Forecasting Russian Macroeconomic Indicators Based on Information from News and Search Queries¹

Filipp Ulyankin, *Russian Presidential Academy of National Economy and Public Administration*
filfonul@gmail.com

Modern economic literature features quite a number of various indices of economic activity. Some of them are based on consumer and business surveys ('manual' indices), while others are based on unstructured data from the Internet ('automatic' indices). However, the question as to which of these approaches is the most effective remains open. In this paper, we compare several different indices of economic activity in terms of their explanatory and predictive power. We build 'automatic' indices using machine learning methods. Search queries, news articles and user comments under news posts from social media are used as source data. The analysis of the resulting indices of economic activity shows that the search and news indices Granger-cause 'manual' indices and also better explain and predict the set of macroeconomic variables selected for research. The good explanatory power of the current values of macroeconomic indicators by means of current indices of economic activity with a lag in the release of macroeconomic statistics makes them suitable for nowcasting.

Keywords: *text analysis, sentiment analysis, economic uncertainty index, data analysis, machine learning*

JEL Codes: C43, C55, C81

Citation: Ulyankin, F. (2020). Forecasting Russian Macroeconomic Indicators Based on Information from News and Search Queries. *Russian Journal of Money and Finance*, 79(4), pp. 75–97.

doi: 10.31477/rjmf.202004.75

1. Introduction

For a timely assessment of the state of the economy and early detection of economic instability, leading indicator systems are used. They help monitor and forecast business activity and reduce the time intervals required for making proactive decisions that are key for the macroeconomic policy. Many such indicators are constructed manually through of consumer and business surveys. With such

¹ The author would like to thank Andrey Polbin (Russian Presidential Academy of National Economy and Public Administration (RANEPA); Gaidar Institute for Economic Policy) for his valuable corrections and comments.

an approach, an index of economic activity can be constructed fairly rarely (once a month or even once a quarter).

Starting with Choi and Varian (2009) and Choi and Varian (2012), many authors suggest that the building of such indicators can be automated by processing large arrays of data generated as a result of human-internet interaction using modern statistical methods. Generally, one of three data sources is used for automation:

- Search queries (search index);
- News stream (news index);
- Stream of comments from social media (sentiment index).

An indicator is constructed from manually selected descriptors. Then an indicator is checked for explanatory power, and it is determined to what extent it helps explain a certain macroeconomic indicator. However, authors generally use only one of the approaches, and it is therefore not clear which approach works best.

In this paper, we construct several low-frequency (monthly) indicators reflecting the current state of the economy and compare them with each other and with the indicators constructed manually.

2. Literature review

2.1. 'Manual' indices

Consumer and business surveys have long been used to build economic confidence indices. Most indices are constructed to reflect the level of confidence: if the index falls, it means that people trust the economic situation less, and if it grows, it means that their confidence in the economic situation is strengthening.

The best known and the oldest of such indices is the University of Michigan Consumer Sentiment Index.² The University of Michigan conducts telephone polls of a sample of 500 consumers twice a month. Each of the consumers is asked five questions concerning their financial standing and their opinion of the current state of the economy. The index value is calculated as the difference between the share of respondents who noted an improvement in the economic environment and the share of those who noted its deterioration.

Such indices are constructed in a similar way in many countries; in particular, in Russia the Michigan Index methodology is used to build the consumer sentiment index³ (CSI), which has been calculated by the Levada Centre since the 1990s. Polls include 1,600 respondents from various regions of Russia. Questions are asking their assessments of their personal financial situation, the current economic situation, and their expectations about the future. After that, the collected responses are aggregated. In 2009–2017, Sberbank's Centre for Macroeconomic Research and the Levada Centre constructed a similar monthly

² See <https://fred.stlouisfed.org/series/UMCSENT>

³ See <https://www.levada.ru/indikatory/sotsialno-ekonomicheskie-indikatory/>

index, the Financial Sentiment Index⁴ (FSI), with a focus on financial rather than consumer behaviour.

Confidence indices are taken as a basis for the development of models identifying the relationship between the index level and various macroeconomic series. For example, Matsusaka and Sbordone (1995) showed that the Michigan Index Granger-causes the US GDP, and Bram and Ludvigson (1998) concluded that the index is statistically significant in explaining the dynamics in several categories of consumer spending. Blanchard (1993) found that there was a sharp, inexplicable collapse in the University of Michigan Consumer Sentiment Index before the recession associated with the invasion of Kuwait and the surge in oil prices. Similar studies have been done for other countries, such as Sweden (Berg and Bergstrom, 1996) and Japan (Utaka, 2003).

Consumer and business survey and manual construction of confidence indices is very resource-intensive. Moreover, such indices are difficult to build on a daily basis. This raises the need for automation of the index building process.

2.2. Search indices

Search engines collect statistics on what users are interested in. Such statistics are partly in the public domain. In 2010, the search engine company Google claimed that it could predict a film's box-office receipts in the early days of its release with 94% accuracy using information from queries for trailers of upcoming films (Goel et al., 2010). In addition, the search queries can predict the spread of influenza epidemics (Ginsberg et al., 2009) and coronavirus outbreaks (Li et al., 2020).

Search data can be useful for building financial indicators. For example, McLaren and Shanbhogue (2011) showed that there is a significant correlation between the number of queries for descriptors related to unemployment and actual unemployment. The authors estimated the second-order autoregression in differences, to which the search index was added as an explanatory variable. Search queries were included in the model without a lag as people are looking for a job today. As a result of the estimation, a significant coefficient for the search index in the regression for unemployment rate was obtained. The model with the index showed lower values of information criteria.

McLaren and Shanbhogue (2011) also found a similarly significant effect for the UK housing prices. In this case, search queries are included in the model with a lag since after a person finds an apartment, they may need some time to negotiate and close a purchase deal. Similar results were obtained earlier by D'Amuri (2009) for Italy and Choi and Varian (2009) for the US.

Choi and Varian (2012) estimated seasonal AR models for such indicators as car sales, applications for unemployment benefits, and air travel distance and obtained improved values for various model characteristics. The information

⁴ See <https://www.sberbank.ru/ru/about/analytics/mood>

criteria indicate the preferability of expanded specifications with the search indices as additional explanatory variables. Also, extended specifications have stronger predictive powers as measured by the Root Mean Square Error (RMSE). The search indices in this paper were interpreted as the desire to find a job, buy a house, etc.

In the context of financial markets, the dynamics of Google queries is capable of recording changes in investor attention to a certain asset. For example, Da et al. (2011) analysed the dynamics of stock quotes of 3,000 public companies in the United States and Google queries on them, and show that the dynamics of search queries positively correlates with the trade volume, and an increase in the number of queries leads to an increase in the market value of the company's stock within two weeks with a subsequent return to the original level. In this paper, the authors use the Structural Vector Autoregression (SVAR) model and analyse impulse responses.

Stolbov (2011) made an attempt to construct a financial sentiment index for Russia based on search queries. The author identified the terms related to financial markets and insurance that were searched most often by people in Russia during the peak of the GFC from September 2008 to June 2009. He selected these descriptors and diluted them with a few more terms.⁵ Search queries for them were then obtained from Google Trends on a monthly basis and were added with certain weights to the final index. As an alternative, the Principal Component Analysis (PCA) method was considered, which makes it possible to single out the first principal component in the index fluctuations. As a result, a significant correlation was established between the constructed index and the RTS dynamics. However, as in Choi and Varian (2009, 2012), the information criteria and the RMSE indicated the preferability of an expanded model.

2.3. News indices

The second source of data for building economic indicators is the stream of news texts. The easiest way to process texts is to build news indices by analogy with search indices. Researchers identify a pool of descriptors and track how often they appear in the media.

This is the approach taken by Baker, Bloom, and Davis (Baker et al., 2016). They construct the Economic Policy Uncertainty Index by tracking the use of such words as 'uncertain', 'uncertainty', 'economy' and a number of others describing economic policy in ten leading American newspapers. The authors demonstrated that before significant events their index took on high values, which signalled an increase in uncertainty.

Over time, Baker, Bloom and Davis made the methodology more complex: they began to include more sources for calculating the frequency index and expanded the list of countries for which it was calculated. For Russia, they construct the index

⁵ Bank, dollar exchange rate, euro, Central Bank, RTS, MICEX, default, devaluation, financial crisis, stock, mortgage, take a loan, unit investment fund, bankruptcy, collateral.

based on the newspaper *Kommersant*. All indices and a detailed description of the methodology are published online⁶ and are updated on a regular basis.

A similar index for Russia is published by Goloshchapova and Andreyev on the website of their Big Data Indicators project.⁷ They calculate how often the words 'crisis', 'recession' and 'decline' appear in the Russian media.

2.4. Sentiment indices

Social media is the third data source for modelling a financial indicator. People leave a huge amount of comments about current events. Their comments reflect the tone of the news flow within which investors are currently making their decisions.

Typically, such indices are built on the basis of the flow of posts on Twitter. Using a pre-trained classifier, researchers assess the sentiment of comments under news of interest. Then, the proportion of negative messages is calculated.

Mao et al. (2012) studied the correlation between the number of negative tweets mentioning the S&P 500 index and its dynamics. It turned out that the data calculated from Twitter can be useful for predicting the stock market prices.

Bollen et al. (2011) decided to use a wider range of emotions to classify tweets rather than just categorising them into negative and positive. They introduced six categories for tweets called Calm, Alert, Sure, Vital, Kind, and Happy, and also constructed an index representing the ratio of positive to negative tweets (OF). It was found that the OF, Calm, and Happy dynamics are Granger-causative of returns. Granger causality was analysed based on a linear model, yet the ratios between public sentiments and stock market prices might be non-linear. To account for the non-linearity, the authors trained a neural network. They concluded that adding the Calm and Happy indicators improved the forecast accuracy on a test sample, while adding other variables (including OF) did not improve the accuracy.

Besides Twitter posts, researchers try to analyse the impact of investor sentiments that are posted on various reputable financial blogs. For example, Frisbee (2010) examines the correlation between market forecasts made by financial bloggers and the S&P index. The author concludes that information from the blogs chosen by the author does not help forecasting in any way.

Goloshchapova and Andreyev (2017) construct an index of uncertainty in inflationary expectations as well as an index describing their intensity based on the sentiment tone of user comments in the media. The indicators they obtained turn out to be relevant to the main macroeconomic trends.

A sentiment index can also be constructed based on news texts. For example, Yakovleva (2018) constructed the news index using thematic modelling. The author collected around 50,000 news items, published in the period from January 2014 through September 2017, from a certain internet resource. She distinguished

⁶ See <http://www.policyuncertainty.com/index.html>

⁷ See <http://bigdata-indicators.com/index.html>

50 key topics using the LDA (Latent Dirichlet Allocation) model. To determine the sentiment of texts, the author took a sample from news and trained the classifier. This allowed her to divide topics into positively coloured and negatively coloured. All neutral news was filtered out. To obtain data on a daily basis, all articles for the day were summarised, after which the frequency of each topic for the day was calculated based on their five most probable words. The data obtained made it possible to determine how often each topic was mentioned during the day. After that, using a moving average, the daily noise was removed, and all topics were aggregated into monthly indicators by calculating the average value for the month. The result was 50 monthly time series, one per topic. Five principal components were identified using PCA. Then, the methodology of Choi and Varian (2012) was used with PMI (Purchasing Managers' Index) taken as a dependent variable. On the test sample, the model taking into account the dynamics of the principal components demonstrated a better predictive result compared to AR.

Recently, there have been quite a lot of research papers on economic activity indices constructed from unstructured data. Most of these papers show that such indicators can be used to explain the dynamics of various macroeconomic and financial variables. However, it remains unclear how the various approaches used in modelling compare against one another.

In this paper, we compare the predictive properties of a number of popular indicators, which were described above, using Russian data, and supplement them with our own indicators.

3. Data

To collect search data, news, and comments, special parser scripts were written in Python. All of the code used in the calculations and collection of information is publicly available on the author's GitHub.⁸ The total amount of collected information turned out to be very large. Upon request, we can provide initial data to other researchers.

Search data. All statistics related to search queries are obtained from the Google Trends service.⁹ The service provides data on search queries in a scaled form. The most frequent query is assigned the value 100. The rest of the frequencies are normalised to this query. No more than five descriptors are loaded at a time. Search data have been available since January 2004 on a monthly basis.

News data. News texts were collected from the websites of a number of major news agencies, such as RBC, RIA Novosti, TASS, Interfax, and Lenta. These sources were selected due to the availability of news for a fairly long period of time and because of the convenient design of their websites, which makes it possible to automate the collection of information. A total of about 5.5 million news entries

⁸ See https://github.com/FUlyankin/uncertainty_index

⁹ See <https://trends.google.ru/trends>

were downloaded. All news items were filtered by categories related to business and politics. Basic information about the news sample is provided in Table. 1.

Table 1. Basic information about news sources

Source	First news item	Last news item	Total news items
RBC	01.01.2012	01.06.2020	330,000
RIA Novosti	01.01.2004	01.06.2020	3,320,000
TASS	01.01.2011	01.06.2020	940,000
Interfax	01.01.2008	01.06.2020	580,000
Lenta	01.01.2008	01.06.2020	590,000

Source: author's calculations

We worked with the texts following a standard process (pipeline). All texts were cleared of obvious errors: news with empty texts or headlines, wrong dates, etc. Each text was reduced to lower case and tokenised – that is, it was split into separate significant units.

We used lemmatisation (reducing words to a standard – that is, dictionary form) as a normalisation algorithm.¹⁰ Stemming (finding the stem of a word) for the Russian language performs worse than lemmatisation since truncation of a word to its root stem does not always work. For example, when using an algorithm implementing stemming, it will not be possible to turn the word 'lyudi' ('people') into the word 'chelovek' ('person'). Lemmatisation has no such disadvantage. However, we should understand that despite the high quality of modern morphological analysers there are ambiguities that cannot be resolved by an algorithm working with only one word. For example, the words that are spelled the same but have different meanings: the word 'stali' can be used as a noun meaning 'steel' ('zavod proizvodit vse bol'she stali' ('the factory produces more and more steel')) and as a verb meaning 'began' ('my stali luchshe uchit'sya' ('we began to study better'))).

Next, all texts were cleared of stop words. A list from the 'nltk' library was used as the list for stop words. All duplicate entries within each source were removed.

Comments. To build the economic sentiment index, we used comments collected from the accounts of the largest news resources in the social network VKontakte. The sample included Interfax, Kommersant, Vedomosti, Lenta, RBC, Meduza, Rossiyskaya Gazeta, RIA Novosti, RT, and TASS. In this case, we were able to use a larger number of media outlets than in the sample for the news index since information from the accounts of these media outlets in the social network could be collected in the same way. In total, the sample included about 42 million comments for the period from 1 January 2012 through 1 June 2020.

Moreover, all news posts were collected from newsgroups, including both headlines and snippets (short descriptions of news for social media). They were also used to construct indices.

¹⁰ For this purpose, we used the mystem library: <https://yandex.ru/dev/mystem/>

‘Manual’ indices. All ‘automatic’ indices are compared to ‘manual’ indices based on surveys. The following were used as ‘manual’ indices:

1. **Levada Centre’s CSI.** In 2009–2017, the index was built on a monthly basis, but since the beginning of 2018 the index frequency has dropped from one month to two. To fill in the gaps, we average the nearest points. For example, if there is a value for June and August, but there is no value for July, the gap is filled with the average value for June and August.
2. **PMI business activity index.** It is based on the results of a survey of purchasing managers and is used to assess changes in the economic activity. We are considering the PMIs of the industrial and service sectors.
3. **Business Confidence Index (BCI).** A qualitative indicator that helps characterise the economic situation by the responses of managers about predicted production output, product balances, and product demand. It is constructed by Rosstat for such categories as Mining, Manufacturing, and Electricity, Gas and Steam Supply. Small business is not included in the survey. The indicator is the arithmetic mean of balances of the responses to the questions.¹¹ In this paper, we consider the index for manufacturing industries as the most non-specific of the three presented. Due to pronounced seasonality, the index is considered in terms of its increase over the same month of the previous year.

Figure 1. Dynamics of ‘manual’ indices

For convenience, all indices are normalised to a scale of 0 to 100. For example, when normalising the CSI, we proceeded from the fact that the point taken by the index compilers as 100 (corresponding to March 2008) is too far from the main data series, so we took an original value of 95 (September 2008) as the maximum index value (100 on our scale), while the minimum value (0 on our scale) corresponds to an original value of 57 (January 2016). ‘Manual’ indices are used as a reference for comparison with ‘automatic’ ones. It should be separately mentioned that most

¹¹ See https://www.gks.ru/free_doc/new_site/metod/prom/met_pred-uver.htm

of the ‘manual’ indices are constructed so as to reflect the confidence of economic agents in the state of the economy. If the index falls, it means that the share of respondents who are pessimistic about the future of the economy are larger than the share of optimists. ‘Automatic’ indices, on the contrary, reflect people’s distrust in the economy and grow in crisis situations.

The dynamics of ‘manual’ indices from 1 January 2012 till present are shown in Figure 1 (hereinafter, the normalised values are used). It can be seen that the indices move quite synchronously and show negative dynamics in crisis periods.

Target variables. The following indicators are used as target variables:

1. RTS index returns. The data were obtained from the Moscow Exchange website.¹² We consider:

- RTS_LAST_VALUE – the last closing price for the month;
- RTS_CLOSE_MEAN – the average closing price of the index during the month;
- RTS_TRADING_VOLUME_MEAN – the average trade volume for the month.

2. Official exchange rates of the Central Bank of Russia.¹³ The average values of the euro (EUR) and dollar (USD) exchange rates were used for the month.

3. Series associated with bank deposits and loans from the EMISS (Unified Interdepartmental Statistical Information System) database:¹⁴

- DEP_FIZ_SUM – household deposits in roubles;
- IPOT_CNT – the number of mortgage loans issued;
- IPOT_VOLUME – the volume of mortgage loans issued in roubles;
- IPOT_DEBT – debt on loans issued.

4. Macroeconomic series from the hse.stat database:¹⁵

- CPI_M_CHI – consumer price index;
- IP2_EA_M – industrial production index as per OKVED 2;
- UNEMPL_M_SH – end of month unemployment rate;
- WAG_M – real wages;
- WAG_C_M – average nominal wages (roubles per month);
- RTRD_M_DIRI – index of real retail turnover;
- RTRD_M – retail trade turnover (current prices, billion roubles);
- INVFC_M – investments in fixed assets (current prices, billion roubles);
- IM_T_M – imports (billion dollars).

For the exchange rate and the RTS index we calculate the rates of growth month-over-month, that is compared to the previous month. For the rest of the variables we calculate the rates of growth year-over-year, that is compared to the same month of the previous year. Similarly to indices, all variables are normalised to a scale from 0 to 100. This is done so that the forecast errors estimated for different variables are comparable in magnitude.

¹² See <https://www.moex.com/ru/index/stat/monthlyhistory.aspx?code=RTSI>

¹³ See https://cbr.ru/currency_base/dynamics/

¹⁴ See <https://www.fedstat.ru/indicator/42475>

¹⁵ See <http://sophist.hse.ru/hse/nindex.shtml>

4. Building the indices

Search indices are constructed using the methodology from Stolbov (2011). It is calculated based on the descriptors identified by the author using the formula:

$$I = w_1 I_1 + \dots + w_{15} I_{15}, \quad (1)$$

where I_i – is the normalised dynamics of queries from Google Trends for the i -th descriptor, and w_i – is the weight for the i -th descriptor calculated by the formula:

$$w_i = \frac{\sum_{i=1}^{15} r_{ij}}{\sum_{i=1}^{15} \sum_{j=1}^{15} r_{ij}}, \quad (2)$$

where r_{ij} – is the sample correlation between the corresponding descriptors.

As an alternative approach, the first principal component was used. The principal component method attempts to linearly combine the data so as to maintain the highest possible variance. For this reason, one of the first components must be highly responsive to peaks in data. It is this component that we take as an index.

Figure 2. Dynamics of ‘manual’ indices and search indices

Figure 2 shows the dynamics of search indices for the period from January 2004 through May 2020 and ‘manual’ indices that are artificially shifted to the range from 100 to 200. As you can see, the search indices constructed on the basis of descriptor correlations and PCA are similar; however, the index based on PCA is more responsive to peaks, as expected. It can be seen that ‘manual’ and ‘automatic’ indices record the same trend in changes in the financial environment.

The search index in Stolbov (2011) is based on the period up to 2012 and is compared with the FSI of Sberbank and Levada Centre. It should be noted that the author did not have the opportunity to compare the dynamics of the search index with the FSI over a long period (since the FSI only started to be published in May 2009) or to validate it on other crises. In this paper, we consider the period up to

2020, and since the FSI of Sberbank and Levada Centre is no longer constructed, we use the Levada Centre's CSI for our analysis. During the existence of both indices the CSI demonstrated dynamics close to, although not entirely coinciding with, the FSI.

If one chooses descriptors based on some major event, such as the 2008 crisis, there may be an accidental 'overtraining' for this event. However, over an extended period, we see a sharp surge in 'automatic' indices not only in 2008 but also in other periods when the Russian economy was in crisis, for example, in 2014–2015 and during the coronavirus crisis in 2020. This allows us to assert that the index has not been overtrained for a specific event.

The increase in the search index (growth of mistrust) coincides overall with the decline in CSI. Judging by the dynamics, the growth of the search index periodically precedes a drop in the CSI. We believe this is because the CSI is based on consumer surveys. Before answering a survey, a respondent needs to formulate their opinion, which he does based on news and information searches on the internet. It turns out that the search index is constructed based on the sources from which people take information to formulate their opinions. Most likely, this is why we can see such a lead. Further, we will see similar behaviour of the news index, and we will test the hypothesis that these two indices Granger-cause their 'manual' analogues.

It should be separately mentioned that the drop in the search index can be interpreted in two ways. On the one hand, it can be regarded as the absence of pronounced crisis expectations, and, on the other hand, as a lack of interest in financial instruments.

Sentiment indices. Sentiment indices are usually constructed from a stream of tweets selected by tags. Since Twitter API has rather strict restrictions on data downloading, another social network, VKontakte, was chosen as the source of comments.

A distinctive feature of the task for determining sentiments is the need for a marked training sample that would show which texts contain positive information, which contain negative information, and which contain neutral information. At the time of writing of this paper, for the Russian language, there existed only a semi-automatic marking of tweets by sentiments.¹⁶ The manual marking of a small random sub-sample shows that the original marking accuracy is about 80%. This marking is not particularly suitable for us because the model was trained on artificially abrupt texts – as at the time of collection of this sample Twitter had a limit of 140 characters for each tweet. VKontakte has no such limitations. Moreover, the commenting mechanism in VKontakte differs from that in Twitter.

Therefore, instead of training our classification algorithm, we used a pre-trained neural network taken from the IPavlov project.¹⁷ It was trained on a sample

¹⁶ See <http://study.mokoron.com>

¹⁷ See <http://docs.deepavlov.ai/en/master/features/models/classifiers.html>

of comments from VKontakte marked up as part of the ‘rusentiment’ project.¹⁸ Unfortunately, the initial sample was removed from the public domain at the request of the VKontakte administration, so we cannot train on it our own model.

Each comment was run through the neural network and received one of the following tags: positive, negative, neutral, speech (greetings, congratulations, thanks) or skip (sentiment is unclear, jokes, not in Russian, nonsense) (Rogers et al., 2018).

The verdicts for extra-long comments (outliers) were removed from the final sample. A post and its comments remained in the sample only if the post text contained at least one of the crisis descriptors mentioned above.¹⁹

We construct indices based on the following methodologies:

1. **Frequency Sentiment Index.** The share of negative and positive comments under each post is calculated, and the second indicator is subtracted from the first, after which the average value for all posts for the month is calculated
2. **Bayesian Sentiment Index.** Based on a small number of comments, it would be incorrect to say that people perceived a particular piece of news negatively. However, the sample contains quite a few posts with a small number of comments. This can cause distortions in the frequency index. To account for them, we will use the methodology proposed by Davidson–Pilon (2015) for sorting comments on Reddit. The number of negative comments appearing under a post is a binomial random variable. Let the parameter p be the probability of success (negative comment). We do not know what kind of comments people are more inclined to, so we use uniform distribution as the *a priori* distribution. We can obtain the *a posteriori* distribution for the parameter p in analytical form. This will be a beta distribution with the parameters $a = 1 + u$ and $b = 1 + d$, where u is the number of negative comments, and d is the number of other comments. The 5% quantile of this distribution (Q_5) will be the final estimate of how negatively people perceived the post. It can be found using the formula

$$Q_5 = \frac{a}{a+b} - 1.65 \cdot \sqrt{\frac{a \cdot b}{(a+b)^2(a+b+1)}}. \quad (3)$$

Next, the average value of Q_5 is calculated for all posts for the month.

The indices obtained as a result of the calculation are shown in Figure 3. The indices are calculated starting from the end of 2011. Because of a small amount of data in 2012 and 2013, the frequency index has a large variance. As expected, the Bayesian index turns out to be more stable over this period.

By the crisis of 2015, it can be seen that ‘manual’ indices are the first to respond to a deterioration in the economic situation. Judging by the most pronounced

¹⁸ See <https://github.com/text-machine-lab/rusentiment>

¹⁹ Bank, dollar exchange rate, euro, Central Bank, RTS, MICEX, default, devaluation, financial crisis, stock, mortgage, take a loan, unit investment fund, bankruptcy, collateral.

peaks, the sentiment indices have a long delay in their dynamics as compared to the ‘manual’ ones.

Figure 3. Dynamics of ‘manual’ indices and sentiment indices

Most likely, this can be attributed to the fact that people’s emotional reaction occurs after they have thought about the events that have taken place. Research (see, for example, Mao et al., 2012) shows that stock market price dynamics Granger-cause sentiment index dynamics. We are probably seeing a similar effect here.

We also construct several versions of **news indices**. From the news corpus, we select news related to politics and economics. We construct the indices using news texts, not readers’ comments on them.

1. Frequency news indices are constructed based on the same crisis descriptors as the search index using the methodology from Stolbov (2011). Similarly to the search index, we count the number of articles that include each of the crisis descriptors during the day and then weight them with weights from the descriptor correlations, same as in (2). The calculation of the share of articles that contained crisis descriptors per month is considered as an alternative approach. The difference from Baker et al. (2016) lies in a broader list of news sources and the selection of crisis descriptors based on Google Trends.

2. The sentiment news index attempts to consider the connotation of each article without being tied to manually picked crisis descriptors. It is impossible to use a classifier trained on comments for sentiment analysis of news because of the different vocabulary used in comments and news. To assess the sentiment connotation of news articles, we use the tonal dictionary of the Russian language²⁰ compiled by the Word Map project using crowdsourcing. In the project, the respondents were asked to rate different words as neutral, positive, or negative. They also could choose the response ‘I don’t know’. Based on the results of this marking, we calculated the indicators responsible for confidence in the verdict. Highly inconsistent verdicts were removed. In practice, high inconsistency occurs in situations where the word score is

²⁰ See https://github.com/dkulagin/kartaslov/tree/master/dataset/emo_dict

highly context-dependent. We also removed the words related to medicine, viruses, and the epidemic. We replace all the words in the title and the text of the article with their 'strength of expression of the emotional and evaluative charge' obtained from the tonal marking. Positive values meant positive connotations of words, while negative values implied negative connotation. The stronger the emotion is the higher is the word's absolute score. Words that are not in the dictionary are assigned zero value. We estimate the final connotation score of the article as the sum of the word scores divided by the number of words in the article. Then the average score is calculated for all articles within one month.

Figure 4. Dynamics of 'manual' indices and news indices

All indices are broken down by different sources (media outlets). After that, we estimate the share of news for each of the media outlets in the sample during the month, multiply the indices by the corresponding shares, and sum them up.

The indices obtained as a result of the estimation are shown in Figure 4. You can see that similarly to the search index the largest peaks in the news index also precede the largest drops in the 'manual' indices. The sentiment news index turns out to be quite noisy as it is impossible to distinguish sharp and easy-to-interpret peaks in its dynamics.

5. Interrelation of indices

To see how the indices relate to each other, we calculate the Pearson correlation. The calculation results (see Figure 5) show that there is a moderate correlation between 'manual' and 'automatic' indices. This suggests that we do not fully reproduce the results of 'manual' indices but at the same time pick up similar trends.

Based on the resulting correlation matrix, several index clusters can be distinguished. For example, the news and search indices correlate quite strongly with each other. This is probably because common descriptors were involved in

their construction. The appearance of crisis descriptors in the news is accompanied by their appearance in the search. Crisis events attract the attention of people and the media at the same time. Hence we have a high correlation. We also see a negative correlation between automatic and ‘manual’ indices. This can be attributed to the fact that in a crisis situation ‘automatic’ indices increase, while ‘manual’ ones, on the contrary, fall.

Figure 5. Correlation between indices

1.00	0.82	0.40	0.22	0.36	0.57	-0.14	-0.21	-0.29	-0.40	-0.31	1
0.82	1.00	0.31	0.17	0.63	0.78	-0.39	-0.20	-0.15	-0.38	-0.47	2
0.40	0.31	1.00	0.77	0.32	0.47	0.40	-0.28	-0.29	-0.30	-0.44	3
0.22	0.17	0.77	1.00	0.20	0.33	0.30	-0.21	-0.19	-0.16	-0.37	4
0.36	0.63	0.32	0.20	1.00	0.75	-0.25	-0.16	-0.07	-0.36	-0.42	5
0.57	0.78	0.47	0.33	0.75	1.00	-0.54	-0.37	-0.30	-0.47	-0.49	6
-0.14	-0.39	0.40	0.30	-0.25	-0.54	1.00	-0.13	0.18	0.07	0.40	7
-0.21	-0.20	-0.28	-0.21	-0.16	-0.37	-0.13	1.00	0.05	0.23	0.33	8
-0.29	-0.15	-0.29	-0.19	-0.07	-0.30	0.18	0.05	1.00	0.77	0.57	9
-0.40	-0.38	-0.30	-0.16	-0.36	-0.47	0.07	0.23	0.77	1.00	0.56	10
-0.31	-0.47	-0.44	-0.37	-0.42	-0.49	0.40	0.33	0.57	0.56	1.00	11
1	2	3	4	5	6	7	8	9	10	11	

1 – Search Index (corr); 2 – Search Index (PCA); 3 – Bayesian Sentiment Index;

4 – Frequency Sentiment Index; 5 – Frequency News Index; 6 – Frequency News Index (corr);

7 – Sentiment News Index; 8 – CSI; 9 – Manufacturing PMI; 10 – Services PMI; 11 – BCI

6. Granger causality

Observing the dynamics of the indices, you can see that the largest peaks in the news and search indices precede the largest drops in the ‘manual’ indices. The news agenda and searches seem to pick up changes earlier than consumer and purchasing manager surveys.

Let’s see if the dynamics of ‘automatic’ indices really contains information that helps predict the dynamics of ‘manual’ indices. To do this, we conduct tests for Granger causality for all pairs of indices (see Figure 6).

In the cells of Figure 6, the p-values are given for testing the hypothesis that the coefficients in the VAR model are insignificant. All cells with a p-value of over 0.05 are empty. The target variable in the equation for the cause turns out to be insignificant in all of the above models. From the series of tests conducted, we can see that the news and search indices are Granger-causative of ‘manual’ indices. The sentiment indices are not Granger-causative of ‘manual’ indices.

People learn about crisis events from the Internet and the news. After that, managers use this information to make purchasing decisions for the next month, and respondents assess the state of the economy in surveys.

The search and news indices focus on the sources from which people learn information. ‘Manual’ indices and sentiment indices are guided by the people’s

existing opinions. Because of this, they lag behind in dynamics and turn out to be less effective indicators of financial conditions. It should be separately noted that such a lag cannot be very large. For this reason and also because we work with monthly data, one lag is selected for Granger causality testing.

Note that Granger causality is not a sufficient condition for causation. It just suggests that the search and news indices contain information that can be used to predict 'manual' indices.

Figure 6. Granger causality

0.0000	0.0000	0.0001	0.0000	Search Index (corr)
0.0001	0.0010	0.0020	0.0000	Search Index (PCA)
				Bayesian Sentiment Index
				Frequency Sentiment Index
				Frequency News Index
0.0014	0.0113	0.0049		Frequency News Index (corr)
				Sentiment News Index
CSI	Manufacturer PMI	Services PMI	BCI	

7. Predictive power of sentiment indices

In this section, we use the methodology from Choi and Varian (2012) with our improvements for all constructed indices and all selected target variables:

1. We split the sample into training and test samples. The test sample includes the last 30% of observations that are present in the data for the respective index and target variable.
2. Using the training sample, we estimate the Autoregressive Integrated Moving Average (ARIMA) model for the target variable. The optimal parameters for the model are selected by minimising the Schwarz information criterion. We then fix model parameters and work with the test sample.
3. We calculate a forecast for the first test point and a forecast error. After adding this observation to the training sample, we train the same model, taking into account the new observation. Now we calculate a forecast for the next observation from the test sample and calculate the forecast error again. We repeat this procedure until we reach the end of the test sample (leave-one-out cross-validation on a rolling basis). We calculate the average forecast error. As a metric for calculating the error, we use the Mean Absolute Error (MAE):

$$MAE = \frac{1}{n} \sum_{i=1}^n |\hat{y}_i - y_i|. \quad (4)$$

This metric was chosen because it is more robust to outliers than RMSE.

4. Using the training sample, we estimate the ARIMA model for the current target variable with an exogenous variable – one of the constructed indices as an indicator of the financial environment. We again select the model parameters based on the Schwarz information criterion.

5. We repeat the same procedure with calculating forecasts and retraining the model on a sliding window. Finally, we calculate the average forecast error.

Next, we only have to compare the model without the index and the one with the index taken as an exogenous variable for the resulting values of the information criterion and the average forecast error. We carry out this procedure for two different situations:

- The current index value is used as an exogenous variable (nowcasting). With this problem formulation, we try to understand whether the indicators of the financial environment explain the current state of the economy.
- The lagged index value is used as an exogenous variable (forecasting). With this problem formulation, we try to understand whether the financial indicators help predict the future values of the target variable.

Before building the model, we normalised all the variables to a range of 0 to 100 so that the error made in the prediction of various variables would be comparable.

To compare the indices in terms of their predictive power, we calculate to what extent on average the forecasts in the expanded model with the index included are better than the forecasts in the model without the index. We rank all the indices presented in the sample from the strongest to the weakest comparing the average change in MAE when the index is added to the model.

A more rigorous testing of the hypothesis of better predictions is possible using the Diebold-Mariano test (Diebold and Mariano, 2002), which makes it possible to compare the quality of time series forecasts obtained by two predictive models. We will conduct it for each pair of models (with and without the index).

Similarly to Gareev (2020), given the rather modest size of the test samples (no more than 25 observations), we will use the Diebold-Mariano statistics adjusted for small samples to test the hypothesis (see Harvey et al., 1997). The zero hypothesis will be that the two models have the same prediction quality. The hypothesis is tested against a two-sided alternative as taking the index into account can either improve or degrade the forecasts.

7.1. Nowcasting

We use the current index value as an exogenous variable in the regression for the target macroeconomic variable. This allows us to understand whether the indices that we have constructed explain the current state of the economy.

The estimation results of the models are shown in Figure 7. The columns feature various variables that we tried to predict, and the lines feature the indices that we used for the prediction. The difference between MAE in the model without the index and MAE in the model with the index is at the intersection. If the addition of the index did not improve the model during the test period, the cell remains empty. One may notice that the expanded model is quite often preferable to the parsimonious one. This occurs in 98 cases out of 198.

Ten percent of the largest improvements in MAE are attributable to nowcasts of the dollar exchange rate, industrial production index, retail turnover, RTS index and household deposits.

Figure 7. Improvement in the MAE when using the current index value as an explanatory variable

1 – RTS, the last closing price for the month; 2 – RTS, the average closing price of the index during the month; 3 – RTS, the average trade volume for the month; 4 – USD/RUB exchange rate, the average value for the month; 5 – EUR/RUB exchange rate, the average value for the month; 6 – Household deposits in roubles; 7 – Consumer price index; 8 – Industrial production index; 9 – Unemployment rate; 10 – Real wages; 11 – Average nominal wages; 12 – Index of real retail turnover; 13 – Retail trade turnover (current prices, billion roubles); 14 – Investments in fixed assets (current prices, billion roubles); 15 – Imports (billion dollars); 16 – Number of mortgage loans issued; 17 – Volume of mortgage loans issued in roubles; 18 – Debt on loans issued

Figure 8. Diebold-Mariano tests for nowcasting

1 – RTS, the last closing price for the month; 2 – RTS, the average closing price of the index during the month; 3 – RTS, the average trade volume for the month; 4 – USD/RUB exchange rate, the average value for the month; 5 – EUR/RUB exchange rate, the average value for the month; 6 – Household deposits in roubles; 7 – Consumer price index; 8 – Industrial production index; 9 – Unemployment rate; 10 – Real wages; 11 – Average nominal wages; 12 – Index of real retail turnover; 13 – Retail trade turnover (current prices, billion roubles); 14 – Investments in fixed assets (current prices, billion roubles); 15 – Imports (billion dollars); 16 – Number of mortgage loans issued; 17 – Volume of mortgage loans issued in roubles; 18 – Debt on loans issued

Let us calculate the average change in the forecast quality for each line from the resulting matrix. The average is obtained across all values, that is even for those values where the model quality deteriorates after adding the index. As a result, we rank the indices by their explanatory power. The calculation results are shown in Table 2.

Figure 8 presents the results of the Diebold-Mariano tests for different models. In each test, we compare the quality of the model without the index and that with the index. If the p-value is higher than 0.05, the difference between the two models is statistically insignificant, and the respective cells in the figure are left blank. Additionally, we ranked the indices only for those models where the quality increase turned out to be significant (see Table 3).

7.2. Forecasting

We use the lagged index value as an exogenous variable in the regression for the target macroeconomic variable. The estimation results for the models are shown in Figure 9. The expanded model is preferable to the parsimonious model in 82 cases out of 198. It is noticeable that the forecast quality improves by smaller values, by 0.085 on average, compared to nowcasting case. Where the current index value is used, the average improvement is 0.132.

Ten percent of the most significant improvements take place in forecasting exchange rates, the industrial production index, unemployment, and retail turnover. For each line, we calculate the average change in the forecast error and rank the indices by their predictive power (see Table 2).

Table 2. Ranking of indices by their explanatory and predictive power (average change in MAE when adding the index as an explanatory variable)

Explanatory power of indices		Predictive power of indices	
Index	Forecast error reduction	Index	Forecast error reduction
Search Index (PCA)	0.212	Frequency News Index (corr)	0.046
Search Index (corr)	0.157	Search Index (corr)	0.016
Frequency News Index (share)	0.124	Search Index (PCA)	0.005
Frequency News Index (corr)	0.083	Frequency Sentiment Index	-0.006
Bayesian Sentiment Index	0.042	CSI	-0.02
CSI	0.004	Sentiment News Index	-0.025
Frequency Sentiment Index	0.003	Frequency News Index (share)	-0.036
BCI	-0.05	Bayesian Sentiment Index	-0.169
Sentiment News Index	-0.066	BCI	-0.188
Manufacturer PMI	-0.082	Manufacturer PMI	-0.198
Services PMI	-0.399	Services PMI	-0.319

Source: author's calculations

Also, by analogy, we conduct the Diebold-Mariano tests for different models and additionally rank the indices by significant results only. The p-values obtained from testing are shown in Figure 10.

Figure 9. Improvement in the MAE when using the lagged index value as an explanatory variable

1 – RTS, the last closing price for the month; 2 – RTS, the average closing price of the index during the month; 3 – RTS, the average trade volume for the month; 4 – USD/RUB exchange rate, the average value for the month; 5 – EUR/RUB exchange rate, the average value for the month; 6 – Household deposits in roubles; 7 – Consumer price index; 8 – Industrial production index; 9 – Unemployment rate; 10 – Real wages; 11 – Average nominal wages; 12 – Index of real retail turnover; 13 – Retail trade turnover (current prices, billion roubles); 14 – Investments in fixed assets (current prices, billion roubles); 15 – Imports (billion dollars); 16 – Number of mortgage loans issued; 17 – Volume of mortgage loans issued in roubles; 18 – Debt on loans issued

Figure 10. Diebold-Mariano tests for forecasting

1 – RTS, the last closing price for the month; 2 – RTS, the average closing price of the index during the month; 3 – RTS, the average trade volume for the month; 4 – USD/RUB exchange rate, the average value for the month; 5 – EUR/RUB exchange rate, the average value for the month; 6 – Household deposits in roubles; 7 – Consumer price index; 8 – Industrial production index; 9 – Unemployment rate; 10 – Real wages; 11 – Average nominal wages; 12 – Index of real retail turnover; 13 – Retail trade turnover (current prices, billion roubles); 14 – Investments in fixed assets (current prices, billion roubles); 15 – Imports (billion dollars); 16 – Number of mortgage loans issued; 17 – Volume of mortgage loans issued in roubles; 18 – Debt on loans issued

7.3. Ranking of indices

When we try to explain the target variable with the current index value, the search indices show the highest forecast quality increase. Behind them are the news indices. The only ‘manual’ index, which does not lead to a deterioration of the forecast on average, is the CSI (see Table 2).

When we use the lagged index value for forecasting, the forecast quality is most strongly improved by the news frequency index constructed through the methodology of Stolbov (2011), followed by the search indices. All other indices on average worsen the forecast quality for the selected set of target variables. PMI indices perform worse than others in both cases. When we rank the indices only for significant changes in metrics, we get a similar picture (see Table 3).

Table 3. Ranking of indices by their explanatory and predictive power, taking into account the Diebold-Mariano test (average change in MAE when adding the index as an explanatory variable only for the statistically significant results)

Explanatory power of indices		Predictive power of indices	
Index	Forecast error reduction	Index	Forecast error reduction
Search Index (PCA)	0.80	Frequency News Index (corr)	0.42
Search Index (corr)	0.36	Search Index (PCA)	0.18
CSI	0.12	Sentiment News Index	0
Sentiment News Index	0.01	Frequency Sentiment Index	0
BCI	0	Services PMI	-0.01
Manufacturer PMI	0	Search Index (corr)	-0.01
Frequency News Index (corr)	-0.08	CSI	-0.02
Bayesian Sentiment Index	-0.12	Bayesian Sentiment Index	-0.12
Frequency News Index (share)	-0.39	Manufacturer PMI	-0.21
Frequency Sentiment Index	-0.47	Frequency News Index (share)	-0.23
Services PMI	-0.67	BCI	-0.95

Source: author's calculations

It turns out that the search and news indices are best used to describe the overall state of the economy. At the same time, various indices are good for working with individual macroeconomic indicators, e.g. the PMI performs much better than other indices when forecasting the industrial production index. Apparently, the PMI methodology makes the index well suited for forecasting this particular macroeconomic variable. We set up the search and news indices so as to characterise the state of the economy ‘on average’; however, these indices can be freely customised for solving specific problems by selecting the appropriate descriptors.

It is also worth noting that among the ‘manual’ indices CSI is the best. However, it is inferior to the search index in terms of forecast improvement.

If the forecast quality measure is changed from MAE to RMSE, there will be no significant changes in the results: the news and search indices still take the first place in the ranking (see Appendix, Figures 11 and 12).

8. Conclusion

In this paper, we constructed several low-frequency (monthly) indices reflecting the state of the economy based on the automatic processing of news stream, search queries, and comments in social media.

The comparison of the dynamics of 'automatic' indices with 'manual' ones obtained through consumer and business surveys shows that both types of indices capture similar trends in public sentiments. However, the search and news indices turn out to be Granger-causative of 'manual' indices. This is due to the fact that people form their opinions based on the stream of information from mass media and the Internet and then express their opinions in surveys. Building indices on an 'automatic' basis makes it possible to turn to the initial sources of information, bypassing the survey stage, and be the first to learn about changes in economic agents' sentiments.

The evaluation of ARIMA models shows that 'automatic' indices make on average a more significant contribution to improving the forecast and explaining the current state of the target variables. Among the indices presented in this paper, the search and news indices turn out to be the most informative. Given this, we can conclude that the search and news indices best describe the state of the economy 'on average'.

A significant advantage of 'automatic' indices is that they can be constructed without regular human labour and high resource costs, and at a higher frequency (daily or weekly) if needed. In this respect, a promising area for further research is the construction and analysis of high-frequency indicators of the financial environment.

Appendix is available at
[http://rjmf.econs.online/en;](http://rjmf.econs.online/en;dx.doi.org/10.31477/rjmf.202004.75)
dx.doi.org/10.31477/rjmf.202004.75

References

- Baker, S. R., Bloom, N. and Davis, S. J.** (2016). Measuring Economic Policy Uncertainty. *Quarterly Journal of Economics*, 131(4), pp. 1593–1636. doi: 10.1093/qje/qjw024
- Berg, L. and Bergstrom, R.** (1996). *Consumer Confidence and Consumption in Sweden*. Department of Economics Uppsala University Working Paper, N 7.
- Blanchard, O.** (1993). Consumption and the Recession of 1990–1991. *American Economic Review*, 83(2), pp. 270–274.
- Bollen, J., Mao, H. and Zeng, X.** (2011). Twitter Mood Predicts the Stock Market. *Journal of Computational Science*, 2(1), pp. 1–8. doi: 10.1016/j.jocs.2010.12.007
- Bram, J. and Ludvigson, S. C.** (1998). Does Consumer Confidence Forecast Household Expenditure? A Sentiment Index Horse Race. *Economic Policy Review*, 4(2), pp. 59–78.
- Choi, H. and Varian, H.** (2009). *Predicting Initial Claims for Unemployment Benefits*. Google Inc. Technical Report.
- Choi, H. and Varian, H.** (2012). Predicting the Present with Google Trends. *Economic Record*, 88(s1), pp. 2–9. doi: 10.1111/j.1475-4932.2012.00809.x
- Da, Z., Engelberg, J. and Gao P.** (2011). In Search of Attention. *Journal of Finance*, 66(5), pp. 1461–1499. doi: 10.1111/j.1540-6261.2011.01679.x

- D'Amuri, F.** (2009). *Predicting Unemployment in Short Samples with Internet Job Search Query Data*. MPRA Paper from University Library of Munich, N 18403.
- Davidson-Pilon, C.** (2015). *Bayesian Methods for Hackers: Probabilistic Programming and Bayesian Inference*. Addison-Wesley Professional.
- Diebold, F. X. and Mariano, R. S.** (2002). Comparing Predictive Accuracy. *Journal of Business and Economic Statistics*, 20(1), pp. 134–144. doi: 10.1198/073500102753410444
- Frisbee, B.** (2010). The Predictive Power of Financial Blogs. *Senior Economics Thesis*, Haverford College.
- Gareev, M.** (2020). Use of Machine Learning Methods to Forecast Investment in Russia. *Russian Journal of Money and Finance*, 79(1), pp. 35–56. doi: 10.31477/rjmf.202001.35
- Ginsberg, J., Mohebbi, M. H., Patel, R. S., Brammer, L., Smolinski, M. S. and Brilliant L.** (2009). Detecting Influenza Epidemics Using Search Engine Query Data. *Nature*, 457(7232), pp. 1012–1014. doi: 10.1038/nature07634
- Goel, S., Hofman, J. M., Lahaie, S., Pennock, D. M. and Watts, D. J.** (2010). Predicting Consumer Behavior with Web Search. *Proceedings of the National Academy of Sciences*, 107(41), pp. 17486–17490. doi: 10.1073/pnas.1005962107
- Goloshchapova, I. and Andreev, M.** (2017). Measuring Inflation Expectations of the Russian Population with the Help of Machine Learning. *Voprosy Ekonomiki*, 6, pp. 71–93. [In Russian]. doi: 10.32609/0042-8736-2017-6-71-93
- Harvey, D., Leybourne, S. and Newbold, P.** (1997). Testing the Equality of Prediction Mean Squared Errors. *International Journal of Forecasting*, 13(2), pp. 281–291.
- Li, Q., Feng, W. and Quan, Y. H.** (2020). Trend and Forecasting of the COVID-19 Outbreak in China. *Journal of Infection*, 80(4), pp. 469–496. doi: 10.1016/j.jinf.2020.02.014
- Mao, Y., Wei, W., Wang, B. and Liu, B.** (2012). Correlating S&P 500 Stocks with Twitter Data. In: *HotSocial '12: Proceedings of the First ACM International Workshop on Hot Topics on Interdisciplinary Social Networks Research*. New York: Association for Computing Machinery, pp. 69–72. doi: 10.1145/2392622.2392634
- Matsusaka, J. G. and Sbordone, A. M.** (1995). Consumer Confidence and Economic Fluctuations. *Economic Inquiry*, 33(2), pp. 296–318.
- McLaren, N. and Shanbhogue, R.** (2011). Using Internet Search Data as Economic Indicators. *Bank of England Quarterly Bulletin*, 51(2), pp. 134–140.
- Rogers, A., Romanov, A., Rumshisky, A., Volkova, S., Gronas, M. and Gribov, A.** (2018). RuSentiment: An Enriched Sentiment Analysis Dataset for Social Media in Russian. In: E.M. Bender, L. Derczynski, and P. Isabelle, eds. *Proceedings of the 27th International Conference on Computational Linguistics*. Stroudsburg: Association for Computational Linguistics, pp. 755–763.
- Stolbov, M.** (2011). Statistics of Search Queries in Google as an Indicator of Financial Conditions. *Voprosy Ekonomiki*, 11, pp. 79–93. [In Russian]. doi: 10.32609/0042-8736-2011-11-79-93
- Utaka, A.** (2003). Confidence and the Real Economy – the Japanese Case. *Applied Economics*, 35(3), pp. 337–342.
- Yakovleva, K.** (2018). Text Mining-based Economic Activity Estimation. *Russian Journal of Money and Finance*, 77(4), pp. 26–41. doi: 10.31477/rjmf.201804.26