Development of methods for constructing high-frequency indicators of economic expectations of the population on the basis of Big Data

Aigul' N. Ardeeva

Undergraduate,
Ufa State Petroleum Technological University,
450000, 4 Pervomayskaya str., Ufa, Russian Federation;
e-mail: ardeeva.a.n@mail.ru

Ramil' R. Zakiev

Undergraduate,
Ufa State Petroleum Technological University,
450000, 4 Pervomayskaya str., Ufa, Russian Federation;
e-mail: mr.ramil98@mail.ru

Pavel V. Petrenko

Undergraduate,
Ufa State Petroleum Technological University,
450000, 4 Pervomayskaya str., Ufa, Russian Federation;
e-mail: pavelpetrenko1999.job@mail.ru

Arsen I. Khasanov

Undergraduate,
Ufa State Petroleum Technological University,
450000, 4 Pervomayskaya str., Ufa, Russian Federation;
e-mail: khasanovarsen@yandex.ru

Artur I. Arduganov

Undergraduate,
Ufa State Petroleum Technological University,
450000, 4 Pervomayskaya str., Ufa, Russian Federation;
e-mail: ardtur2@mail.ru

Abstract

The expectations of economic agents are the basis for economic theories explaining the processes of consumption, investment, money demand, inflation, etc. In the scientific literature, the economic expectations are defined as expectations of economic agents – households, firms, government and the expert community – concerning future events in the economy that significantly affect decision making. In the modern practice of economic research, there are two

key approaches to the assessment of economic expectations: the approach based on surveys of economic agents (survey-based approach) and the approach based on the analysis of exchange parameters (market-based approach). However, these approaches have a number of significant drawbacks. Since 2015, the official monetary policy regime in Russia has been inflation targeting, the transition to which was actually initiated by the regulator in September 2013. During this period, the level of inflation expectations of the population is an important object of research for the Bank of Russia. The regulator's numerical benchmark for inflation under the current monetary policy regime is intended to have a significant impact on inflation expectations of the population. The latter, in turn, can affect the price level in the long term. The development of a methodology of building indicators of economic expectations on the basis of large data sets is due to their significant influence on economic processes, and the ability to overcome key limitations of current methods for the measurement of expectations using the methods of Big Data analysis.

Introduction

We present data on the share of Russian households with Internet access, based on the official statistics of Rosstat (figure 1), statistical tests of the relationship between the intensity coefficient of inflation expectations of the Russian population based on big data and the dynamics of the median FOM for short-term numerical inflation expectations of the Russian population based on the econometric formula for calculating the initial data (table 1), the statistical tests of the relationship between the dynamics of the indicator of the intensity of inflationary expectations of the population of Russia on the basis of big data and the dynamics underlying consumer price index, based on Rosstat data by processing the original Rosstat data for 2011-2018 (table 2), the data of the test of Granger causality index of intensity of inflation expectations from different economic parameters from monthly data - in this case, we took the figures for the latest 2018 from the original Rosstat data (table 3).

Main part

The choice of media sources can be carried out by analyzing the ratings of media sources. As a basis for the selection of the largest and most important economic media in Russia, the open ratings of the company "Medialogia" were determined. This company provides open access to monthly reports on the largest and most influential citation Russian media sources, obtained on the basis of an automatic monitoring system of media in real time. The coverage of media sources for monitoring is extremely wide and amounts to more than 43 thousand media, citation of which is estimated with the help of more than 800 million social media. The rating sections include lists of the most cited newspapers, magazines and news agencies based on their representation on the Internet (analysis of traffic to official web pages and official media groups in social media). In addition, open ratings include regular identification of

the leading TV channels and radio stations, as well as the most popular blogs.
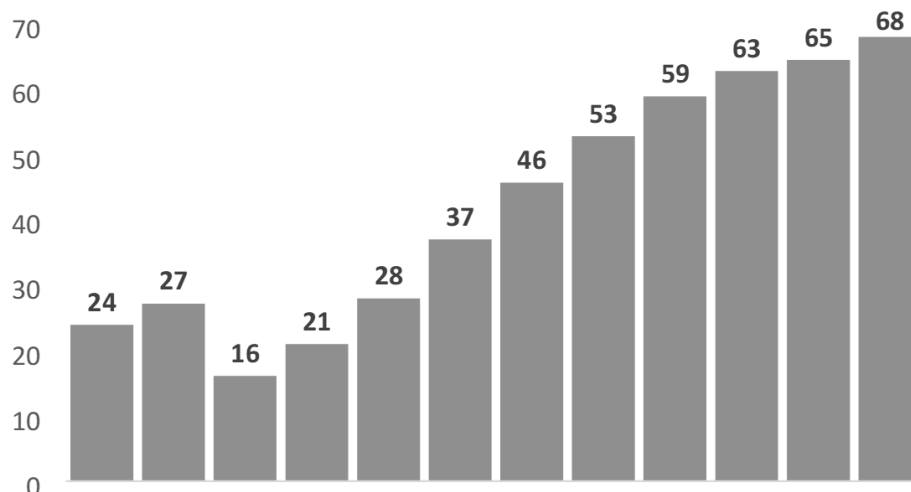

Figure 1 - Share of Russian households with Internet access

Table 1 - Statistical tests of the relationship dynamics indicator of the intensity of inflation expectations of the Russian population based on Big Data and dynamics of the FOM median indicator for short-term numerical inflation expectations of the Russian population

Statistical test of the relationship		Correlation with indicator FOM	Mutual information with indicator FOM
The intensity of inflation expectations of the Russian population on the basis of Big Data	UN sludge = 0	0.52****	0.24*
	UN sludge = 1	0.67****	0.22
	UN sludge = 2	0.72****	0.36***
	UN sludge = 3	0.70****	0.27**
	UN sludge = 4	0.69****	0.22
	UN sludge = 5	0.61****	0.24*
	UN sludge = 6	0.48****	0.16

Notes: the following designations are used for different levels of significance: * – 20%, ** – 15%, *** – 5%, **** – 1%.
Source: compiled by the author.

Table 2 - Statistical tests of the relationship dynamics indicator of the intensity of inflation expectations of the Russian population based on Big Data and the dynamics of the basic consumer price index based on Rosstat data

Statistical test of the relationship		Correlation with baseline CPI	Mutual information with basic CPI	The Granger test for the effect of CPI on the base
Intensity of inflation expectations Russian population based on Big Data	UN sludge = 0	0.51****	0.26*	0.54
	UN sludge = 1	0.55****	0.28*	0.88
	UN sludge = 2	0.59****	0.27*	0.87
	UN sludge = 3	0.62****	0.24	0.87
	UN sludge = 4	0.63****	0.22	0.19
	UN sludge = 5	0.58****	0.22	0.25
	UN sludge = 6	0.52****	0.21	0.52
	UN sludge = 7	0.45****	0.18	0.37
	UN sludge = 8	0.40****	0.40**	0.54

Statistical test of the relationship		Correlation with baseline CPI	Mutual information with basic CPI	The Granger test for the effect of CPI on the base
	UN sludge = 9	0.37***	0.27*	0.69
	UN sludge = 10	0.37***	0.22	0.06**
	UN sludge = 11	0.33***	0.12	0.04***
	UN sludge = 12	0.24*	0.18	0.05***

Table 3 - Data of the Granger test on the cause-effect dependence of the indicator of the intensity of inflation expectations on various economic parameters according to monthly data

Procedure for the test Grainger (time lag)	Consumer price index (%)	Brent oil price growth rate	The rate of growth of the ruble against the dollar	Change key Bank rates Russia (PP.)	Industrial production index (y / y, %)	Level change unemployments (y / y, PP.)	Real growth rate salary (%, y/y)
1 month	**0.20***	0.32 **VAR: 0.19***	0.48 VAR: 0.62	**0.04*** VAR: 0.20***	0.74 VAR: 0.97	0.89 VAR: 0.78	0.96 VAR: 0.93
2 months	0.25 VAR: 0.63	0.64	0.56	0.15**	0.85	0.99	0.29
3 months	0.53	0.81	0.80	0.36	0.93	0.80	0.57
4 months	0.72	0.83	0.79	0.56	0.98	0.83	0.66
5 months	0.81	0.72	0.92	0.75	0.98	0.30	0.76
6 months	0.92	0.78	0.91	0.54	0.94	0.49	0.80
7 months	0.74	0.89	0.96	0.70	0.79	0.39	0.86
8 months	0.19*	0.65	0.59	0.11**	0.96	0.52	0.57
9 months	0.32	0.82	0.74	0.20*	0.71	0.57	0.70
10 months	0.62	0.78	0.72	0.28	0.88	0.69	0.70
11 months	0.37	0.29	*0.19**	0.14**	*0.17**	0.46	0.47
12 months	0.26	0.15**	0.40	0.20*	0.16*	0.65	0.51

Based on the analysis of monthly reports "Medialogia" for the period 2016-2017 ratings in the media for all the above sections and given profiles, publications, namely the relevance of their publications on economic issues, was compiled by the author's map of the largest Russian economic media that are relevant to the analysis of economic expectations of the Russian population.

The relevance of the obtained set of media was verified and validated using a survey of experts in the field of economic analysis (CMASF, MSU, October, 2016) and submission of the scientific community [Apokin, 2015]. It should be noted that the majority of experts named the following four sources as the key major Russian media for the analysis of economic expectations of the Russian population: Vedomosti, RBC, Expert, Kommersant.

At the same time, to expand the coverage of the relevant audience – interested in the domestic economic situation, reading and commenting on economic news on the Internet – respondents listed additional important media sources, which were also taken into account in the preparation of the

General map of sources.

Within the framework of the developed methodology, the basis for assessing the economic expectations of the Russian population is the opinion of active Internet users regarding economic events, data and trends covered by the largest economic media in articles on the subject of the considered areas of the economy.

It is assumed that the expectations of active Internet users interested in economic events can determine the dynamics of economic expectations of the population as a whole. Thus, according to Rosstat, the share of households with Internet access in Russia as a whole in 2016 was 68% (see figure 1). The trend of steady and rapid growth of the indicator is observed in the last 10 years. So, in 2007 its value was 16%. At the same time, according to the latest available data, the trend continues at the present time. According to the President of PJSC "Rostelecom" M. Oseevsky from October 2017 to currently, the share of households with access to the Internet has exceeded 70%. In the future, the company plans to increase this figure to 90%. Thus, aggregated indicators of economic expectations, obtained on the basis of collecting and processing the opinions of Russian Internet users, can be representative to reflect the aggregate economic expectations of the Russian population.

The technique involves extracting the opinions of Internet users in certain areas of the economy from the comments of Internet users to news articles on a given area of economic expectations in a certain set of economic media at the first step [Goloshchapov, Andreev, 2017].

Since user comments are text messages, aggregated indicators of economic expectations are obtained using text processing techniques.

Three types of sources are used to collect comments from Internet users on articles in economic media: official web pages of the media on the Internet, official media groups in social networks Vkontakte (vk.com) and Facebook (facebook.com).

The first group of sources – the official web pages of the media-aims to reach users who are mainly focused on reading specific sources, and presumably do not use social networks. Social networks are selected based on the analysis of the number of user comments among Internet users [Kaftannikov, Parasic, 2016]. As access to the Internet spreads among households, subscriptions to social media pages and discussion of published materials are becoming more popular. According to the data of the analysis on the basis of the data collected by the minimum representative sample of the media (Kommersant, RBC, Expert, Vedomosti), the ratio of the number of comments in social networks to comments on the official web pages of the media is twelve to one.

It should be noted that as a significant addition to the base of social networks can act as another common network in Russia – "Odnoklassniki" (ok.ru). The inclusion of this source in the analysis may to some extent increase the reach of Internet users to assess indicators of economic expectations.

Thus, the lists of sources of economic media and resources for collecting information about the opinion of Internet users in relation to the published news, representative for assessing the economic expectations of the Russian population in the framework of the author's methodology, were formed.

The next stage of the method is the calculation of indicators of the intensity of expectations of the population for different areas of the economy, given by the researcher. Correct execution of this step allows to obtain a high-frequency approximation of Internet expectations users representing a significant proportion of the Russian population, relative to various economic phenomena [Vorontsov, 2016].

The correct implementation of this stage of the technique involves three steps: (1) calculation of the indicator intensity of expectations; (2) assessment of the relevance of the resulting indicator to a given type of expectations; (3) testing the stability of the data to the rules of calculation of the indicator.

The indicator responsible for approximating the economic expectations of the population is the indicator of the intensity of expectations. The latter shows how high or" intense" are the expectations of a given change on the part of the population at any given time – for example, expectations of a rise in the price level in the economy. This makes it possible to approximate changes in population expectations over time.

The calculation of the indicator is based on determining the number of comments that are directly related to the measured economic expectations. The selection of comments is carried out at the previous stage of the methodology – among the media articles relevant to a given area of Economics, the measurement of economic expectations in which is the aim of the researcher.

At any given time, the indicator is calculated as the ratio of the number of relevant comments to the total number of media articles mentioning expectations similar to those sought in the comments:

$$\sum_{T}^{l+} = \frac{\sum_{t=1}^{T}\sum_{l=1}^{L} Number\ of\ comments}{\sum_{t=1}^{T}\sum_{l=1}^{L} Number\ of\ articles}$$

where T is the time period for calculating the indicator (day, month, etc.);

l+ is the economic expectations set by the researcher (for example, expectations of price growth);

L is a set of articles on a specific field of Economics, the measurement of economic expectations in which is the aim of the researcher (media articles selected in the second stage of the methodology).

According to the presented formula, the numerator of the fraction includes the total number of Internet users comments for period T, mentioning the given economic expectations l+ in accordance with the system of selection rules, to each of the media articles on the topic L.

In this case, the denominator of the fraction includes the total for period T number of media articles on the topic L, mentioning the given economic expectations l+, according to the system of rules developed for the selection of comments. In other words, the selection of media articles for the calculation of the denominator of the fraction of formula is carried out in accordance with the rules or regular expressions identical to those used for the selection of Internet users' comments. This allows you to get a sample of media articles that mention the economic expectations set by the researcher. This step is necessary in order to adjust the economic expectations of Internet users to media expectations or to the possible artificial popularization of the topic by the media. The intensity indicator of economic expectations will tend to reflect the intensity of the "true" concerns a specific issue or problem of the economy – with the exception of the influence of media resources.

Conclusion

The author has developed a method of constructing a high-frequency indicators of economic expectations of the population, based on five key stages: (1) identify sources of information and assessment of the representativeness of the sample for analysis; (2) selection of news articles selected set of Russian economic media; (3) selection of review Internet users to a selected set of news articles economic the media on the subject of specific areas of economic expectations; (4) calculation of the intensity indicator of a given type of economic expectations, including the assessment of the relevance of the obtained indicator to the objectives of this article and verification of the stability of the data to the rules of calculation; (5) identification of factors of change in the intensity of a given type of economic expectations over time.

Verification of the developed method is carried out on the example of constructing a high-frequency indicator of inflation expectations of the Russian population.

The developed technique was implemented on the basis of text data of news articles of four largest economic mass media of Russia (Vedomosti, Kommersant, RBC, Expert) and comments of Internet users to them on official web pages of mass media and in social networks Facebook and Vkontakte.

References

1. Apokin A. et al. Monetary policy: work on mistakes // Questions of Economics. 2015. No. 9. P. 136-151.
2. Bank of Russia. On the methodology of the study of inflation expectations and consumer sentiment of the population // Materials of the official website of the Bank of Russia. Available at: https://www.cbr.ru/Content/Document/File/17324/FOM_meth.pdf [Accessed 16/05/2020].
3. Bank of Russia. The method of quantification of inflation expectations of the population of the Bank of Russia // Materials of the official website of the Bank of Russia. 2017. Available at: https://www.cbr.ru/Content/Document/File/17325/Inflation_expectations_gu ide.pdf [Accessed 12/05/2020].
4. Goloshchapov I., Andreev A. Assessment of inflationary expectations of the Russian population using machine learning // Questions of economy. 2017. No. 6. P. 71-93.
5. Kaftannikov I.L., Parasic A.V. Problems of formation of the training sample in machine learning problems // Bulletin of the South Ural state University. Series: computer technology, control, electronics. 2016. Vol. 16. No. 3. P. 15-24.
6. Krapcev D., Porshakov A., Seleznev S., Sinyakov A. The equilibrium interest rate: estimates for Russia // Bank of Russia, a Series of reports on economic research. 2016. No. 13. Available at: https://www.cbr.ru/Content/Document/File/16727/wps_13.pdf [Accessed 18/05/2020].
7. Perevyshin Yu.N., Rykalin A.S. Modeling of inflation expectations in the Russian economy. Moscow, 2018. 384 p.
8. Pestova A.A. et al. Development of a system of indicators of financial instability based on high-frequency data // Money and credit. 2017. No. 6. P. 49-58.
9. Stolbov M. Google search Statistics as an indicator of the financial situation // Questions of Economics. 2018. No. 11. P. 79-93.
10. Vorontsov K.V. Probabilistic thematic modeling. 2013. Available at: http://www.machinelearning.ru/wiki/images/2/22/Voron-2013-ptm.pdf [Accessed 16/05/2020].
11. Zhemkov M., Kuznetsova O. Measurement of inflation expectations of financial market participants in Russia // Questions of economy. 2017. No. 10. P. 111-122.

Разработка методов построения высокочастотных индикаторов экономических ожиданий населения на основе больших данных

Ардеева Айгуль Наилевна

Магистрант,
Уфимский государственный нефтяной технический университет,
450000, Российская Федерация, Уфа, Первомайская ул., 14;
e-mail: ardeeva.a.n@mail.ru

Закиев Рамиль Раилевич

Магистрант,
Уфимский государственный нефтяной технический университет,
450000, Российская Федерация, Уфа, Первомайская ул., 14;
e-mail: mr.ramil98@mail.ru

Петренко Павел Витальевич

Магистрант,
Уфимский государственный нефтяной технический университет,
450000, Российская Федерация, Уфа, Первомайская ул., 14;
e-mail: pavelpetrenko1999.job@mail.ru

Хасанов Арсен Ильшатович

Магистрант,
Уфимский государственный нефтяной технический университет,
450000, Российская Федерация, Уфа, Первомайская ул., 14;
khasanovarsen@yandex.ru

Ардуганов Артур Ильдарович

Магистрант,
Уфимский государственный нефтяной технический университет,
450000, Российская Федерация, Уфа, Первомайская ул., 14;
e-mail: ardtur2@mail.ru

Аннотация

Ожидания экономических агентов являются основой для экономических теорий, объясняющих процессы потребления, инвестиций, спроса на деньги, инфляции и т.д. В научной литературе экономические ожидания определяются как ожидания экономических агентов – домашних хозяйств, фирм, правительства и экспертов. сообщества – относительно будущих событий в экономике, которые существенно влияют на принятие решений. В современных экономических исследованиях существуют два ключевых подхода к оценке экономических ожиданий: подход, основанный на опросах экономических агентов (подход, основанный на опросе), и подход, основанный на анализе параметров обмена (рыночный подход). Однако эти подходы имеют ряд существенных недостатков. С 2015 года официальным режимом денежно-кредитной политики в России является таргетирование инфляции, переход к которому фактически был инициирован в сентябре 2013 года. В течение этого периода уровень инфляционных ожиданий населения является важным объектом исследования для Банка. России. Предполагается, что численный ориентир регулятора для инфляции в текущем режиме денежно-кредитной политики окажет значительное влияние на инфляционные ожидания населения. Последние, в свою очередь, могут повлиять на уровень цен в долгосрочной перспективе. Разработка методологии построения показателей экономических ожиданий на основе больших наборов данных необходима в связи с их значительным влиянием на экономические процессы и способностью преодолевать основные ограничения существующих методов измерения ожиданий с использованием методов анализа больших данных.

Библиография

1. Апокин А. и др. Денежно-кредитная политика: работа над ошибками // Вопросы экономики. 2015. № 9. С. 136-151.
2. Банк России. Метод количественного определения инфляционных ожиданий населения Банка России // Материалы официального сайта Банка России. 2017. URL: https://www.cbr.ru/Content/Document/File/17325/Inflation_expectations_gu ide.pdf
3. Банк России. О методологии изучения инфляционных ожиданий и потребительских настроений населения // Материалы официального сайта Банка России. URL: https://www.cbr.ru/Content/Document/File/17324/FOM_meth.pdf
4. Воронцов К.В. Вероятностное тематическое моделирование. 2013. URL: http://www.machinelearning.ru/wiki/images/2/22/Voron-2013-ptm.pdf
5. Голощапов И., Андреев А. Оценка инфляционных ожиданийроссийского населения методами машинного обучения // Вопросы экономики. 2017. № 6. С. 71-93.
6. Жемков М., Кузнецова О. Измерение инфляционных ожиданий участников финансового рынка в России // Вопросы экономики. 2017. № 10. С. 111-122.
7. Кафтанников И.Л., Парасик А.В. Проблемы формирования обучающей выборки в задачах машинного обучения // Вестник Южно-Уральского государственного университета. Серия: компьютерные технологии, управление, электроника. 2016. Т. 16. № 3. С. 15-24.
8. Крапцев Д., Поршаков А., Селезнев С., Синяков А. Равновесная процентная ставка: оценки для России // Банк России, Серия отчетов по экономическим исследованиям. 2016. № 13. URL: https://www.cbr.ru/Content/Document/File/16727/wps_13.pdf
9. Перевышин Ю.Н., Рыкалин А.С. Моделирование инфляционных ожиданий в российской экономике. М., 2018. 384 с.
10. Пестова А.А. и др. Разработка системы индикаторов финансовой нестабильности на основе высокочастотных данных // Деньги и кредит. 2017. № 6. С. 49-58.
11. Столбов М. Статистика поиска Google как индикатор финансовой конъюнктуры // Вопросы экономики. 2018. № 11. С. 79-93.