This approach develops a dynamic factor model using the state-space technique and the Kalman filter to estimate factors in two stages. First, model parameters are estimated using the principal component technique, after which the findings are used to re-evaluate factors through the Kalman filter, thus mitigating the problem of missing data for indicators. The Kalman filter enables not only the estimation of factors, but also their adjustment with account of gradually released statistics, owing to which this algorithm may be extensively exploited for unbalanced samples.

Generally, the model presented in the state-space form may be expressed as follows:

$$X_t = \lambda f_t + \varepsilon_t \qquad (2)$$

$$f_t = \psi f_{t-1} + \eta_t \qquad (3)$$

where X_t – the matrix of high-frequency observations, f_t – the matrix of unobserved common factors.

Nonetheless, there is no apparent preference in the academic literature to either of the techniques that may be employed to extract factors. For instance, Marcellino and Schumacher (2010) state that the method for extracting unobserved common factors does not affect the accuracy of GDP forecasting. According to Bernanke et al. (2005), the two-step method delivers more realistic dynamics of factors, although both approaches generate similar results. Boivin and Ng (2006) discuss this topic in detail. The use of this method in our study has several undeniable advantages: it does not only address the RE problem, but also makes it possible to forecast all unobserved factors over a given forecast horizon within one model.

As was noted above, including the extracted factors in a model may involve certain problems. One of them is differences in the frequencies of predictors. Standard regression models impose the restriction implying the same frequency of explained dependent variables, and this requirement is rarely met in practice. Traditionally, there are two opposite approaches to dealing with this problem: aggregating high-frequency indicators into low frequency ones (the aggregation approach) and using individual coefficients for each high-frequency indicator (the individual coefficients approach). The individual coefficients approach ensures high flexibility, but requires analysts to estimate a vast amount of parameters. The aggregation method estimates a much smaller number of parameters, but researchers may lose additional information from the aggregation. The MIDAS technique offers a balance between these two approaches: on the one hand, high-frequency indicators are estimated with various weights in a model; on the other hand, mixed data sampling makes it possible to estimate a moderate number of unknown parameters (Ghysels, Santa-Clara, and Valkanov, 2004; Ghysels, Santa-Clara and Valkanov, 2006; Andreou, Ghysels, and Kourtellos, 2010).

In its general form, this technique may be presented as follows:

$$y_t = X_t'\beta + f\left(X_{\frac{t}{s}}^H, \theta, \lambda\right) + \varepsilon_t \qquad (4)$$

where y_t – a low-frequency dependent variable over the time period t, X_t – a set of explanatory indicators having the same frequency as the dependent variable, $X_{\frac{t}{s}}^H$ – a set of high-frequency explanatory indicators (S – the number of high-frequency values in one low-frequency value), f – the function describing the effect of high-frequency regressors on a dependent variable (the weighting function), β, θ, λ – estimated parameters.

Of major interest in this class of models is the method employed to estimate the weighting function, which restricts the number of estimated parameters. There are several approaches to estimating this function: step weighting, Almon (PDL) weighting, exponential Almon weighting, and Beta weighting. As shown in a number of studies (Mikosch and Solanko, 2019; Mikosch and Zhang, 2014), the use of non-exponential Almon polynomials improves forecasting accuracy, as compared to exponential or other non-linear lag polynomials.

By aggregating the above-described DFM and MIDAS models using non-exponential Almon polynomials, we will obtain a variation of the FaMIDAS approach (Marcellino and Schumacher, 2010) which is implemented in this study. Unobserved common factors are extracted from three data groups: real sector, financial and survey indicators (for details, refer to the Data section). The model will include the first two unobserved factors from each of the groups, which is quite a widespread practice in the academic literature. For instance, according to Stock and Watson (2005), larger improvement in terms of prediction accuracy is observed when the first (and most important) principal components are incorporated in a model.

The last third group comprises simple models without exogenous indicators (autoregressive, random walk, and unobserved trend models). We suppose that such processes may sometimes give better description of the movements of GDP by expenditure components. The preliminary analysis of the data proves that naive and autoregressive forecasts may have dominating weights for individual GDP components, e.g. for general government final consumption (Figure 2). Moreover, a range of studies demonstrates that models for short-term GDP forecasting very often cannot outperform simpler models, or even naive forecasts in some cases. Thus, Giannone, Reichlin, and Small (2008) show that model-based GDP forecasts enhance prediction accuracy compared with naive forecasts only over very short time horizons, and most often within the current quarter.

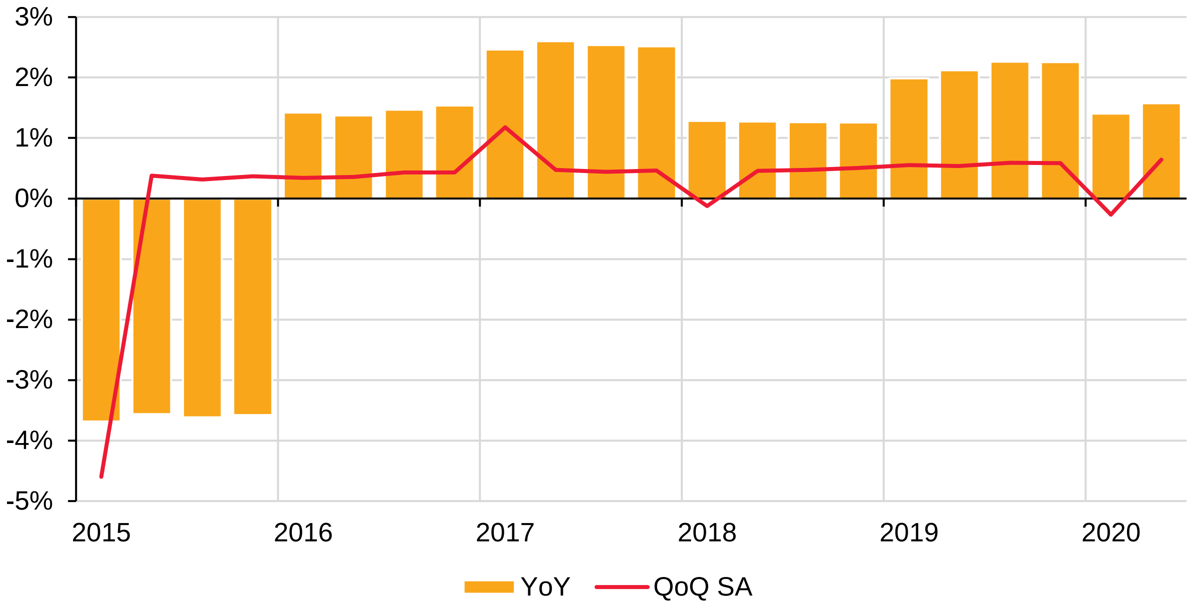

Figure 2 General government final consumption expenditure

In general, the forecast combination using re-evaluated weights both within a group of models and between combinations is expressed as:

$$Y_t^f = \sum_{i=1}^{V} w_i^G \sum_{j=1}^{N} w_j^M \, y_{ij,\,t}^f \qquad (5)$$

where , Y_t^f – the forecast of the growth GDP by expenditure component, w_i^G – the weight of an individual group of models ($w_i^G = \frac{1/MSFE_i^G}{\sum_{i=1}^{V} 1/MSFE_i^G}$, $\sum_{i=1}^{V} w_i^G = 1$), w_i^M – the weight of an individual combination within the group of models ($w_j^M = \frac{1/MSFE_j^M}{\sum_{j=1}^{N} 1/MSFE_j^M}$, $\sum_{j=1}^{N} w_j^M = 1$), $y_{ij,\,t}^f$ – the forecast of the growth rate of a GDP component from an individual combination of each of the groups.

Thus, we use nearly 500 models to predict GDP through forecast combination (Figure 3).

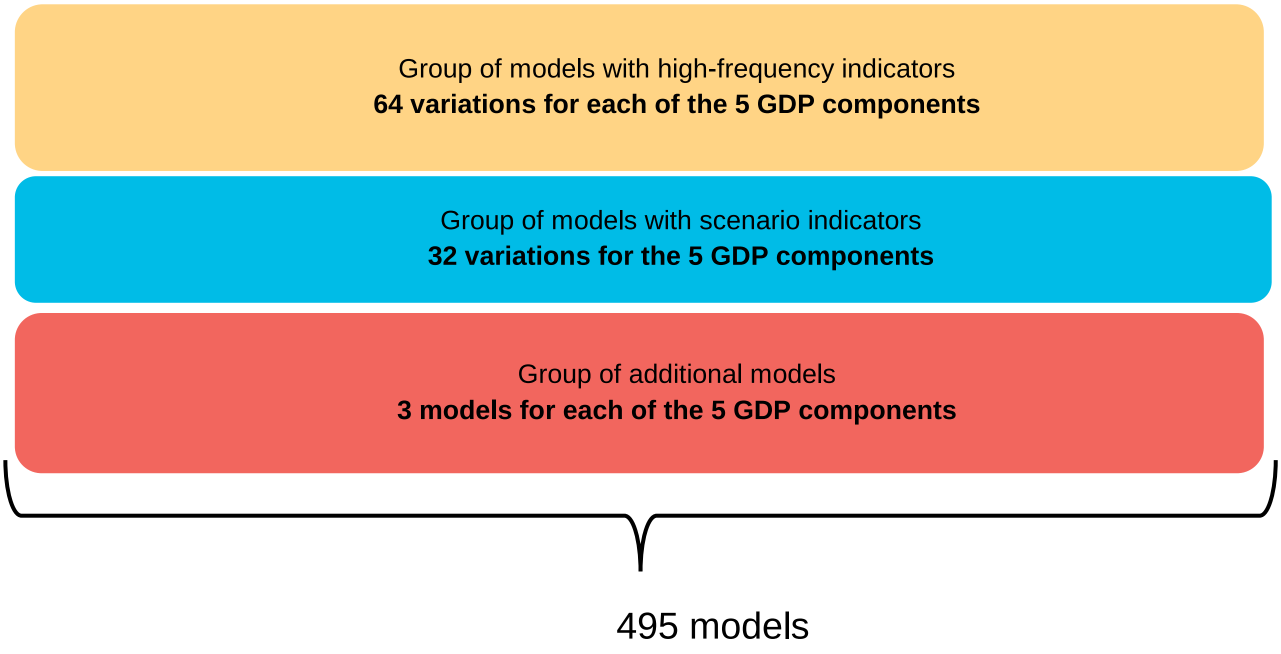

Figure 3 Number of models used in combining GDP forecasts

The academic literature describes several approaches to estimating the weights of models: relative performance, shrinking relative performance, highlighting recent performance, trimming, etc. Nonetheless, it is rather difficult to find a theoretical justification for the selection of a particular method to estimate the weights of models. This study uses the weight estimation approach similar to the one suggested by Kuzin et al. (2011) and Mikosch and Solanko (2019). For every individual expenditure component of GDP, a particular date and forecast horizon, we calculate the mean squared forecast error (MSFE) for a rolling period which is normalised through division by the sum of forecast errors of all possible combinations of models. This approach ensures the high adaptability of our model. If one of the combinations of models demonstrates better forecast accuracy over a particular period, the weight of this model in the aggregate forecast will be higher compared to all others. Conversely, if one of the combinations loses prediction accuracy as new data become available; its weight in the aggregate forecast will considerably decrease.

In order to verify the quality of our technique, we conduct a pseudo-real time estimation of the accuracy of the out-of-sample forecasts of this model and several benchmark models with a rolling estimation window and for various forecast horizons. We select the most widespread benchmark models for the short-term prediction of GDP used in the academic literature: a dynamic factor model (DFM), a factor-augmented vector autoregressive (FAVAR) approach, dynamic model averaging/switching (DMA/DMS) and several standard simple models (ARMA, RW). Among these models, the dynamic model averaging/switching approach should be explained in detail. This approach was first suggested by Raftery et al. (2010), and then applied by Koop and Korobilis (2012) to forecast inflation in the USA. It implies that the structure of a model and estimated parameters are time varying and that a final forecast is compiled from a combination of forecasts based on several models weighted depending on their prediction accuracy over time. Styrin (2019) gives an example of using this technique based on Russian statistics.

In its general form, the process of dynamic model averaging may be presented as follows:

$$y_t = z_t^k \theta_t^k + \varepsilon_t^k \qquad (6)$$

$$\theta_{t+1}^k = \theta_t^k + \eta_t^k \qquad (7)$$

where θ_t^k – a vector of time-varying estimated parameters, and $\varepsilon_t^k \sim N(0, H_t^k)$, $\eta_t^k \sim N(0, Q_t^k)$

DMA estimates $\Pr(L_t = k | y^{t-1})$ for each of k=1,…,K models, where $L_t \in \{1,2,\dots,K\}$ and $y^t \in (y_1, \dots, y_t)'$, and then combines the forecasts based on the estimated probabilities. DMS is similar to DMA, but instead of combining forecasts, it selects the only model with higher probability. Koop and Korobilis (2012) describe the estimation technique in detail.

4. Data

In this study we use the chain indices of GDP by expenditure components in fixed prices and of real-time indicators (as % QoQ SA for quarterly predictors and as % MoM SA for monthly predictors). We seasonally adjust all series through the X-13 ARIMA-SEATS method. The series of GDP by expenditure and scenario indicators are used at a quarterly frequency for the period from 2003 Q2 through 2020 Q1, with the observations in the sample numbering 68. Real-time indicators are specified at a monthly frequency for the period from January 2003 through March 2020, with the number of observations totaling 207. The monthly indicators are broken down into three groups: real sector indicators, financial indicators, and agent's expectations. These groups comprise:

1) High-frequency **real sector** indicators (39 predictors):
   - Industrial output
   - Labour market indicators
   - Business surveys: current situation
2) High-frequency **financial and external** indicators (18 predictors)

- Deposit rates
- Loan rates
- Stock indices
- Monetary aggregates
- External sector indicators

3) High-frequency **expectations** indicators (28 predictors):
   - Companies' monitoring
   - PMI
   - Leading indicators
   - Diffusion indices

A more detailed breakdown of the predictors into the groups is presented in Appendix 1.

5. Results

We compare the accuracy of the pseudo-real-time nowcasting of GDP yielded by the forecast combination technique with similar forecasts generated by the benchmark models. All the models are evaluated on data for the period from 2003 Q2 through 2014 Q4; and their out-of-sample mean squared forecast error (MSFE) for a rolling period is estimated over the period from 2015 Q1 through 2020 Q1.

Figures 4–6 demonstrate the unobserved common factors extracted in our model from high-frequency indicators. For visualization purposes, the figures show them as a three-month moving average (3MMA), comparing it with seasonally adjusted quarterly GDP growth in fixed prices. It seems that all the extracted factors have intuitively correct dynamics. The factor of the real sector has a strong correlation with the GDP growth rate (Figure 4), this may significantly enhance the accuracy of GDP forecasts due to earlier publication of the real sector statistics. The expectations factor reflects not only changes in economic activity, but also individual institutional factors (e.g. the increase in the VAT base rate in Russia in early 2019; Figure 5), which should also improve the accuracy of model-based forecasting. Finally, the financial sector factor may be a leading indicator since the effect of economic shocks is reflected in them considerably earlier than formal statistics on GDP are released (Figure 6).

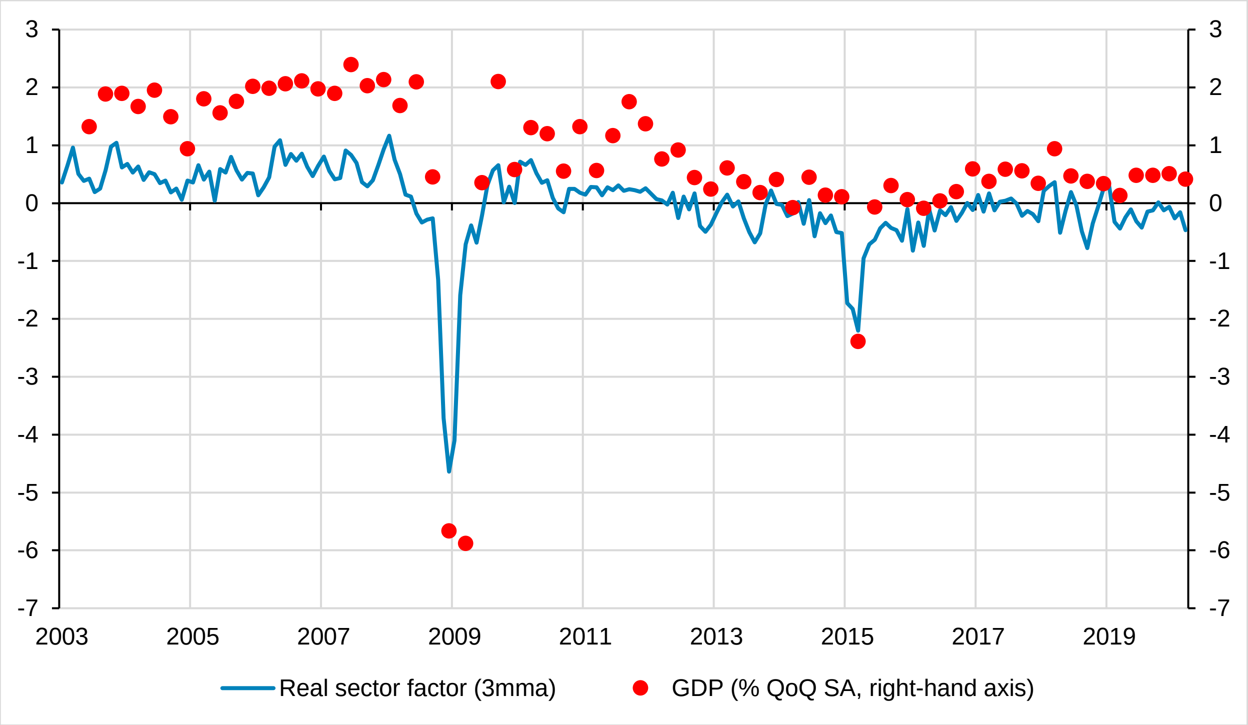

Figure 4 Real sector factor and GDP growth

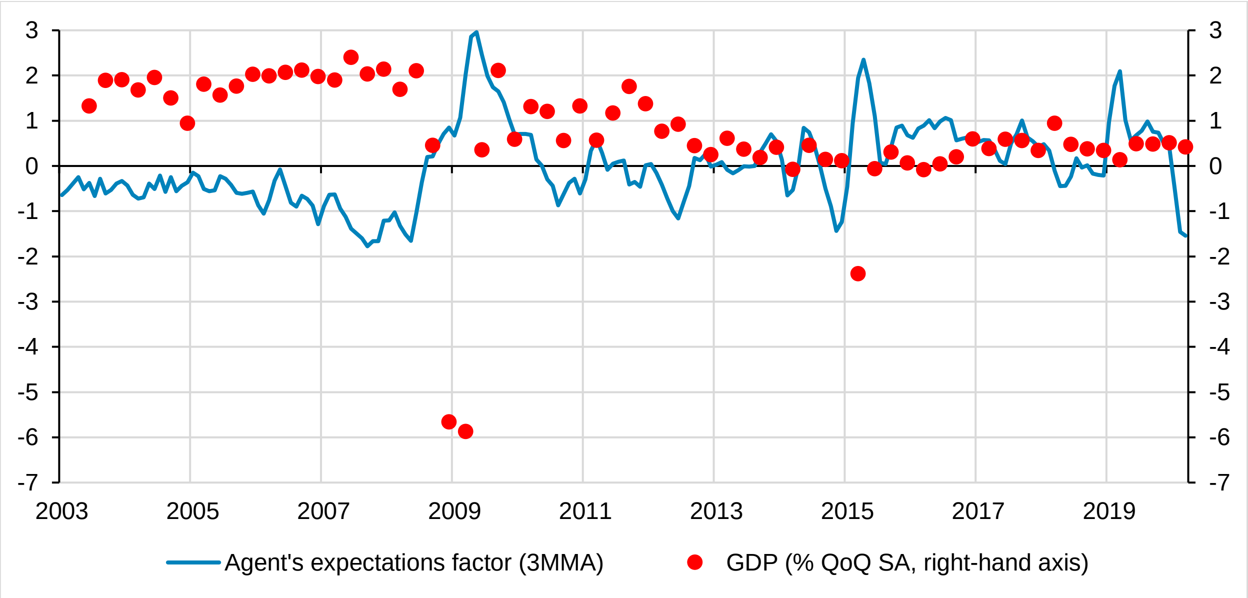

Figure 5 Agent's expectations factor and GDP growth

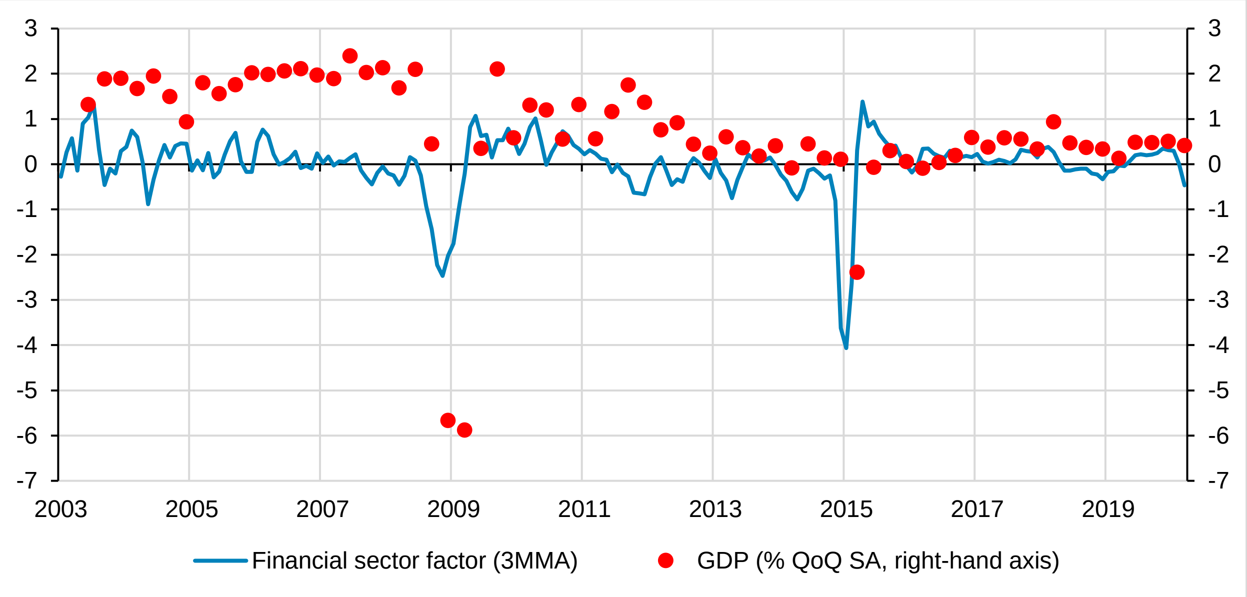

Figure 6 Financial sector factor and GDP growth

Table 1 presents the main results of our research. In addition to the benchmark models proposed above, the accuracy of the forecast combination model is also compared against its two additional variations – forecast combination without additional models (ARMA, RW, UC) and forecast combination using equal weights for all models. The values in the table denote the ratio of the MSFE of a given model and forecast horizon to the MSFE of the naive forecast. Accordingly, the values exceeding one imply that the naive forecast was more precise, while the values below one mean that a particular model produced a more accurate forecast. The period of four quarters was chosen as the longest forecast horizon. We suppose that it useless to prepare longer-term forecasts using this technique since our model, just as any other stationary one, tends to converge to the sample mean.

Table 1 Main results

Period	ARMA	DFM	DMA	DMS	FAVAR	Forecast combination (without additional models)	Forecast combination	Forecast combination: equal weights
1	0.86	0.88	0.81	1.13	1.79	**0.78**	0.79	0.80
2	0.87	0.87	0.94	1.10	2.35	0.91	**0.81**	0.85
3	0.90	0.90	1.48	1.97	2.38	0.89	**0.81**	0.83
4	0.88	0.87	1.99	1.66	2.21	0.84	**0.75**	0.79

The forecast combination technique suggested in this research demonstrates better prediction accuracy compared to all reviewed benchmark models. Moreover, the truncation of additional models improves the accuracy of the forecast for one quarter, while their incorporation increases the accuracy of forecasting for two to four quarters. This is in line with a study by Giannone, Reichlin, and Small (2008) stating that factor models yield a higher prediction accuracy compared to simple models only for the nearest quarter. Furthermore, the forecast combination using weights based on the historical accuracy of the models enhances the accuracy of forecasts for two to four quarters, compared to models with equal weights. Nonetheless, the difference in the accuracy of forecasts for one quarter yielded by these variations is insignificant (0.78 against 0.80). As compared with the alternative models, forecast combination outperforms all other benchmark models in terms of prediction accuracy. This difference is significant over the forecast horizon from two to four quarters, but in one-quarter forecasting the difference between the prediction accuracy yielded by our technique and the dynamic model averaging approach is not that considerable (0.78 against 0.81, respectively).

An important result of the evaluation of the forecasting quality is that the technique suggested does not generate a systematic error (Figure 7). This may evidence that the specification of our equations is correct and that we have adequately considered structural shifts in the economy and interdependencies between indicators. Overall, the average deviation of a forecast from actual data gradually diminishes when the periods from 2010 to 2020 are compared. The deviation of the 2020 Q1 forecast from actual data could be caused by dramatic changes in the indicators of external conditions and expectations amid the outbreak of the coronavirus infection in western countries, which however did not have a serious impact on the indicators of Russia's real sector in Q1.

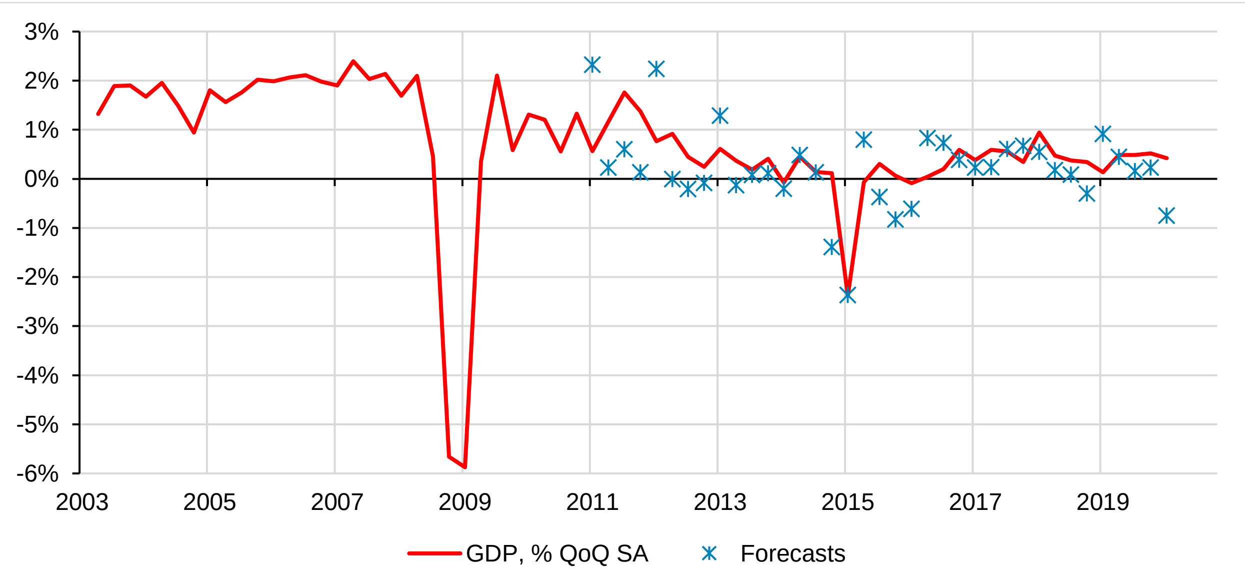

Figure 7 Out-of-sample pseudo real-time forecasts

6. Conclusions

Our approach yields better accuracy of out-of-sample GDP forecasts in Russia for the period from 2011 to 2020 in comparison with the alternative benchmark models. Nonetheless, the difference in the prediction accuracy over a one-quarter forecast horizon is insignificant, and therefore, as the sample expands in the future, it may be necessary to reassess the quality of models. The truncation of additional models from the forecast combination improves the quality of forecasts, but only over a short-term horizon. An important result is that our technique does not generate a systematic forecast error, which may evidence for adequate consideration about structural shifts in the economy and interdependencies between indicators. However, our research involves certain restrictions. When a process generating data is known for researcher and it is possible to build an 'ideal' model, forecast combination may yield less accurate estimates. The uncertainty regarding the weights of selected models and the methods used to calculate the model weights might also entail a certain error in forecasting. A combined forecast may not be decomposed by key indicators (because of the combinations of multiple models with various factors, the estimates of their contributions may be shifted).

This research provides an extensive coverage of the topic, yet there is a tremendous potential for its further elaboration. Possibly, the vintage data could be the most important improvement of this study. This should enhance the integrity of the experiment, since macroeconomists most often have to compare their forecasts against the earliest GDP estimates, and not the final measurements revised later. However, it is rather hard to predict how this may affect the assessments of our model. Another important feature is the usage of not only the GDP by expenditure components, but also of the supply-side components of GDP. Currently, the usage of these components is limited due to the insufficient size of the sample, which is because the codes in the OKVED system were changed and certain data series are incomparable. Another area in nowcasting is the use of mixed-frequency factor-augmented vector autoregressive (MF-FAVAR) models. A serious disadvantage of this class of models is the need to estimate a large number of parameters, which causes certain problems when a data series is relatively short. In our case, we were unable to get any sustainable estimates with MF-FAVAR model. However, in the future, as the sample expands, it will be possible to carry out this assessment again.

Annex 1. Data used in the paper

Group 1: Real sector indicators

1. Retail turnover
2. Retail turnover: food products
3. Retail turnover: non-food goods
4. Commercial services
5. Nominal wages
6. Real wages
7. Wholesale turnover
8. Construction: completed works
9. Freight turnover
10. Agricultural output index
11. Unemployment rate
12. Key Industry Index
13. Industrial production index
14. Industrial production index: mining and quarrying
15. Industrial production index: manufacturing
16. Industrial production index: electric power, gas and water production and distribution
17. Labour force size
18. Labour demand
19. Railway cargo shipments
20. Diffusion index of wages: actual changes
21. Diffusion index of employment: actual changes
22. Diffusion index of output: actual changes
23. Diffusion index of order portfolio: actual changes
24. Diffusion index of finished goods inventory: actual changes
25. Diffusion index of output/input prices ratio: actual changes
26. Diffusion index of equipment purchases: actual changes
27. Capacity utilisation rate
28. Labour utilisation rate
29. Finished goods inventory
30. Order portfolio
31. Debt to banks
32. Portion of companies having ‘good’ or ‘normal’ financial standing
33. Percentage of companies not purchasing equipment for two or more consecutive months
34. Interest rates on bank loans (in rubles) to be raised in the next three months
35. Percentage of companies not indebted to banks and not going to be indebted in the next three months
36. Percentage of companies not going to raise new bank loans in the next three months

Group 2: Financial sector indicators

1. Interest rate on ruble-denominated household deposits: for up to one year
2. Interest rate on ruble-denominated household deposits: for more than one year
3. Interest rate on ruble-denominated non-financial organisations’ deposits: for more than one year
4. Interest rate on ruble-denominated household loans: for up to one year, including demand loans
5. Interest rate on ruble-denominated household loans: for more than one year
6. Interest rate on ruble-denominated loans to non-financial organisations: for up to one year, including demand loans

7. Interest rate on ruble-denominated loans to non-financial organisations: for more than one year
8. Weighted average actual credit rate (MIACR): on overnight loans
9. RTS Index
10. MOEX Index
11. Interest rate on ruble-denominated household loans: for up to one year, including demand loans
12. Interest rate on ruble-denominated household loans: for more than one year
13. Interest rate on ruble-denominated loans to non-financial organisations: for up to one year, including demand loans
14. Interest rate on ruble-denominated loans to non-financial organisations: for more than one year
15. Interest rate on ruble-denominated mortgage loans: on average, year-to-date
16. International reserves

Group 3: Agents’ expectations

1. Companies’ price expectations
2. Companies’ price expectations: mining and quarrying
3. Companies’ price expectations: manufacturing
4. Companies’ price expectations: industrial production
5. Companies’ price expectations: agriculture, forestry, hunting, fishing and fish-breeding
6. Companies’ price expectations: construction
7. Companies’ price expectations: wholesale and retail trade; repair of motor vehicles and motorcycles
8. Companies’ price expectations: wholesale trade, except of motor vehicles and motorcycles
9. Companies’ price expectations: retail trade, except of motor vehicles and motorcycles
10. Companies’ price expectations: transportation and storage
11. Companies’ price expectations: services
12. PMI Composite SA
13. PMI Manufacturing SA
14. PMI Services SA
15. PMI Composite New Orders SA
16. PMI Manufacturing New Orders SA
17. PMI Services New Business SA
18. PMI Composite Employment SA
19. PMI Manufacturing Employment SA
20. PMI Services Employment SA
21. PMI Composite Input Prices SA
22. PMI Manufacturing Input Prices SA
23. PMI Services Input Prices SA
24. PMI Composite Output Prices SA
25. PMI Manufacturing Output Prices SA
26. PMI Services Prices Charged SA
27. CEIC Leading Indicator: Russia
28. Higher School of Economics: leading indicator
29. Diffusion index of output prices expected changes
30. Diffusion index of input prices: expected changes
31. Diffusion index of wages: expected changes
32. Diffusion index of employment: expected changes
33. Diffusion index of output: expected changes
34. Diffusion index of equipment purchases: expected changes

35. Diffusion index of financial situation: expected changes
36. Diffusion index of order portfolio: expected changes
37. Diffusion index of debt to banks: expected changes

References

Achkasov Yu. Nowcasting of the Russian GDP using current statistics: approach modification //The Bank of Russia. Working Paper Series. – 2016. – No. 8.

Andreev A. Inflation forecasting in the Bank of Russia by combining forecasts //The Bank of Russia. Working Paper Series. – 2016. – V. 14. – P. 2–11.

Andreou E., Ghysels E., Kourtellos A. Regression models with mixed sampling frequencies //Journal of Econometrics. – 2010. – V. 158. – No. 2. – P. 246–261.

Baffigi A., Golinelli R., Parigi G. Bridge models to forecast the euro area GDP //International Journal of Forecasting. – 2004. – V. 20. – No. 3. – P. 447–460.

Bernanke B. S., Boivin J., Eliasz P. Measuring the effects of monetary policy: a factor-augmented vector autoregressive (FAVAR) approach //The Quarterly Journal of Economics. – 2005. – V. 120. – No. 1. – P. 387–422.

Boivin J., Ng S. Are more data always better for factor analysis? //Journal of Econometrics. – 2006. – V. 132. – No. 1. – P. 169–194.

Clements M. P., Galvão A. B. Macroeconomic forecasting with mixed-frequency data: Forecasting output growth in the United States //Journal of Business & Economic Statistics. – 2008. – V. 26. – No. 4. – P. 546–554.

Doz C., Giannone D., Reichlin L. A two-step estimator for large approximate dynamic factor models based on Kalman filtering //Journal of Econometrics. – 2011. – V. 164. – No. 1. – P. 188–205.

Fair R. C., Shiller R. J. Comparing information in forecasts from econometric models //The American Economic Review. – 1990. – P. 375–389.

Ghysels E., Santa-Clara P., Valkanov R. The MIDAS touch: Mixed data sampling regression models. – 2004.

Ghysels E., Santa-Clara P., Valkanov R. Predicting volatility: getting the most out of return data sampled at different frequencies //Journal of Econometrics. – 2006. – V. 131. – No. 1–2 – P. 59–95.

Giannone D., Reichlin L., Small D. Nowcasting: The real-time informational content of macroeconomic data //Journal of Monetary Economics. – 2008. – V. 55. – No. 4. – P. 665–676.

Hendry D. F., Hubrich K. Combining disaggregate forecasts or combining disaggregate information to forecast an aggregate //Journal of Business & Economic Statistics. – 2011. – V. 29. – No. 2. – P. 216–227.

Hubrich K. Forecasting euro area inflation: Does aggregating forecasts by HICP component improve forecast accuracy? //International Journal of Forecasting. – 2005. – V. 21. – No. 1. – P. 119–136.

Koop G., Korobilis D. UK macroeconomic forecasting with many predictors: Which models forecast best and when do they do so? //Economic Modelling. – 2011. – V. 28. – No. 5. – P. 2307–2318.

Kuzin V., Marcellino M., Schumacher C. MIDAS vs. mixed-frequency VAR: Nowcasting GDP in the euro area //International Journal of Forecasting. – 2011. – V. 27. – No. 2. – P. 529–542.

Marcellino M., Porqueddu M., Venditti F. Short-term GDP forecasting with a mixed-frequency dynamic factor model with stochastic volatility //Journal of Business & Economic Statistics. – 2016. – V. 34. – No. 1. – P. 118–127.

Marcellino M., Schumacher C. Factor MIDAS for nowcasting and forecasting with ragged-edge data: A model comparison for German GDP //Oxford Bulletin of Economics and Statistics. – 2010. – V. 72. – No. 4. – P. 518–550.

Mikosch H., Zhang Y. Forecasting Chinese GDP growth with mixed frequency data: Which indicators to look at?. – 2014.

Mikosch H. and Solanko L. Forecasting quarterly Russian GDP growth with mixed-frequency data //Russian Journal of Money and Finance. – 2019. – V. 78. – No. 1. – P. 19–35.

Pinkwart N. Short-term forecasting economic activity in Germany: A supply and demand side system of bridge equations. – 2018.

Porshakov A., Ponomarenko A., Sinyakov A. Nowcasting and short-term forecasting of Russian GDP with a dynamic factor model //Journal of the New Economic Association. – 2016. – P. 60.

Raftery A. E., Kárný M., Ettler P. Online prediction under model uncertainty via dynamic model averaging: Application to a cold rolling mill //Technometrics. – 2010. – V. 52. – No. 1. – P. 52–66.

Schumacher C., Breitung J. Real-time forecasting of German GDP based on a large factor model with monthly and quarterly data //International Journal of Forecasting. – 2008. – V. 24. – No. 3. – P. 386–398.

Stock J. H., Watson M. W. Implications of dynamic factor models for VAR analysis. – National Bureau of Economic Research, 2005. – №. w11467.

Styrin K. Forecasting inflation in Russia using dynamic model averaging //Russian Journal of Money and Finance. – 2019. – V. 78. – No. 1. – P. 3-18.

Timmermann A. Forecast combinations //Handbook of Economic Forecasting. – 2006. – V. 1. – P. 135–196.

Wallis K. F. Forecasting with an econometric model: The ‘ragged edge’ problem //Journal of Forecasting. – 1986. – V. 5. – No. 1. – P. 1–13.

Zheng I. Y., Rossiter J. Using monthly indicators to predict quarterly GDP. – 2006.