Elizaveta A. Golovanova
E-mail: golovanova-ea@ranepa.ru

Andrei V. Zubarev
E-mail: zubarev@ranepa.ru

Russian Presidential Academy of National Economy and Public Administration
Address: 82, build. 1, Prospect Vernadskogo, Moscow 119571, Russia

Abstract

Text analysis with machine learning support can be implemented for studying experts’ relations to the Bank of Russia. To reach macroeconomic goals, the communication policy of the bank must be predictable and trustworthy. Surveys addressing this theme are still insufficient compare to the theoretical studies on the subject of other bank tools. The goal of this research is to analyze the perception of uncertainty by economic agents. For that purpose, we built an uncertainty indicator based on news sources from the Internet and on textual analysis. The dynamics of the indicator reflect unexpected statements of the Bank of Russia and events affecting monetary policy. Financial theory links monetary policy and stock prices, so we used this fact to examine the impact of the uncertainty indicator on the MOEX and RTS indices. We tested the hypothesis that our indicator is significant in GARCH models for chosen financial series. We found out several specifications in which our indicator is significant. Among the specifications considered, the uncertainty indicator contributes the most to explaining variances of the RTS index. The obtained uncertainty indicator can be used for forecasting of different macroeconomic variables.

Key words: uncertainty; Bank of Russia; news sources; data analysis; machine learning; word cloud; stock index.

Introduction

On the one hand, expectations of economic agents are one of the policy references for central banks, but on the other hand they also influence the effectiveness of their policy. Such expectations can be traced following the opinions of experts published in the mass media on the issue of central bank policy. That news can reflect the perception by economic agents of certain central bank policy measures, as well as influence this perception [1, 2]. In this study, we are making an attempt to quantify these expectations.

Machine learning methods help to process large amounts of information, significantly simplify textual analysis, and allow us to get transparent results in aggregated form. Nowadays, machine learning methods for news analysis have a wide range of applications and can be found in a variety of areas. Machine learning is used to analyze the texts of news sources to predict an election victory [3] and to detect fake news [4]. Comments analysis from financial microblogs and Twitter is used to predict the volatility of securities [5]. The tourism sector, where continuous improvement of service is required (e.g. the restaurant and hotel business), actively uses the analysis of comments and reviews about their businesses on websites [6, 7].

Analysis of news source texts can be useful in the context of reviewing research on the policies of central banks. Blinder et al. [8] have shown that the Central Bank’s communication policy is a powerful tool, as it can improve the predictability of monetary policy and has the potential to achieve macroeconomic goals such as low and stable inflation.

By processing information from news sources using machine learning methods, it is possible to estimate the level of uncertainty in the expectations of economic agents at each moment. For example, uncertainty in macroeconomic news can have a negative impact on financial markets [1, 2]. However, there is empirical evidence that in some countries monetary authorities are responding to investor sentiment. For example, the study [9] states that the Reserve Bank of Australia lowers the interest rate in response to a higher level of uncertainty amongst experts from the Shadow Board1.

Hansen and McMahon [10] investigated how news released by the Federal Open Market Committee influence market and real economic variables. In this study, 76 macroeconomic variables were examined. Using FAVAR models, the authors found out that shocks to forward guidance are more important than the FOMC communication of current economic conditions in terms of their effects on certain market variables, such as stock indices, and macroeconomic variables, such as unemployment and CPI.

Cieslak and Schrimpf [11] calculated the importance of non-monetary news of Central Banks reports. The authors found a significant difference in the news content depending on the communication channel used by central banks. According to their estimates, non-monetary news prevails in about 40% of FED and ECB statements, and in the context of press conference news this share is especially high.

Despite central bank communication becoming an integral component of the set of tools for conducting monetary policy, the range of research on this topic is still quite limited compared to other macroeconomic policy tools. The goal of this study is to describe the economic agents’ perception of uncertainty in relation to Bank of Russia policy, and analyze its impact on financial market indicators.

---

1 The Shadow Board brings together professional macroeconomists who make recommendations for interest rates changes in the week before each meeting of the actual Reserve Bank Board.
To achieve this goal, we construct an indicator of uncertainty in relation to the policy of the Bank of Russia based on news from the Internet and test the hypothesis of the indicator’s significance for explaining the volatility of Moscow Stock Exchange indices returns.

1. Construction of an uncertainty indicator

In this study, we considered the news from leading Russian media writing about the economy: RBC, Gazeta.Ru, Finmarket, TASS and Kommersant. The data covers the period from 01.01.2014 to 31.05.2020. The following words and phrases were used as queries for the archival search: “Bank of Russia,” “Nabiullina,” “CB of RF.” These requests were formed based on their popularity and direct relationship to the Bank of Russia. The TASS news could not be searched by archives, so it was searched by headlines. Accordingly, the list of keywords was expanded with the following words: “Central Bank,” “rouble exchange rate,” “Yudaeva,” “Skorobogatova,” “Tulin,” “Shvetsov,” “vice chairman of CB.” Headlines were searched only for the “politics” and “economics” categories. After processing all sources, 22,156 articles were received. Figure 1 shows the monthly dynamics of the number of all news items.

Figure 2 shows a diagram of the number of articles in each source. The leading position is occupied by RBC, and the lowest — by TASS. Data are unequally distributed by sources, and this is taken into account when constructing the uncertainty indicator.

To build the uncertainty indicator, it is necessary to find out what topics are contained in the collected corpus of texts in order to select the ones that directly relate to the economic policy of the Bank of Russia. Before this, the data needs to be pre-processed. First, we split each text into a list of words and symbols (tokenization). Then we reduced all words to lowercase so that “CB” and “cb” are not considered as different words. To decrease the variety of words, lemmatization is required: we put all nouns in the nominative case and made them singular, all adjectives in the masculine gender and made them singular as well, put verbs in the infinitive form, and so on. The Python 3 package pymorphy was used for this purpose. Punctuation marks and numbers have been removed from the texts, because they do not contain any useful information without context. Also, conjunctions, prepositions, and particles that are quite common in the text, but useless separately from it, were excluded from the list of words. With such text processing, it is assumed that the word order in the text does not change...
its sentiment. This method is named a “bag of words.” Each text represented a vector showing how many times each word from the dictionary appears in it. The resulting number of unique words was 269,611. However, this list contained words that occur only once in the text, or, conversely, too often. Therefore, we filtered rare and frequently occurring words, assuming that the word should appear in the total array of texts more than three times, but not more often than in 40% of the texts. The resulting number of unique words after filtering was 52,073.

Afterwards the resulting list of words was divided into thematic lists in order to visualise them using the word cloud and to determine what issues are covered in articles for the queries selected. It is worth noting that each text from the text corpus can have several topics, despite the fact that they were received by queries using the same keywords. Thus, it would be incorrect to assign each text to one specific topic, and it is necessary to use a different approach.

A hierarchical Bayesian model was constructed to identify topics found in texts. On the first level we set a prior parameter that determines the number of the themes splitting ($T$). On the second level there is a multinomial variable with a prior Dirichlet distribution which determines the probability of a word’s relation to a predetermined theme in the document. This model is also called the latent Dirichlet allocation (LDA). The formal definition is as follows: a dictionary is a list of words that are pre-filtered by the frequency of occurrence in the text $\{1, \ldots, V\}$. Each word is a vector $w, w_i \{1, \ldots, V\}$ where exactly one component is equal to 1. Each text is a sequence of $N$ words $w$. We consider a corpus of $M$ texts $D = \{w_d \mid d = 1..M\}$.

Assume that the number of topics $T$ is set exogenously. Each document has a distribution of topics within it $\theta \sim p(T|d)$. Next, the probability that a word appears in the document is calculated. One of the topics is selected randomly. Each word is included in the selected document based on the word distribution $\beta \sim p(w|T)$.

Then we form $T$ hypotheses that the word $w$ in document $d$ if it belongs to topic $t_1$ or topic $t_2$, and so on up to topic $T$. The total probability of the word appearing in the document can be calculated using the formula:

$$D_i(w|T) = \sum_{t \in T} p(w|T) \cdot p(T|d) = \sum_{t \in T} \theta \cdot \beta.$$  

Using the latent Dirichlet allocation algorithm, all words were divided into seven top-

---

2 A document is a sequence of events in the multinomial model. Each event is considered as a random selection of one word from the “bag of words”
ics\(^3\): money and payment systems, financial regulation, social sphere, international relations, stock market, economic forecasts, and property transactions (Table 1). Several iterations were made (different values of \( T \) were tried) until clusters of words were formed that could be unambiguously divided into topics. The names of these topics were set post factum based on these lists of words. It is also interesting to look at what percentage of topics each source contains. To do this, we recalculate separately the words within each source that relate to each topic. The results are shown in Figure 3. It can be seen that economic forecasts take up a significant part in each source (13–18\%).

### Table 1. List of topic names and keywords

<table>
<thead>
<tr>
<th>Topic name</th>
<th>Keywords</th>
</tr>
</thead>
<tbody>
<tr>
<td>Money and payment systems</td>
<td>stock, asset, investor, bargain, security, bond, million, investment, instrument, sberbank, large, cryptocurrency, sale, capital, project, shareholder, valuable, gold, exchange, group</td>
</tr>
<tr>
<td>Financial regulation</td>
<td>system, organization, instrument, banking, customer, million, credit, report, law, information, license, decision, amount, operation, claim, card, activity, regulator, number, case</td>
</tr>
<tr>
<td>Social sphere</td>
<td>person, project, thousand, job, region, business, work, nizhegorodsky, most, new, development, city, region, other, million, center, very, country, money, place</td>
</tr>
<tr>
<td>International relations</td>
<td>president, country, sanction, government, head, putin, announce, vladimir, power, question, minister, attitude, against, economic, say, council, state, ukraine, federation, word</td>
</tr>
<tr>
<td>Stock market</td>
<td>currency, dollar, oil, week, foreign exchange, level, analyst, rate, decline, barrel, index, investor, fall, petroleum, quote, mark, American, factor, expect, country</td>
</tr>
<tr>
<td>Economic forecasts</td>
<td>rate, economy, inflation, level, decline, economic, forecast, key, head, nabiullin, policy, consider, regulator, increase, declare, say, decision, situation, elvira, estimate</td>
</tr>
<tr>
<td>Property transactions</td>
<td>loan, rate, credit, mortgage, thousand, hypothecate, million, borrower, lending, program, business, region, income, client, make up, housing, amount, condition, real estate, oblast</td>
</tr>
</tbody>
</table>

\(^3\) The gensim library is used to implement the latent Dirichlet allocation algorithm in Python
\(^4\) The table shows the first 20 words generated by the LDA algorithm for each topic
\(^5\) As the texts are in Russian, all words on the picture are nothing else but the translation.
“cast,” “risk,” “scenario” or “expectation,” as well as their derivatives were eliminated. Further, texts that do not contain the word “economy” and its derivatives were excluded from the resulting text array. Finally, texts that are not directly related to the Bank of Russia’s policy were eliminated, according to the words “income,” “inflation,” “rate” and their derivatives. It should be noted that the order of filtering texts by keywords does not matter, but selected texts have to contain at least one word from all three groups of keywords. According to the results of the dropout, 4,691 texts remained, which is about 21% of the total text corpus.

In addition, in the work of Baker, Bloom, and Davis [12] a certain correction controlling the number of articles in each source was made. Authors divided the number of articles with keywords by the total number of articles for the period and for each source. Then, within
each source standardisation was made. In other words, the mean value was subtracted, and the series was divided by the standard deviation. Thereafter, the data was combined into a single array and was summed up over the period. Finally, the array was scaled so that the minimum value is zero.

Since the distribution of articles by source was unequal in the afore-mentioned study, we applied a similar correction. It is important to take this correction into account, because each source reviews the news with different frequency and, furthermore, may have a biased attitude to some issues related to the policy of the Bank of Russia, as well as may cover this topic with different intensity.

The indicator does not reflect the attitude to the Central Bank of Russia in terms of “good” and “bad,” but shows an increase of the “discussion” around the Bank of Russia’s policy in the context of a certain set of topics related to uncertainty. Thereby it reflects an increase or decrease in the interest in monetary policy. Thus, the constructed indicator reflects the perception of uncertainty in the Bank of Russia’s policy by the expert community. It is worth mentioning that this perception is influenced not only by direct actions of the Central Bank, but also by various economic and political shocks, such as sanctions that were imposed on Russia more than once during the period under consideration. The resulting indicator is shown in Figure 5.

In Figure 5 we can observe three periods with an increase in the value of indicator: in 2015, 2016, and 2019. The first increase occurred on January 2015, when the Bank of Russia unexpectedly lowered the key rate [13] after its sharp increase in December 2014 [14], when strict measures were undertaken to stabilise the situation in the foreign exchange market, which arose due to a combination of negative economic factors (oil prices, sanctions, speculative component in the financial market). At the same time, the increase of the uncertainty indicator was observed throughout the second half of 2014. That corresponds to the dynamics of the exchange rate and the overall increase in uncertainty. In March 2015, the Bank of Russia met the market’s expectations for monetary policy easing [15], but again we can see the peak of the indicator, which corresponds to the animated debates about the regulator’s further actions. At the end of June 2015, inflation decreased dramatically, and the Central Bank lowered the key rate [16]. This was completely at
odds with the forecasts of experts and the comments of the Bank of Russia. At that moment, we can observe high values of the uncertainty indicator, comparable to the peak in March. In September, October and December 2015, there was much discussion about the destabilization of the situation in the foreign exchange market due to the rise in oil prices and the key rate being unchanged. Further, rate constancy was predictable in September and October, but not in December, when the Bank of Russia promised to reduce the key rate by 0.5 pp. [17]. As a result, we can observe a new peak of the indicator in December 2015, which is slightly larger than the values in June and March.

The Central Bank of the Russian Federation also refused to lower the key rate in March 2016, which was in line with the forecasts of many analysts [18]. In June [19] and September [20] 2016, economic sanctions were strengthened, and the key rate was lowered, which led to another increase in uncertainty about the policy of the Bank of Russia.

Since September 2018, the Central Bank had started raising the key rate [21]. By the end of that year, the Bank of Russia announced the resumption of currency purchases from January 2019 and a new increase in the key rate [22], which also came as a surprise to many experts and reflected the growth of the uncertainty indicator. In June 2019, an economic forum took place in St. Petersburg (accompanied by a number of negative political news, such as the ongoing imprisonment of Michael Calvey). The Bank of Russia also lowered the key rate for the first time after its gradual increase since September 2018 [23]. At that moment, there was a sharp increase in the dynamics of the indicator, which corresponds to a higher degree of uncertainty in the economy, that is also transferred to uncertainty in economic policy. The Bank of Russia also commented that the key rate was likely to be lowered during the next meeting, which happened in July and did not cause a surge of uncertainty. In September, the rate was lowered for the third time during the year [24], but this was associated with a new round of economic sanctions, which eventually led to another peak of the indicator.

The current situation differs significantly from the considered peaks of the uncertainty indicator. The coronavirus pandemic has led to a global economic crisis. The decline in economic activity and the downward shift in the aggregate demand curve also affected oil prices, which led to a decrease in the rouble exchange rate. However, maintaining the exchange rate in this case does not make sense precisely because of the nature of the global shock, against which economic policy should be addressed. In this regard, many experts expected softening (or at least not tightening) of monetary policy in order to stimulate aggregate demand. Thus, on April 24, 2020, the Bank of Russia decided to reduce the key rate by 0.5 pp. It is clear that although this decision slightly increased the level of uncertainty, the indicator stayed at a moderate level. In early May, uncertainty began to subside due to the strengthening of the rouble and rising oil prices. However, by the end of May, it was actively discussed that the Central Bank was going to significantly reduce the key rate in June [25]. As a result, dynamics of the constructed indicator reflect the events that took place in the economy; its high values correspond to the increase of uncertainty in terms of perception of the Bank of Russia’s policy and falls within the economic logic.

2. The models with the indicator of uncertainty

Uncertainty in monetary policy can affect the volatility of the exchange rate, which is one of the factors of stock price movement in the framework of financial theory as extended model of the CAPM. In this regard, it is interesting to analyze the impact of the constructed indicator on the dynamics of market indices.

We considered models with MOEX and RTS stock indices for the period from 01.01.2014 to 31.05.2020. The hypothesis of significance of the
constructed indicator in terms of the impact on the variance of the financial market was tested using GARCH models. To take into account more information, we used the two-week frequency indicator (instead of the monthly version). Figure 6 shows that the two-week uncertainty indicator generally has dynamics similar to its monthly counterpart. However, the considered peaks are sometimes higher, since they contain information on the events described above and are not smoothed out by other observations within a month.

Firstly, models without the indicator were selected to describe the volatility of each index, then the indicator was added to the indices’ variance equations. The GARCH(\(\alpha, \beta\)) process with the mean ARMA(\(p, q\)) equation is generally represented as follows:

\[
y_t = c + \sum_{j=1}^{k} p_j y_{t-j} + \sum_{j=1}^{l} q_j \epsilon_{t-j},
\]

\[
\epsilon_t = \sqrt{h_t} \cdot \nu_t, \nu_t \sim \text{i.i.d. } N(0,1),
\]

\[
h_t = \omega + \sum_{j=1}^{m} \alpha_j \epsilon_{t-j}^2 + \sum_{j=1}^{r} \beta_j h_{t-j},
\]

where \(h_t\) — conditional variance;
\(c\) — constant in the mean equation;
\(\omega\) — constant in the variance equation;
\(p_j\) — lags coefficients in the mean equation;
\(q_j\) — residuals lags coefficients in the mean equation;
\(\alpha_j\) — lags coefficients in the variance equation;
\(\beta_j\) — residuals lags coefficients in the variance equation;
\(y_t\) — current value of the series.

Pearson’s chi-squared test was used to analyze whether the selected distributions fit the innovations in models with the uncertainty indicator. To test this hypothesis, the sample is divided into several intervals. Let \(n_i\) be the number of elements that fall into the \(i\)-th interval, and the probability of a random variable falling into the \(i\)-th interval is \(p_i\). The deviation of the sample distribution from the theoretical one is determined by the formula:

\[
\sum_{i=1}^{k} \frac{(n_i - np_i)^2}{np_i}.
\]

The sum has an asymptotic \(\chi^2\) distribution with degree of freedom \(f = k - c - 1\), where \(c\) is the number of model parameters determined from the sample. If \(\hat{\chi}^2\) is less than the \(\chi^2_{c,f}(f)\) from the table of critical distribution values of \(\chi^2\) distribution, then the null hypothesis is not rejected.

![Fig. 6. Dynamics of the two-week uncertainty indicator](image-url)
Empirically, the GARCH(1,1) model was selected for the MOEX stock index with the ARMA(1,1) mean equation and the skewed normal distribution of innovations, i.e. $v_t \sim \text{i.i.d.} \mathcal{SN}(0,1)$. A comparison of the normal distribution and the skewed normal distribution of standardized innovations is illustrated in Figure 7. It is noticeable that the distribution of innovations is asymmetric, that should be taken into account when estimating the model.

To confirm the selected distribution of innovations, the Pearson’s test was used (Table 2).

### Table 2.

<table>
<thead>
<tr>
<th>Groups</th>
<th>Statistics</th>
<th>$p$-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>20</td>
<td>0.2867</td>
</tr>
<tr>
<td>2</td>
<td>30</td>
<td>0.5179</td>
</tr>
<tr>
<td>3</td>
<td>40</td>
<td>0.6733</td>
</tr>
<tr>
<td>4</td>
<td>50</td>
<td>0.8725</td>
</tr>
</tbody>
</table>

The $p$-values are always higher than the significance level $\alpha = 0.05$, so we can conclude that the distribution of innovations for this model with addition of the uncertainty indicator is selected adequately. The results of GARCH models for the MOEX stock index are presented in Table 3.

GARCH(1,1) models with ARMA(1,1) and skewed normal distribution of innovations were also chosen to describe the volatility of the RTS stock index on two-week data. Figure 8 shows that the distribution of innovation is also skewed to the left.

The selected model with the indicator was also tested to confirm the selected distribution of innovations according to the Pearson’s test. The results are presented in Table 4.

The $p$-value is always higher than the significance level $\alpha = 0.05$, so the distribution of innovations in the model for RTS with the addition of the uncertainty indicator is selected adequately, according to the Pearson test. The results of GARCH models for the RTS stock index are shown in Table 5.
### Table 3.
Comparison of GARCH models for the MOEX index with and without the uncertainty indicator on two-week data

<table>
<thead>
<tr>
<th>Coefficients</th>
<th>Model with uncertainty indicator</th>
<th>Model without uncertainty indicator</th>
</tr>
</thead>
<tbody>
<tr>
<td>$c$</td>
<td>0.0046*** (–0.0002)</td>
<td>0.0046*** (–0.0001)</td>
</tr>
<tr>
<td>$p$</td>
<td>0.8545*** (–0.039)</td>
<td>0.8232*** (–0.0528)</td>
</tr>
<tr>
<td>$q$</td>
<td>–0.9999*** (–0.0024)</td>
<td>–1*** (–0.0023)</td>
</tr>
<tr>
<td>$\omega$</td>
<td>0 (–0.0000)</td>
<td>0.0003 (–0.0002)</td>
</tr>
<tr>
<td>$\alpha$</td>
<td>0.1275 (–0.0828)</td>
<td>0.1016** (–0.0414)</td>
</tr>
<tr>
<td>$\beta$</td>
<td>0.3643 (–0.2804)</td>
<td>0.7009*** (–0.1639)</td>
</tr>
<tr>
<td>Uncertainty indicator</td>
<td>0.0003** (–0.0001)</td>
<td></td>
</tr>
<tr>
<td>Coefficient of skewness of the error distribution</td>
<td>0.5681*** (–0.1129)</td>
<td>0.6124*** (–0.1255)</td>
</tr>
<tr>
<td>Schwartz information criterion</td>
<td>–3.7484</td>
<td>–3.7417</td>
</tr>
</tbody>
</table>

**Fig. 8.** Comparison of standardised innovations of the RTS series with normal distribution and skewed normal distribution
3. Analysis of models with the uncertainty indicator

Tables 3 and 5 show that the Schwartz information criterion decreases when the indicator is added. The coefficient of the uncertainty indicator in Table 3 has a positive sign, which indicates that the MOEX variance goes up by about 0.0003 as the uncertainty indicator increases by one. Although the uncertainty indicator is significant at 5% in this specification, its impact on the stock market is limited.

In Table 5, the information criterion also decreased when the indicator was added. The coefficient of the uncertainty indicator is significant at the level of 5% and has a positive sign. This indicates that RTS variance increases by about 0.0011 as the uncertainty indicator increases by one. This result is significantly higher than for the MOEX, which may suggest that the uncertainty indicator has a greater impact on the RTS’s currency component than on shares.

### Table 4. Results of the Pearson’s test for the RTS model with the uncertainty indicator

<table>
<thead>
<tr>
<th>Groups</th>
<th>Statistics</th>
<th>ρ-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>20</td>
<td>24.42</td>
</tr>
<tr>
<td>2</td>
<td>30</td>
<td>33.56</td>
</tr>
<tr>
<td>3</td>
<td>40</td>
<td>43.43</td>
</tr>
<tr>
<td>4</td>
<td>50</td>
<td>53.43</td>
</tr>
</tbody>
</table>

### Table 5. Comparison of GARCH models for the RTS index with and without the uncertainty indicator on two-week data

<table>
<thead>
<tr>
<th>Coefficients</th>
<th>Model with uncertainty indicator</th>
<th>Model without uncertainty indicator</th>
</tr>
</thead>
<tbody>
<tr>
<td>c</td>
<td>0.0010 (-0.0044)</td>
<td>0.0004 (-0.0053)</td>
</tr>
<tr>
<td>p</td>
<td>-0.7837*** (-0.1154)</td>
<td>-0.8065*** (-0.1104)</td>
</tr>
<tr>
<td>q</td>
<td>0.8163*** (-0.0882)</td>
<td>0.8356*** (-0.0899)</td>
</tr>
<tr>
<td>ω</td>
<td>0 (-0.0000)</td>
<td>0.0003 (-0.002)</td>
</tr>
<tr>
<td>α</td>
<td>0.0369 (-0.0784)</td>
<td>0.0748** (-0.0375)</td>
</tr>
<tr>
<td>β</td>
<td>0.2438 (-0.2455)</td>
<td>0.8529*** (-0.0667)</td>
</tr>
<tr>
<td>Uncertainty indicator</td>
<td>0.0011** (-0.0005)</td>
<td></td>
</tr>
<tr>
<td>Coefficient of skewness of the error distribution</td>
<td>0.7971*** (-0.1051)</td>
<td>0.7329*** (-0.0869)</td>
</tr>
<tr>
<td>Schwartz information criterion</td>
<td>-2.7201</td>
<td>-2.7038</td>
</tr>
</tbody>
</table>
Conclusion

In this study we constructed the uncertainty indicator that reflects the experts’ perception of the Bank of Russia’s policy. To build the indicator, more than 22,000 articles from five news sources were used. Those texts were preprocessed, and all words collected from them were divided into topics using the latent Dirichlet allocation algorithm. The words from the key topic were visualized using Word Cloud. In addition, normalization on the number of articles in the sources was made. The dynamics of the uncertainty indicator were juxtaposed with the key events and statements of the Bank of Russia.

To check the validity of the constructed indicator, we employed GARCH models that explain the volatility of the MOEX and RTS stock indices based on two-week data. The coefficients of the uncertainty indicator obtained were significant in the GARCH(1,1) models with the ARMA(1,1) mean equation and a skewed error distribution for describing both MOEX and RTS volatility.

The indicator we constructed can be used for forecasting individual macroeconomic variables and conducting monetary policy by the Bank of Russia, since it reflects the expectations of economic agents. A possible direction for future research is the construction of indicators of exchange rate and inflation expectations.

Acknowledgement

The article was written on the basis of the RANEPA state assignment research programme.

References


17. TASS (2015) The Bank of Russia kept the key rate at the level of 11%, but promised to reduce it soon. TASS, 30 October 2015. Available at: https://tass.ru/ekonomika/2393054 (accessed 1 September 2020).


25. RBC (2020) Nabiullina did not see a reason for an unplanned reduction in the key rate. RBC, 22 May 2020. Available at: https://www.rbc.ru/finances/22/05/2020/5ec78c639d79472a021822e (accessed 1 September 2020).


About the authors

Elizaveta A. Golovanova  
Junior Researcher, Institute of Applied Economic Research, Russian Presidential Academy of National Economy and Public Administration, 82, build. 1, Prospect Vernadskogo, Moscow 119571, Russia;  
MSc Student, Department of Economics, Mathematics and Informational Technology, Russian Presidential Academy of National Economy and Public Administration, 82, build. 1, Prospect Vernadskogo, Moscow 119571, Russia;  
E-mail: golovanova-ea@ranepa.ru  
ORCID: 0000-0002-9073-7553

Andrei V. Zubarev  
Cand. Sci. (Econ.);  
Senior Researcher, Institute of Applied Economic Research, Russian Presidential Academy of National Economy and Public Administration, 82, build. 1, Prospect Vernadskogo, Moscow 119571, Russia;  
Lecturer, Department of Economics, Mathematics and Informational Technology, Russian Presidential Academy of National Economy and Public Administration, 82, build. 1, Prospect Vernadskogo, Moscow 119571, Russia;  
E-mail: zubarev@ranepa.ru  
ORCID: 0000-0003-2945-5271