Information Processing and Management 59 (2022) 103043

Contents lists available at ScienceDirect

Information Processing and Management

journal homepage: www.elsevier.com/locate/infoproman

Topics of the nationwide phone-ins with Vladimir Putin and their role for public support and Russian economy

Ivan Savin[a,b,*], Nikita Teplyakov[b]

[a] Institute of Environmental Science and Technology, Universitat Autònoma de Barcelona, Spain
[b] Graduate School of Economics and Management, Ural Federal University, Yekaterinburg, Russian Federation

ARTICLE INFO

Keywords:
"Direct line with Vladimir Putin"
Machine learning
Presidential address
Public approval
Russia
Topic modelling

ABSTRACT

The addresses of national leaders can affect their public support and spur changes in the country's economy. To date, very few studies exist establishing these relationships, and no research has been done on the addresses from Vladimir Putin. In this paper we fill this knowledge gap by analysing the nationwide phone-ins of Putin, a special annual format where he addresses the public, and using structural topic modelling studying their topics over time. Furthermore, we relate these topics to public approval of the president and the government as well as to some Russian macroeconomic indicators such as inflation and budget expenditures. Based on our data containing 1938 responses and almost 250 thousand words, we identify 16 main topics covering areas from healthcare and education through economics to elections and legislation. We find that the topic of foreign affairs has gained in popularity over time the most (from around 4.5% at the beginning to more than 10% starting from 2014). Another topic, consistently gaining weight in the president's statements, is related to solving particular problems of the general public (from 8% to 12.5%) and is significantly correlated with subsequent decrease in the country's unemployment (Pearson's correlation coefficient -0.502). We also find that when the government's support is decreasing, Putin tends to discuss more socially significant topics (e.g., inflation, healthcare, Pearson's coef. around -0.5), while when the support is rising, he speaks more about foreign affairs (Pearson's coef. 0.773). Our study provides first evidence that Vladimir Putin may adapt the content of his phone-in meetings to gather public support and influence the country's economy.

1. Introduction

The addresses of national leaders give important signals for officials, business representatives and ordinary citizens. They can set long-term trends and encourage immediate actions (Peake & Eshbaugh-Soha, 2008). This is particularly true for Russia, where the role of one man, Vladimir Putin, is exceptionally important (Rotkirch et al., 2007; Petrov et al., 2014). In our study we aim to contribute to the literature addressing the public speeches of national leaders and discussing their role on the public approval within the country (MacKuen, 1983; Jennings & John, 2009; Berlemann & Enkelmann, 2014; Dybowski & Adämmer, 2018). To this end, we study all nationwide phone-ins with Vladimir Putin (also known as "Direct Line with Vladimir Putin"), over the period 2001-2020. Nationwide phone-ins with Vladimir Putin is a nationally broadcasted TV program, where usually once a year he answers the questions from the

* Corresponding author at: ICTA, Edifici Z, UAB Campus, Bellaterra 08193, Spain.
E-mail address: ivan.savin@uab.cat (I. Savin).

https://doi.org/10.1016/j.ipm.2022.103043
Received 18 February 2022; Received in revised form 4 July 2022; Accepted 20 July 2022
Available online 29 July 2022

public, which are preselected by the presidential administration. We classify the content of the phone-ins into distinct themes ("topics") and show how the importance of those topics was changing from year to year.

We apply advances from natural language processing on the collected textual data, the method called topic modelling (TM), that has been developed in the last two decades thanks to the rapid development of computer technology and machine learning algorithms (Albalawi et al., 2020; Bing et al., 2020).[1] TM has been applied to text classification in a wide range of areas including patent data (Venugopalan & Rai, 2015; Chen et al., 2017; Savin et al., 2022c) and scientific publications (Chen et al., 2020; Savin & van den Bergh, 2021; Savin & Teplyakov, 2022), political debates and petitions (Hagen, 2018; Wei et al., 2020), survey open-questions (Tvinnereim and Fløttum, 2015; Tvinnereim et al., 2017, 2021; Liu et al., 2021; Savin et al., 2021), firm descriptions in business platforms (Savin et al., 2022a; Żbikowski and Antosiuk, 2021) and publications in mass media (Lenz & Winker, 2020; Park et al., 2016).

In the natural language processing literature, there are many TM algorithms for extracting topics from textual data. The traditional latent Dirichlet allocation (LDA) method is still very popular because of its simplicity and interpretability (Momtazi, 2018; Obadimu et al., 2021). Because of certain limitations, LDA has seen several extensions such as correlated topic modelling, weighted LDA and structural topic modelling (Daenekindt & Huisman, 2020; Takahashi et al., 2020; Roberts et al., 2014). LDA also has been combined with neural networks (Zhao et al., 2021). Apart from developing new algorithms, researchers also experimented with representations of texts. Curiskis et al. (2020) use TF-IDF, word2vec and doc2vec text embeddings and cluster them with machine learning techniques, such as k-means or hierarchical clustering. Bagheri et al. (2018) experiment with the neural text representations and show that they are superior to word embeddings and inferior to entity embeddings on the data used. In this paper we employ a structural topic modelling (STM) algorithm (Roberts et al., 2019), as it allows for a more realistic representation of topics in short documents (responses) and can incorporate additional information about the statements from Vladimir Putin.

To our knowledge, this is the first study that applies topic modelling to textual responses of the Russian president. This is also one of the first studies using computational-linguistic analysis to explore public speeches by the heads of the state. Earlier, Ahmadian et al. (2017) examined the campaign speeches of the US presidential candidates judging them by grandiosity, informality and dynamism, while Ficcadenti et al. (2019) proposed methods for processing public speeches of US presidents using rank-size laws. While other studies typically focus on the USA and aggregate public addresses over several politicians (Dybowski & Adämmer, 2018; Ruhl et al., 2018), we contribute to the literature by following the rhetoric of one person and considering a different country's context.

This analysis delivers a model with 16 topics based on the highest prediction power and topic exclusivity achieved by this model. We interpret and label these topics based on the most frequent and exclusive words and their illustrative statements with the highest share of those topics. Furthermore, we calculate the prevalence of topics over time and test for the presence of a trend for each of them. We find that economic and social issues are addressed less frequently lately, while sports and foreign affairs are becoming more popular. This way, we observe an important shift in the focus of nationwide phone-ins, which may be contributing to changes in the approval of authorities or be related to economic changes in the country.

To investigate the hypothesis that the public approval of the president, government, and governors as well a some macroeconomic indicators of the country are related to the prevalence of certain topics, we apply a series of correlation tests. Interestingly, some of the topics having positive association with the public approval gained in popularity over time, but topics having a negative association have not increased. This can be interpreted as first evidence that, along with answering the questions from the general public, the nationwide phone-ins serve the purpose of gathering support of the president.

The remainder of the paper is organized as follows. In Section 2 we present our research questions. Section 3 describes our dataset and the methodology for topic modelling. Section 4 provides the results of the topic analysis. Section 5 discusses the relationships between the topics and public approval of the president and the federal and regional governments as well as some macroeconomic indicators of the country. Section 6 discusses implications of our research and concludes.

2. Research questions

In this paper we aim to understand the content of nationwide phone-ins with Vladimir Putin, classify this into distinct topics and explore how popularity of these topics over time is related to the public approval of the president and the government as well as some macroeconomic indicators of the country such as inflation, unemployment, and budget expenditures. More particularly, our research is focused on the following four research questions (RQs):

RQ1: How many main topics are present in the nationwide phone-ins, and what are they about?

RQ2: How did the shares (prevalences) of those topics change over time? Was their increase (or decrease) statistically significant?

RQ3: Do changes in absolute levels of public approval of the president, the federal government and governors[2] correlate with the shares of these topics? If yes, how this can be interpreted?

RQ4: Do changes in the Russian macroeconomic indicators correlate with those topic shares? How can these correlations be meaningfully justified?

[1] According to Chowdhary (2020), one of the main NLP tasks is discourse analysis, where topic modelling is frequently used (Jacobs & Tschötschel, 2019; Savin et al., 2022b).

[2] Henceforth, when talking about public approval, we will use the terms "governors" and "regional governments" interchangeably, as implied in the data source.

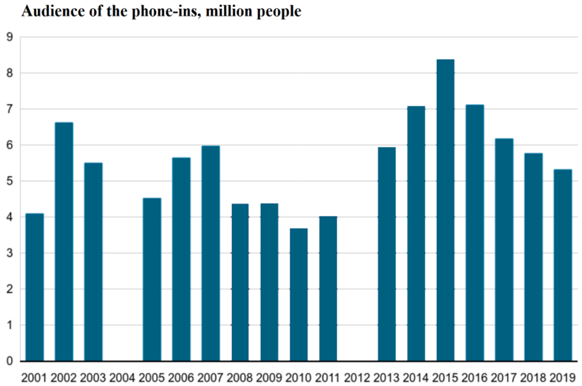

Fig. 1. Audience of the nationwide phone-ins with Vladimir Putin (in million viewers).
Source: BBC News Russian Service (2019).

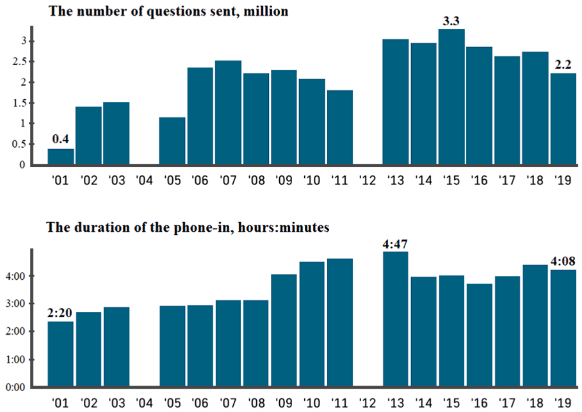

Fig. 2. Statistics on nationwide phone-ins with Vladimir Putin.
Source: TASS, news agency (2019). Note that the notation of the y-axes is provided on the top of the charts, while x-axis shows different years.

3. Data and methodology

3.1. Data description

In the Russian Federation, nationwide phone-ins with Vladimir Putin started in 2001 and have been regularly conducted on a yearly basis since then (with three exceptions: 2004, 2012 and 2020).[3] In 2004 and 2012, the president took a break from communication through a nationwide phone-in with citizens on the air and instead the Kremlin organized an extensive press conference of Vladimir Putin. In 2020 the event planned for June was cancelled due to the coronavirus pandemic, and Vladimir Putin again gave a press conference in December of that year answering questions from journalists. In Putin's own words during the phone-in in 2015, the nationwide phone-ins serve as "the most powerful sociological survey that allows citizens to express their position and assess the country's leadership" (Official network resources of the President of Russia, 2015). This Q&A television program is typically broadcasted by the main federal TV channels such as "Channel One Russia" and "Russia-1" securing a wide television audience. Fig. 1 shows that the TV audience of the phone-ins reached a peak in 2015, gathering more than 8 million people. Since then, this number has declined to a little of over 5 million, which is about 3% of the Russian population. In fact, the audience of the nationwide phone-ins is much higher than the number of TV viewers alone, thanks to the use of the Internet and digital social networks. For example, in 2018, the audience of the respective phone-in was more than 40 million people, which is about 27% of the population of Russia at that time (Gazeta.ru, 2018). This figure is distributed as follows: (1) 16 million viewers were received from Instagram influencers; (2) the official broadcast on the Odnoklassniki social network collected 10 million views; (3) the VKontakte service provided 3.9 million viewers; (4) TV broadcast provided further 5.8 million viewers; (5) the source of attracting the rest of the audience is not specified.

According to Komsomolskaya Pravda (2017), the procedure for selecting questions for nationwide phone-ins works as follows.

[3] Between 2008 and 2011 Putin was participating in the event as a prime minister, maintaining the format of the event. However, henceforth we use the word "president" as a synonym to Vladimir Putin.

Shortly before and during the broadcast, any volunteer can send a question by phone or via the Internet to a special center created with the participation of the two main federal channels (mentioned above). It carries out initial moderation of the questions and their filtering for the purpose of further transfer to the presidential administration. Next, the staff of the presidential administration groups the pre-selected messages by topic and issue. Then, for each of the topics, several questions (typical and "bright") are selected for broadcast. These questions are returned to TV channels to organize the departure of film crews. Such a process of selecting questions suggests that a nationwide phone-in is not very different from the usual address of the president, where only preselected topics are addressed. However, we want to highlight a few key differences: (1) the initial agenda for each of the nationwide phone-ins, despite the many stages of moderation, is co-created by the general public increasing its attractiveness to the population; (2) during the nationwide phone-in Putin can receive feedback on his statements, and a dialogue may be initiated. Based on this, we conclude that nationwide phone-ins with Vladimir Putin represent a special form of presidential address, where he discusses some selected problems raised by the general public.

Fig. 2 shows that the number of questions received, and the duration of the phone-ins broadcasted live varied over time. The number of submitted questions increased from 0.4 million in 2001 to 2.2 million in 2019. The duration of the broadcasts increased over the period 2001-2013 from 2 hours and 20 minutes to over 4 hours of screen time, and since then remained quite stable.

The media shorthand the statements from all participants in the event and put them in the form of text interviews. We used materials from the Kremlin's official website and "Rossiyskaya Gazeta", which is published by the Russian Government (Official network resources of the President of Russia, 2015; Rossiyskaya Gazeta, 2019). We accessed these sources in April 2020 to collect and process the textual data. To avoid possible translation bias, we conduct our analysis on the original text (in Russian), but provide results translated in English.

The original texts of nationwide phone-ins contain the statements from various actors which can be grouped in three categories: the president, citizens and moderators (TV presenters from the federal TV channels). We consider the president's responses only for several reasons. First, as explained earlier, the questions have been thoroughly preselected through multiple filters. Second, many statements by moderators and citizens are not informative, as they contain organizational cues (like "you can see how many people have gathered here to ask the president a question") and tend to devote many words to presenting themselves. Finally, our aim is to understand the content and role of speeches from Putin on the socio-economic life of the country. As one can see in Fig. A1 in the Appendix, the president's answers are a major part of the phone-ins being around two to three times as long as statements from the citizens and the moderators combined.

Before proceeding with the description of the textual data, we would like to point out that information on public approval data and macroeconomic indicators can be found in the respective Sections 5.1 and 5.2. They contain information on data sources, methodology for calculating the original indicators, and our strategy for their further processing. The differences calculated on the basis of these indicators, which we use in our analysis, can be found in the Appendix (Tables A4–A7). The data on public approval and macroeconomic indicators contain 17 observations corresponding to the number of nationwide phone-ins considered in the paper.

Fig. A2 in the Appendix provides further information on the length of the phone-ins. In line with Fig. 2 on the duration of the phone-ins, the length of the events has become longer, also as the number of responses increased over time, starting with 63 in 2001, peaking with 193 in 2013, and stabilizing afterwards in the range 120 to 180 (see left chart in Fig. A2). At the same time, the average length of the responses ranges from 1 word to 1187, having a right-skewed shape of the distribution (right chart in Fig. A2). In total, we analyse 1938 responses from Vladimir Putin in total containing 244927 words.

We pay particular attention to text pre-processing as it is essential for a good topic model (Kannan & Gurusamy, 2014). In this paper we use the following procedure for text pre-processing:

1 *Tokenize* responses, i.e. split them into distinct text units. In our case, these are words and punctuation marks. For example, the response "Эта задача перед ПравительствоМ поставлена, и Мы ее будеМ решать" will be represented like ['Эта', 'задача', 'перед', 'правительствоМ', 'поставлена', ',', 'и', 'Мы', 'ее', 'будеМ', 'решать'].[4]Convert all capital letters into lower case (including proper names and toponyms).
2 *Lemmatize words, i.e.* transform words to their initial dictionary form. Lemmatizing is preferable to stemming (that reduces words to their stems, base or root form by deleting endings and suffixes) because after lemmatizing one can still see the word in full and avoid ambiguous shortenings. We used the original 'mystem' algorithm for Russian language introduced by Segalovich (2003), Zelenkov et al. (2014)[5].
3 *Remove punctuation and non-letters.*
4 *Remove stop words* (such as "Мы", "и")[6] extended with names of the moderators and citizens, common Russian names and our custom stop word dictionary. This had the biggest impact on the reduction of the length of responses.

[4] "This task has been set before the Government, and we will solve it" will be represented like ['This', 'task', 'has', 'been', 'set', 'before', 'the', 'Government', ',', 'and', 'we', 'will', 'solve', 'it'].

[5] Working with Russian scientific and technical texts in 1995-1996, Segalovich found out that more than half of the words are not present in grammatical dictionaries. The solution, created by Ilya Segalovich and his team, allows one to generate "synthetic" morphological descriptions for unknown words. The algorithm is implemented in the "pymystem" module for the Python language.

[6] Stop words like "we" and "and".

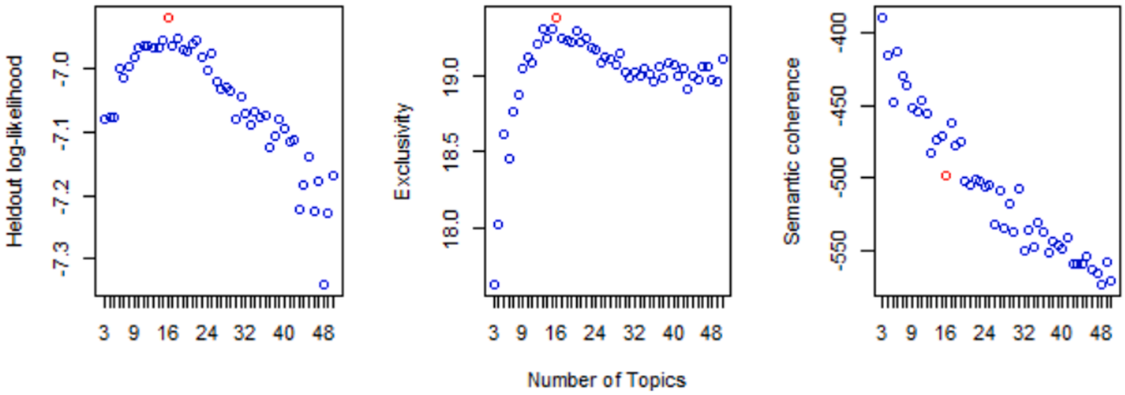

Fig. 3. Model performance for distinct numbers of topics.

5 Retrieve multi-word expressions. Using Normalized Pointwise Mutual Information (Bouma, 2009) score, we declared certain sequences of words appearing predominantly together as single words and combined them with underscores (e.g. "farthest_east" and "maternity_capital").
6 Remove *rare words* which occur less than 5 times in all responses (Savin et al., 2020).

Text pre-processing considerably reduced our data set making it easier to classify by a computer. The example shown in Table A1 in the Appendix illustrates how raw text data is transformed into clean tokens for building topic models. The average length of the responses reduced from 126 to 36 words; the maximum fell from 1187 to 351 terms.

3.2. *Structural topic modelling*

The topics of the president's responses are formed by means of the topic modelling (TM) method. This is a computer-based approach developed at the intersection of machine learning and natural language processing that allows to discover distinct topics presented in text. In formal terms, a topic model is an outcome of Bayesian inference of words related to a given topic and the topics being discussed in a given text, based on texts observed so far. Put differently, a topic model backtracks from the texts the topics that are likely to have generated them. The advantage of TM over counting keywords is that it considers words not in isolation but in the context of other words they appear with allowing words to have different meanings depending on their context. TM assumes that each word in the texts is generated through a multiple repetition of a two-step process: first, that each text has its own topic distribution, and a topic is randomly drawn from it; second, that each topic has its own word distribution, and a word is randomly drawn from this distribution for the topic selected in the first step. As a result, each response by Putin exhibits multiple topics in different proportions.

TM discovers the topic distribution for each response and the word distribution of each topic iteratively, by fitting this two-step procedure to the textual data until it finds the best model that describes the underlying data. Similar to cluster analysis, TM reduces the dimensionality of linguistic data from words to topics based on their co-occurrence in a collection of responses to infer the underlying topics in those texts and the weight of each topic in each individual response. For example, if we observe the word "ukraine" in the topic 3 that we label as "Foreign Affairs", it implies that this word appeared more frequently and exclusively in combination with other words in this topic, and that Vladimir Putin used it more often in this context.

An important advantage of structural topic modelling (STM) over classical TM is that it incorporates additional information about the responses (in our case, year of the phone-in). Hence, instead of assuming that topical prevalence (i.e. the degree to which a single response belongs to a given topic) and topical content being constant across years, we use the information as updated Bayesian priors in which we expect to see variance between responses. This has proven to result in higher quality topic identification when textual responses are relatively short (Roberts et al., 2014). Furthermore, unlike basic TM, STM allows for topics to be correlated within responses, which results in a more realistic representation of the textual data. This is in line with the idea that people tend to connect certain topics (Blei & Lafferty, 2007). For example, people discussing social security issues are more likely to address the problem of inflation and unemployment, than, for example, foreign affairs.

4. Results from topic modelling

Determining the number of topics k can be a challenging task. Ideally, this number should maximize model's prediction power, topic's exclusivity and semantic coherence. The first metric stands for the accuracy of the model in forecasting words from a sample that has been excluded (held-out) from the model estimation step; exclusivity measures the likelihood of seeing the topic given the words (i.e. whether most frequent words from topics do not overlap), while semantic coherence captures the extent to which words belonging to the same topic tend to appear in the same responses. Fig. 3 shows the performance of alternative model specifications (k ranges from 3 to 50). We decided for 16 topics as this allows us to reach simultaneously highest prediction power and exclusivity while

Table 1
Overview of topics (the topics are ordered in the decreasing order of their prevalence).

	Topic label	Top 10 words by frequency and exclusivity [RU, EN]	Topic share
1	Solving Particular Problems	problem, specific, solve, help, situation, minister, probably, medicine, need, carefully *проблеМа, конкретный, решать, поМогать, ситуация, Министр, наверное, лекарство, нуЖдаться, вниМательно*	10,25%
2	Legislation & Migration	take, responsibility, respond, duma, appropriate, decision, order, proposal, law, promise *приниМать, ответственность, отвечать, дуМа, соответствующий, решение, порядок, предлоЖение, закон, обещать*	9.05%
3	Foreign Affairs	ukrainian, ukraine, crimea, international, american, unite_state, find, dialogue, position, attitude *украинский, украина, крыМ, МеЖдународный, аМериканский, соединять_штат, находить, диалог, позиция, отношение*	8,83%
4	Elections & Regional Government	authority, elections, parliament, governor, party, react, try, homeland, leader, letter *власть, выборы, парлаМент, губернатор, партия, отреагировать, постараться, родина, руководитель, письМо*	8,35%
5	Law & Criminality	crime, drug, law_enforcement_agency, court, violation, fight, employee, prosecutor, criminal, struggle *преступление, наркотик, правоохранительный_орган, суд, нарушение, бороться, сотрудник, прокуратура, уголовный, борьба*	7,34%
6	Economic & Political Development	state, consolidation, achieve, consistently, result, national, economic, own, definite, many *государство, укрепление, добиваться, последовательно, результат, национальный, эконоМический, собственный, определенный, Многий*	6,78%
7	Sports & Infrastructure	sports, object, infrastructure, farthest_east, sochi, build, sports, construction, construct, football *спорт, объект, инфраструктура, дальний_восток, сочи, построить, спортивный, строительство, строиться, футбол*	6,58%
8	Healthcare & Education	healthcare, medical, teacher, school, municipal, institution, specialist, education, medicine, university *здравоохранение, Медицинский, учитель, школа, Муниципальный, учреЖдение, специалист, образование, Медицина, вуз*	6,11%
9	Social Security	family, category, retirement, veteran, benefit, age, pensioner, pension, parent, maternity_capital *сеМья, категория, пенсионный, ветеран, пособие, возраст, пенсионер, пенсия, родитель, Материнский_капитал*	5,57%
10	Housing & Agriculture Support	allocate, housing, house, bank, rate, billion, pay, ruble, agriculture, mortgage *выделять, Жилищный, Жилье, банк, ставка, Миллиард, выплачивать, рубль, сельский_хозяйство, ипотечный_кредит*	5,38%
11	Inflation, Unemployment & Income	growth, inflation, wage, percentage, increase, small, salary, indicator, grow, increase *рост, инфляция, заработная плата, процент, повышение, небольшой, зарплата, показатель, расти, вырастать*	5,34%
12	Armed Forces	armed, weaponry, combat, navy, army, military, defense, weapons, ministry, unit *вооруЖенный, вооруЖение, боевой, флот, арМия, военный, оборона, оруЖие, Министерство, подразделение*	4,57%
13	Taxation & Entrepreneurship	small, business, capital, create, condition, factor, production, investment, high-tech, export *Малый, бизнес, капитал, создавать, условие, фактор, производство, инвестиция, высокотехнологичный, экспорт*	4,38%
14	Public Procurement	enterprise, order, plan, factory, congratulate, technology, airplane, transportation, russian_railways, settlement *предприятие, заказ, план, завод, поздравлять, технология, саМолет, перевозка, оао_рЖд, поселок*	4,08%
15	International Trade	belarus, foreign, integration, sanction, partner, manufacturer, product, process, production, kazakhstan *белоруссия, иностранный, интеграция, санкция, партнер, производитель, продукт, процесс, продукция, казахстан*	3,90%
16	Fossils & Monetary Reserves	price, gas, oil, tax, currency, dollar, tariff, sell, company, payment *цена, газ, нефть, налог, валюта, доллар, тариф, продавать, коМпания, платеЖ*	3,52%

semantic coherence is at its intermediate level.

In line with Tvinnereim et al. (2017), after the optimal number of topics has been selected, ten STM runs were performed with different seed values for the algorithm, which produced slightly different topic models. These models have been subsequently evaluated by both authors independently to identify the topic model that is meant to be the best expression of the diversity of topics identified through the manual reading of the whole textual corpus. Both authors ranked the seventh out of the ten STM runs as the best in representing the main themes from the nationwide phone-ins, and this model is discussed below.

Afterwards, we proceed to the analysis of their content, starting preparations for assigning labels to topics. It is a subjective step, as it depends on our interpretation of the model's output. We have selected the topic labels to reflect their main content in a clear and concise way after carefully examining each topic's top words and illustrative responses (i.e. responses with highest topic prevalence).

In Table 1 we provide topic labels together with 10 words scoring highest in terms of their frequency and exclusivity. The words or word sequences in each set are arranged in a descending order of the FREX[7] metric score. For instance, the word "crime" in the fifth topic about law and criminality is more frequent and exclusive in this topic than the word "struggle".

In Table A2 in the Appendix for each of the topics we report a fraction from one of the illustrative responses to better exemplify their subject. For instance, in the seventh topic Putin talks about mega-events in sports held in Russia and the infrastructure that has been

[7] This metric is called FREX (i.e., FRequency and EXclusivity). The weights of frequency and exclusivity here are the same and equal 50%.

Fig. 4. Word cloud representation of the topics.
Note: The font size corresponds to the probability (weight) of the respective word given the topic, while the colour of the word corresponds to its exclusivity (the darker the colour, the more exclusive are the words).

constructed to host those (e.g. football fields or Olympic village). The ninth topic is about material assistance to the population in the form of maternity capital or pension. The rest of the topics also fit well to the labels, revealing information about migration, law, armed forces and other tissues. Topics 1 and 6 are perhaps the most general. While the latter covers discussion about changes in the economic and political system of the country (e.g., cultural differences between Russia and Europe), the former answers to various personal requests from citizens that did not fall within the other topics (for example, storage and recycling of waste, fires and floods, public transport).

We construct word clouds to graphically represent the topics' most frequent and exclusive words (Fig. 4), where font size indicates word frequency and darkness exclusivity. For example, in topic 3 the equally frequent words "ukraine" and "attitude" have different exclusivity, which shows that the latter word tends to appear in other topics more frequently. This nuance is not visible from an aggregated FREX measure.

The clouds also show that while some topics are largely dominated by few if not single words (topics 1, 6, 11, 14 and 16), others constitute combinations of words with more equal distribution of weights among them. These combinations of words provide insights that go beyond what one can learn from the topic label. For example, in topic 13 the words "development", "economic", "create", "condition", "small", "business" come out strongly, which suggests attention to the role of small business in the economy during the phone-ins.

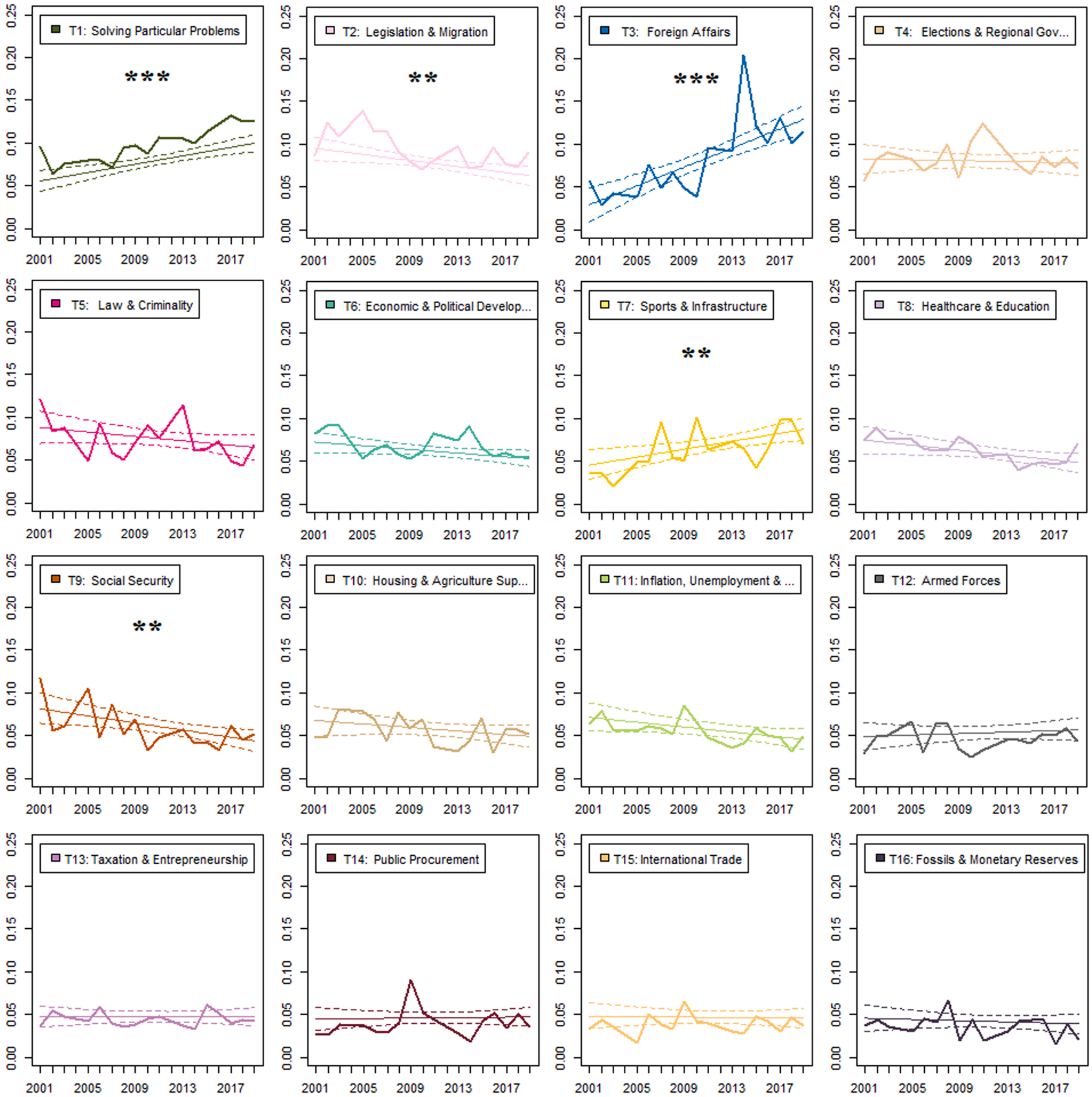

Fig. 5. Prevalence of topic over time.
Note: Values generated by a regression where the outcome variable is the proportion of each president's response dedicated to each topic, given the selected STM model. The panel shows actual topic prevalences for each year, point estimates and confidence intervals of the effects of time on topic prevalence. Confidence intervals plotted as dashed lines indicate the 95% uncertainty range and include both regression and measurement uncertainties associated with the STM model. *** and ** denote 0.1% and 1% significance level, respectively.

We move to analysing topic prevalence (i.e. share of textual responses by Vladimir Putin classified to a particular topic) across all the responses (last column in Table 1). The most popular topic (T1[8]) is almost three times more prevalent than the rarest one (T16). It is a rather trivial conclusion that the most common topic concerns particular problems as this is assumed by the format of the phone-ins. This topic is also less coherent than others and covers very different problems from deafblind people and availability of drugs in pharmacies to shallowing of the Volga River and releasing killer whales). While the topics related to internal and foreign politics (e.g., T2-T5) have the highest prevalence, economic topics (e.g., T13-T16) have the lowest.

Next, we consider how the relative shares of topics change over time. The main question here is which topics have gained in (relative[9]) popularity over time and which ones have lost. To establish the presence of a time trend, we estimate the linear model used to build our STM topic model for each of the topics indexed by i as follows:

[8] Henceforth, we refer to Topic X by TX.
[9] Relative, because the prevalences of all topics in each year must some up to unity.

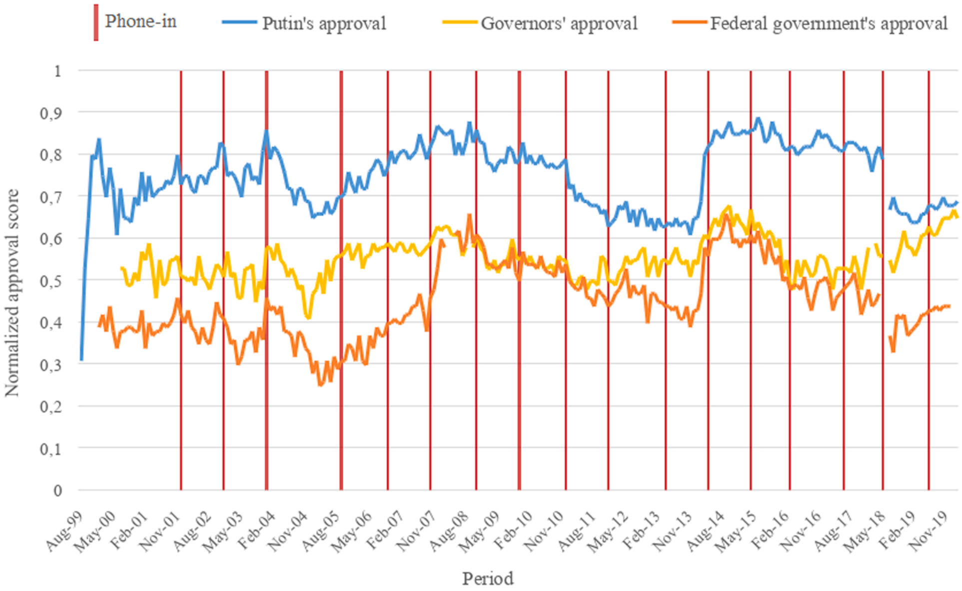

Fig. 6. Public approval of the Russian authority over time.
Source: Levada Center (2020).

$$Topic\ Prevalence_i \sim Constant_i + Year + Residual_i$$

Fig. 5 illustrates the results of a topic prevalence analysis, where for five out of 16 topics the presence of a time trend is significant at 5% level or better. We also provide regression coefficients in Table A3 in the Appendix. The results show that over time the prevalence of T3 on Foreign Affairs, for example, was growing on average by 0.6%, starting from around 4.5% in years 2001-2002 and consistently exceeding 10% starting from 2014. This is the largest increase among all the topics we have identified. Another topic that is constantly gaining popularity is related to solving particular problems of the general public, the prevalence of which is growing by an average of 0.2% per year (from 8% to 12.5%). On the other hand, T2 and T9 on Migration and Social Security demonstrate the maximum loss of on average 0.2% of their prevalence per year.

5. Correlation analysis

In this section, we describe the relationships that emerge between the topics we identified in the phone-in meetings with Vladimir Putin and the country's economic and political indicators. Public approval of Vladimir Putin, the government and governors are used to characterize the political situation in the country. Inflation rate, unemployment, real wage index and budget expenditures are considered as macroeconomic variables. Data sources and indicator descriptions are given below in the respective subsections. We review the preceding and subsequent changes in the above-mentioned indicators, trying to relate them to the dynamics in topic prevalence.

5.1. Relationship between topics and public approval

Public approval is one of the key indicators of political stability in a country. The authorities can assess their decisions by how their public approval has changed. Furthermore, they can foresee changes in their public approval and strategically adjust their actions to avoid losing public support.

Approval indicators in this work are based on social surveys by the Levada Research Center (Levada Center, 2020), which has been conducting regular monthly surveys since the early 2000s and considered as an independent pollster (Freedom House, 2022). This non-governmental research structure was created by former employees of the state-owned polling institution, the All-Russian Public Opinion Research Center (VCIOM). Today, Levada Center is one of the largest research companies in Russia. The main partners and customers of its analytics and research are companies, universities, research, and non-profit organizations. The Levada Center's network of interviewees includes over 100 regional partners located in more than 50 regions of Russia. Surveys are conducted at 137 survey points, including 97 urban areas and 40 rural areas, stratified by federal districts and population. As a result, the total sample of 1600 respondents above 18 years old is divided into 38 strata, each of which recruits respondents in proportion to the adult population of Russia in each stratum. For a more detailed description of the methodology, please refer to the Levada Center website (Levada Center, 2022).

The respondents were asked whether they generally approve of the activities of Vladimir Putin, government or regional governor. The final approval score is the proportion of people who answered yes.[10] The dynamics of the three selected approval scores is displayed in Fig. 6. Vladimir Putin has a consistently higher approval rate than the federal and regional governments during the entire period under review. All three approval ratings are subject to significant fluctuations, the strongest of which occurred against the backdrop of the 2014–2015 conflict in Ukraine.

The direction of the relationship between the prevalence of topics and changes in public approval (or changes in economic indicators) can in principle go both ways. On the one hand, previous fluctuations in public approval or macroeconomic variables can be an incentive to focus more or less on certain topics in the upcoming national phone-in. On the other hand, topics raised during a live broadcast may influence subsequent changes in the approval rates or the economic performance. We test both options, because significant correlations, despite the possible influence of other factors, are worth mentioning.

One problem with our data is that it is impossible to determine the exact days of the survey, i.e. whether the approval rating of the president, the federal or regional governments in a certain month was measured before or after the phone-in that took place in the same month. Therefore, we prefer to ignore the rating which is measured during the same period as the phone-in. Instead, we look either on the difference between preceding and subsequent months to test if the phone-ins had a certain effect on the ratings, or on the difference in ratings between the two preceding months to see if change in ratings has any statistical relation to the topic prevalences.

We proceed by testing the sensitivity of public approval to changes in prevalence of different topics. On the left chart of Fig. A1 we report the identified Pearson correlations, where the Y-axis contains the topic labels, while the X-axis shows approval ratings of Vladimir Putin, the federal government, and regional governments. We discover a positive correlation at the 5% significance level between the topic about Housing & Agriculture Support (T10) and the approval of the governors. In other words, if the president talks more about public aid to housebuilders and farmers, this tends to positively affect public approval of the regional governments. As an example, in the Supplementary information we provide a full response of Putin on building new houses in the Kuzbass region: "I hope that together with the region administration and the governor we will be able to continue these programs, because many emergency buildings still need to be resettled. Kuzbass still has a lot of this "stuff"." Furthermore, during the phone-in meetings, a lot of attention is paid to the issues of agriculture, mortgages, and affordable housing. In particular, farmers ask the president to support domestic producers while families seek for an assistance in purchase of housing. The approval of the governors may be strengthened by the president, who promises that local authorities will resolve these issues. At the same time, the president's approval is not sensitive to this topic.

On the right chart of Fig. A3 we look at the difference in approval in the two periods preceding the nationwide phone-in. This means that if the fluctuations in the approval of the authorities are noticed before the phone-ins, the president can adapt the content of his public addresses. Here, in contrast to the case where we tested changes in public approval to changes in shares of topics, the number of significant correlations is higher, although still relatively small.

There is a strong positive correlation between the difference in the public approval of Vladimir Putin and the share of the topic about Economic and Political Development (T6). This topic can be characterized as general because there the president talks more about the development of the country as a whole, without focusing on specific problems. For example, Putin says here "… if we divide Europe, European values and European people, if we engage in separatism in the broad sense of the word, we will all be unimportant, uninteresting players and will have no influence on world development or even on our own." When Vladimir Putin's approval is growing, it is logical to talk about Russia's development strategy, and when it falls, pay attention to the reasons of the decline.

The president can also react to changes not only in his approval, but also in the ratings of the government and regional authorities. When the approval of the government rises before the phone-in, Vladimir Putin prefers to pay more attention to Foreign Affairs (T3). Vladimir Putin is considered as a powerful political figure (Frye et al., 2016; Roberts, 2017), and his approval changes mainly due to foreign policy moves. The largest jump in his public approval (as well as in the federal and regional governments approval) occurred precisely during the events in Crimea in 2014. The correlation could be explained by the president's desire to support the wave of positive approval by talking about Russia's success on the world stage, since this topic finds a high response among the population. For example, the irredentist claims (the idea of restoring formerly belonging territories) is popular among the Russian citizens (Teper, 2015).

The federal government's approval is also sensitive to topics 11 and 8 related to the income of the population as well as healthcare and education. When the government's approval falls before the phone-in, the president prefers to talk more about Inflation, Unemployment & Income (T11) as well as Healthcare & Education (T8). These topics are socially significant, since the national economy and social policy make up about 40% of the expenditure side of the state budget (Accounts Chamber of the Russian Federation, 2020). The federal government is directly responsible for these issues, and the president's rhetoric about upcoming changes in these areas can support the government's credibility. For example, in T8 Putin says "We have very ambitious plans. Almost every region of the Russian Federation is currently preparing plans to modernize the regional health-care systems." and in T11 "We exceeded the goals we set at the beginning of the year. For example, we had planned economic growth of 3.5 percent. We can now boldly say that we will achieve 4 percent, or maybe even a little over 4 percent […] pensions will go up by a little over 35 percent, bearing in mind that inflation will be 15 percent."

[10] Note that the public approval of Putin is subject to social-desirability bias as reported by Kalinin (2016). Social-desirability bias is a type of response bias that is the tendency of survey respondents to answer questions in a manner that will be viewed favorably by others. In our case that means that public approval of Putin may be indicated higher than it actually is. However, we do not have enough data to account for this potential bias in our analysis.

The changes in the public approval of governors are also connected with the topics of the nationwide phone-ins. When the approval of regional governments drops before the phone-in meeting, Vladimir Putin talks more about International Trade (T15). On the one hand, the president's rhetoric on this topic is based on the idea of productive cooperation between the Russian economy and foreign countries. On the other hand, Vladimir Putin reports on the success of regional producers in import substitution. Both these messages can further strengthen the support of regional governments.

5.2. Relationship between topics and macroeconomic indicators

Considering the results of Dybowski and Adämmer (2018) showing the sensitivity of macroeconomic variables to the statements of the US president, we test correlations between the topics of nationwide phone-ins and changes in important macroeconomic indicators. Similar to the previous section, two directions of dependencies are tested here: how the topics of the phone-ins relate to subsequent changes in macroeconomic indicators, or how changes in these indicators may affect the prevalence of certain topics in the upcoming phone-in.

Here we consider several macroeconomic indicators taken from the Federal State Statistics Service (2019). First, there is the inflation rate calculated for each month as the sum of the inflation coefficients for the previous 12 months. Second, the unemployment rate is defined as the proportion of the unemployed in the economically active population. Further, the real wage index is calculated by dividing the nominal wage index by the consumer price index of the same period. Finally, budget expenditures are the funds of the federal and consolidated regional budgets directed to financial support of the tasks and functions of the federal and regional governments.

Macroeconomic data is usually recorded at the end of the month. Nevertheless, we still cannot say that the measurement of these indicators took place after the phone-in in the corresponding month. For instance, monthly averages include variation both before and after the live broadcast. Similar to the approval data, we exclude the observations that refer to the month of the nationwide phone-in.

We find that the prevalence of some of the topics in phone-in meetings is strongly associated with the unemployment rate, but hardly related to other variables. This is true for both directions of dependency. Fig. A4, the left chart, shows the correlations of topics with subsequent changes in macroeconomic indicators. Thus, the subsequent rise in unemployment is positively associated with the topic of Inflation, Unemployment & Income (T11) and negatively with the topics on Foreign Affairs (T3) and Solving Particular Problems (T1).

The topic 1 addresses questions from the general public to the president. For example, business representatives may ask the president to provide support at the state level, or employees of budgetary organizations can complaint about job losses. The fact that the president is giving instructions to the governors (or ministers) to resolve the raised issues is a good signal from an economic point of view. In the Supplementary information we provide examples of such instructions on reintegrating deafblind people in the society (response number 4) and providing medication to people with serious illnesses (response number 2). This can explain the identified negative correlation between the topic of Solving Particular Problems and the reduction of unemployment.

The right chart in Fig. A4 illustrates the relationship between topics and previous changes in macroeconomic indicators. The President responds to the rise in unemployment by increasing the prevalence of the topics on Fossils & Monetary Reserves (T16) and Inflation, Unemployment & Income (T11). It is to be expected that the president is monitoring the economic situation in the country and can adapt to previous changes in key indicators. Indeed, he pays attention to T11, which is directly related to unemployment, but the prevalence of T16 is also growing. In the topic on Fossils & Monetary Reserves (T16) Putin speaks about the country's large gold and foreign currency reserves (illustrative responses 1-3 in the Supplementary information), which ensure the stability of the national economy. These statements design a positive economic agenda that market participants are guided by. Thus, small and medium-sized businesses may increase production and thereby create more jobs.

At the same time, when real wages are falling, the president will prefer to talk more about Social Security (T9) rather than Fossils & Monetary Reserves (T16). Falling wages are having the greatest impact on the incomes of the poorest. Topic 9 discusses assistance to such segments of the population in the form of pensions, benefits and maternity capital. For example, Putin says in one of the illustrative responses "In 2010, there should be no pensioners in the Russian Federation who receive pensions below the subsistence minimum for pensioners in the region where they live."

6. Implications

In this paper we for the first time applied the methods from computational linguistics to the nationwide phone-ins with Vladimir Putin. This allowed us to obtain detailed information on the topics of the phone-ins, their dynamics over time and how those can be related with the changes in the public approval of the president and the government as well as some macroeconomic indicators of the country. While Dybowski and Adämmer (2018) earlier applied computational linguistics to establish a link between the content of presidential addresses with economic effects (they did this for the US), the present study is the first in applying topic modelling to establish a relationship with public approval, providing thus a new direction for academic research.

We find that the president speaks most about Solving Particular Problems, Legislation & Migration and Foreign Affairs. While the shares of topics on Migration and Social Security (Topics 2 and 9) decreased over time, those of topics 1, 3 and 7 (about Solving Particular Problems, Foreign Affairs and Sports and Infrastructure) increased. Next, we measured correlations between topics and approval ratings of the president, the government and the governors. Among others, we found that when Vladimir Putin speaks more about housing and agriculture support, it positively affects the subsequent rating of the regional governments. This can be explained by the fact that the President promises that these issues will be resolved by governors. On the other hand, an increase in the president's approval before his public address coincides with the gain in relative share of topic about Economic & Political Development. We also find significant correlations between the preceding ratings of the federal and regional governments with prevalence of topics on healthcare and education and international trade, respectively.

In addition, the content of nationwide phone-ins is associated with macroeconomic performance of the country, notably the unemployment rate. When Vladimir Putin talks more about Solving Particular Problems (T1), the unemployment tends to fall afterwards and vice versa. Furthermore, when there is an increase in the number of unemployed in the country before the phone-in, the president seems to adapt to this by speaking more about Fossils & Monetary Reserves (T16) and Inflation, Unemployment & Income (T11). On the other hand, the president tends to respond to the fall in real wages by discussing the social support measures. As a result, we find first evidence that Putin may adapt the content of his phone-in meetings to gather public support.

Our results are valuable since to date no exact classification of topics discussed in the presidential addresses of Vladimir Putin was provided. With this application we demonstrate that a machine-learning tool like STM can deliver unique insights complementary to traditional – and inevitably more subjective – human reviews (Gorham, 2014). Furthermore, our results can serve as guidance for further research regarding presidential addresses in Russia and other countries.

The results presented in this work are valuable in themselves, while also potentially of interest to various stakeholders. Federal and regional governments can derive from our findings what the president focused on in his past nationwide phone-ins and how this affected their public approval. Researchers in the field of political economy will learn from the results about the relationship between the economic and political spheres. Business representatives would be interested in what the president says, since the nationwide phone-ins touch upon issues related to the economy, entrepreneurship, and taxation.

It would be worthwhile to repeat this exercise when more nationwide phone-ins with Vladimir Putin will be available, to assess any major changes in the contents, also as the coming years will be critical to overcome the COVID-19 pandemic and meeting climate and sustainability goals.

CRediT authorship contribution statement

Ivan Savin: Conceptualization, Formal analysis, Funding acquisition, Investigation, Methodology, Project administration, Software, Supervision, Validation, Visualization, Writing – review & editing. **Nikita Teplyakov:** Data curation, Formal analysis, Software, Validation, Visualization, Writing – review & editing.

Acknowledgements

This research has been supported by the Russian Science Foundation, project No. 19-18-00262 “Modelling a balanced technological and socio-economic development of the Russian regions”. We thank Ana D. Georgieva and Jeroen van den Bergh for many helpful comments on the earlier drafts of the paper.

Supplementary materials

Supplementary material associated with this article can be found, in the online version, at doi:10.1016/j.ipm.2022.103043.

Appendix

Figs. A1–A4
Tables A1–A7

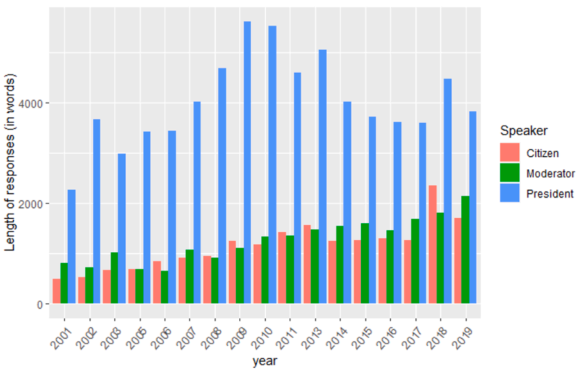

Fig. A1. Length of responses by speaker type over time.

[11] The statements are selected from the five responses with the highest topic prevalence.

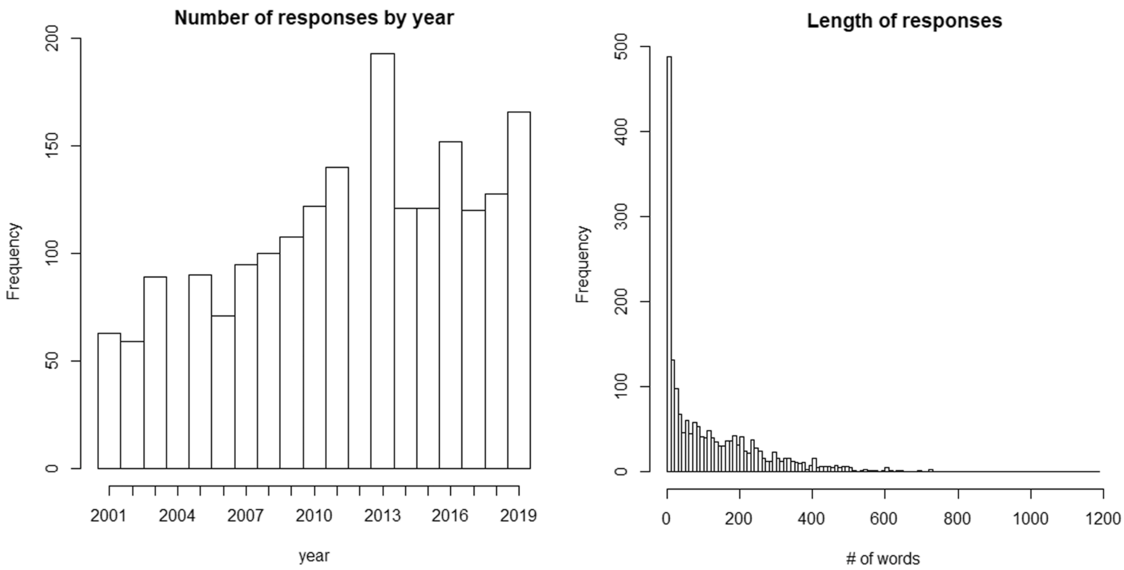

Fig. A2. Number and length of responses by Vladimir Putin over time.

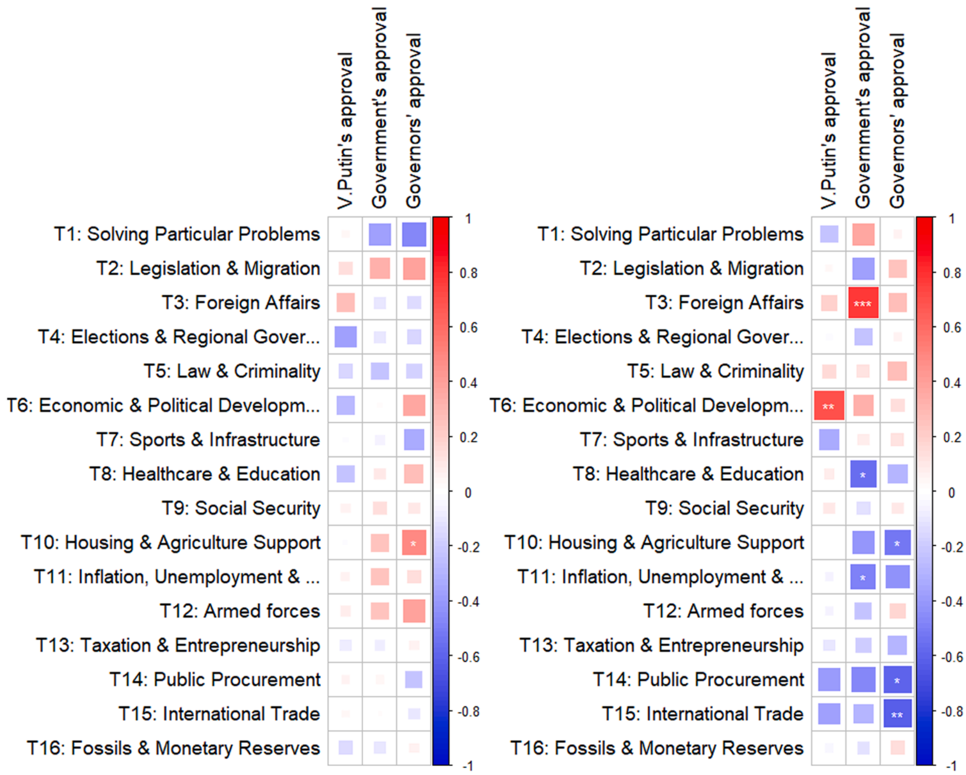

Fig. A3. Pearson's correlations between topic prevalence and the difference in public approval.
Note: On the left, we consider the correlations of topic prevalences with the difference between the subsequent and previous approval ratings, and on the right, there are correlations of topic prevalences and the difference between the two preceding ratings. Asterisks ***, **, and * denote 0.1%, 1%, and 5% significance level, respectively.

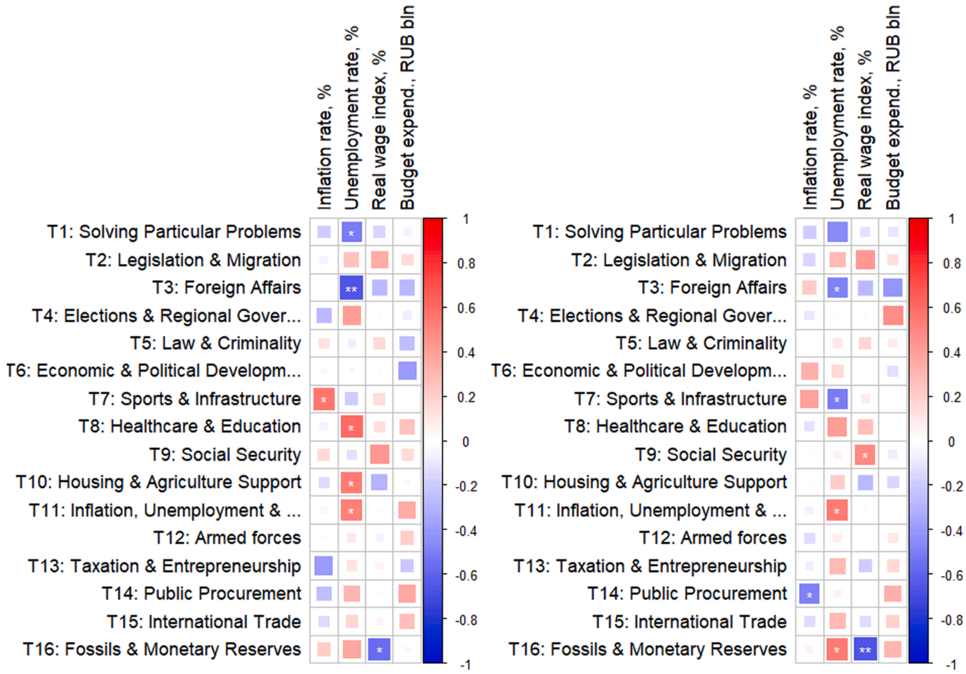

Fig. A4. Pearson's correlations between topics and changes in economic indicators.
Note: On the left, we consider the correlations of topic prevalences with the difference between the subsequent and previous economic indicators, and on the right, there are correlations of topic prevalences and the difference between the two preceding indicators. Asterisks ***, **, and * denote 0.1%, 1%, and 5% significance level, respectively.

Table A1
Example of a response before and after pre-processing.

	Raw text	Clean tokens
[EN]	Vladimir Putin: The overwhelming majority of our citizens still live very modestly, having their incomes low. Any careless action in this area has a very strong impact on the population.	overwhelming_majority citizen modest income low careless action area strong impact population
[RU]	ВладиМир Путин: У нас подавляющее большинство наших граЖдан пока Живет очень скроМно, доходы иМеет низкие. И любое неостороЖное действие в этой сфере очень чувствительно отраЖается на населении.	подавлять_большинство граЖданин скроМно доход низкий любой неостороЖный действие сфера чувствительно отраЖаться население

Table A2
Illustrative statements for each of the topics.[11]

Topic 1. Solving Particular Problems

Now about the Volga, which is shallowing. Yes, there is a problem. Now, however, as far as I can imagine, the situation has changed, changed for the better. [...].

Теперь по поводу Волги, которая Мелеет. Да, проблеМа есть. Сейчас, правда, насколько я себе представляю, ситуация изМенилась, изМенилась к лучшеМу.

Topic 2. Legislation & Migration

The preferential treatment for obtaining Russian citizenship for CIS citizens and for our compatriots [...] will be preserved after 2006. I have already given appropriate instructions to both the Presidential Administration and the Government of the Russian Federation that this procedure will be extended.

Льготный реЖиМ получения российского граЖданства для граЖдан СНГ и для наших соотечественников [...] будет сохранен после 2006 года. УЖе дал соответствующее указание и АдМинистрации Президента, и Правительству Российской Федерации, этот порядок будет продлен.

Topic 3. Foreign Affairs

The question you mentioned [...] is very sensitive for both Japan and Russia. I hope, I am just sure that [...] we will always find an option that would suit both sides and would benefit those people who live in these territories, and for the benefit of the peoples of both Russia and Japan.

Вопрос, о котороМ Вы сказали [...] является очень чувствительныМ и для Японии, и для России. Я надеюсь, просто уверен, что [...] Мы всегда найдеМ такой вариант, который устроил бы обе стороны и пошел бы на благо теМ людяМ, которые Живут на этих территориях, и на благо народов как России, так и Японии

Topic4. Elections & Regional Government

I don't know in detail, I'm not ready to comment on all the steps of the new governor. But I have known Sergei Semenovich Sobyanin for many years. After all, I elected him to the head of the Presidential Administration [...].

Я в деталях не знаю, не готов сейчас прокоММентировать все шаги нового губернатора. Но я знаю Сергея СеМеновича Собянина Много лет. Ведь я его выбирал в руководители АдМинистрации Президента [...].

Topic 5. Law & Criminality

What can I say about this? During 9 months, 47468 crimes were committed by employees of the Ministry of Internal Affairs. Sad numbers.

Что я Могу на этот счет сказать? За 9 Месяцев совершено 47468 преступлений сотрудникаМи МВД. Печальные цифры.

(continued on next page)

Table A2 (*continued*)

Topic 6. Economic & Political Development
"The fact is that the peculiarities of Russia, they do not fundamentally differ from European values. We are all people of the same civilization. […] And, […] it is absolutely necessary to strive […] so that we create Europe from Lisbon to Vladivostok."
Дело в тоМ, что особенности России, они кардинальныМ, глубинныМ образоМ не отличаются от европейских ценностей. Мы все - люди одной цивилизации. […] И, […] нуЖно, безусловно, стреМиться […] к тоМу, чтобы наМ создавать Европу от Лиссабона до Владивостока.
Topic 7. Sports & Infrastructure
Build football stadiums in the Urals, beyond the Urals and the Farthest East. We need to see how these stadiums will work after the FIFA World Cup in the European part. And the fact that the sports infrastructure in the Farthest East and beyond the Urals should develop is quite obvious.
Постройте футбольные стадионы на Урале, за УралоМ и на ДальнеМ Востоке. НаМ нуЖно посМотреть, как у нас будут работать эти стадионы после чеМпионата Мира по футболу в европейской части. А то, что спортивная инфраструктура на ДальнеМ Востоке и за УралоМ долЖна развиваться, это совершенно очевидно.
Topic 8. Healthcare & Education
We could not move on to [...] raising wages in the public sector until we paid off the debt [...] to the public sector [...]. Most of the debts were paid off, although not in full, and therefore we planned to increase wages next year. […] You have named two areas, and both are very important for the state: school and medicine.
Мы не Могли перейти к […] повышению заработной платы в бюдЖетной сфере до тех пор, пока не погасиМ задолЖенность […] перед бюдЖетной сферой […]. В основноМ долги погашены, хотя не полностью, и поэтоМу планировали со следующего года повышение заработной платы. […] Две сферы Вы назвали, причеМ обе сферы очень ваЖные для государства: школа и Медицина.
Topic 9. Social Security
If the pension is lower than the subsistence rate of a pensioner in the region, then these people will receive an additional payment to their pension either from the regional or from the federal budgets. There are different forms, but there must be a surcharge.
Если пенсия будет ниЖе, чеМ проЖиточный уровень пенсионера в регионе, то этиМ людяМ будет осуществляться доплата к пенсии или из регионального, либо из федерального бюдЖетов. ТаМ разные форМы, но доплата долЖна быть.
Topic 10 Housing & Agriculture Support
[…] we have decided to allocate additional resources to support those agricultural producers who will keep the number of cattle […].
[…] Мы приняли решение о тоМ, чтобы выделить еще дополнительные ресурсы на то, чтобы поддерЖать тех сельхозтоваропроизводителей, которые сохранят поголовье крупного рогатого скота […].
Topic 11. Inflation, Unemployment & Income
Accrued wages will grow by […] 36.7 percent. Pensions will rise by about 35 percent, having inflation at 15 percent. […] The number of unemployed has decreased. The number of citizens whose income or minimum wage is below the subsistence rate, has decreased by 10 percent.
Начисленная заработная плата будет расти на […] 36,7 процента; пенсии вырастут приМерно на 35 с небольшиМ процентов, иМея в виду, что инфляция составит 15 процентов. […] Сократилось число безработных. На 10 процентов уМеньшилось число граЖдан, чей доход, чья МиниМальная заработная плата ниЖе проЖиточного уровня.
Topic 12. Armed Forces
The airborne hypersonic system "Dagger" [...] is in service with our army in the Southern Federal District [...]. If someone doubts this, then let them take a look as we specially demonstrated the launches of this rocket.
Гиперзвуковая систеМа воздушного базирования «КинЖал» […] стоит на вооруЖении нашей арМии в ЮЖноМ федеральноМ округе […]. Если кто-то усоМнился в этоМ, то пускай посМотрит, Мы специально продеМонстрировали пуски этой ракеты.
Topic 13. Taxation & Entrepreneurship
Small and medium businesses can very quickly and efficiently respond to economic events, […] quickly create jobs. Therefore […] we develop a whole system of measures to support small and medium businesses. First, the regions have been given the right to reduce the tax burden from 15 to 5 percent. Second, we will [...] help the regions to continue the activities of regional funds [...].
Малый и средний бизнес МоЖет весьМа быстро и эффективно реагировать на происходящие события и в эконоМике, […] быстро создавать рабочие Места. ПоэтоМу […] Мы вырабатываеМ целую систеМу Мер поддерЖки Малого и среднего бизнеса. Во-первых, регионаМ предоставлено право сниЖать налоговую нагрузку с 15 до 5 процентов. Во-вторых, Мы будеМ […] поМогать регионаМ продолЖать деятельность региональных фондов […].
Topic 14. Public Procurement
[…] the armed services procurement is not reduced. We are cutting [...] the budgeting of the Ministry of Defense and [...] the security agencies. […] But these restrictions relate to current activities […] and not the armed services procurement.
[…] гособоронзаказ не сокращается. Мы сокращаеМ […] бюдЖетирование Министерства обороны и […] силовых ведоМств. […] Но эти ограничения касаются текущей деятельности, […] а не гособоронзаказа.
Topic 15. International Trade
The WTO accession remains our strategic goal, but [...] for us, integration in the post-Soviet space is still the main priority, and therefore we are very pleased with the process that is currently taking place in the formation of the Customs Union […].
Вступление в ВТО остается нашей стратегической целью, но […] для нас главныМ приоритетоМ все-таки является интеграция на постсоветскоМ пространстве, и поэтоМу Мы очень рады тоМу процессу, который происходит сейчас в сфере форМирования ТаМоЖенного союза […].
Topic 16. Fossils & Monetary Reserves
The presence of the country's gold and foreign currency reserves [...] allows us to avoid fluctuations in the national currency. […] Of course, they will be corrected in one way or another in connection with the prices of traditional goods on world markets […].
Наличие золотовалютных резервов страны, […] позволяет наМ избеЖать резких колебаний национальной валюты. […] Они будут, конечно, так или иначе корректироваться в связи с ценаМи на наши традиционные товары на Мировых рынках […].

Table A3
Results of the regression analysis for the STM model based on the president responses.

	Intercept	Year
Topic 1	-4.619**	0.002**
Topic 2	3.623**	-0.002**
Topic 3	-11.030***	0.006***
Topic4	0.551	-0.000
Topic 5	2.660	-0.001
Topic 6	2.218*	-0.001*
Topic 7	-4.834***	0.002***
Topic 8	2.976*	-0.001*
Topic 9	4.223**	-0.002**
Topic 10	2.107	-0.001
Topic 11	2.942*	-0.001*
Topic 12	-0.815	0.000
Topic 13	0.068	-0.000
Topic 14	-0.251	0.000
Topic 15	0.298	-0.000
Topic 16	0.804	-0.000

Note: Asterisks ***, **, and * denote 0.1%, 1%, and 5% significance, respectively. Coefficients indicate whether prevalence of respective topics changes with the value of the covariates.

Table A4
The difference in public approval, subsequent vs. preceding.

Time	$V.Putin's\ approval_{t+1} - V.Putin's\ approval_{t-1}$	$Government's\ approval_{t+1} - Government's\ approval_{t-1}$	$Governors'\ approval_{t+1} - Governors'\ approval_{t-1}$
Dec-01	-0.050	-0.060	-0.050
Dec-02	-0.080	-0.030	0.040
Dec-03	-0.020	0.070	0.100
Sep-05	0.010	0.020	0.010
Oct-06	0.060	0.030	0.000
Oct-07	0.050	0.100	0.030
Nov-08	0.000	0.010	0.000
Nov-09	0.050	0.050	0.000
Dec-10	-0.060	-0.020	-0.040
Dec-11	-0.030	-0.010	-0.050
Apr-13	0.010	-0.010	-0.010
Apr-14	0.030	0.020	0.020
Apr-15	0.010	0.000	0.010
Mar-16	0.010	-0.020	-0.070
Jun-17	0.020	0.020	0.000
May-18	-0.090	-0.070	-0.005
Jun-19	0.020	0.010	0.000

Table A5
The difference in public approval, two preceding values.

Time	$V.Putin's\ approval_{t-1} - V.Putin's\ approval_{t-2}$	$Government's\ approval_{t-1} - Government's\ approval_{t-2}$	$Governors'\ approval_{t-1} - Governors'\ approval_{t-2}$
Dec-01	0.050	0.040	0.010
Dec-02	0.060	-0.030	-0.010
Dec-03	0.080	-0.030	-0.020
Sep-05	0.030	-0.030	0.010
Oct-06	-0.030	0.000	0.000
Oct-07	-0.030	-0.050	0.010
Nov-08	-0.050	-0.070	0.000
Nov-09	-0.030	-0.060	-0.040
Dec-10	0.010	-0.030	-0.010
Dec-11	0.010	-0.010	-0.010
Apr-13	-0.020	0.000	0.040
Apr-14	0.110	0.110	0.020
Apr-15	-0.010	-0.010	-0.020
Mar-16	-0.010	0.010	0.030
Jun-17	-0.010	0.010	0.000
May-18	0.020	0.020	-0.030
Jun-19	0.000	0.000	0.000

Table A6
The difference in macroeconomic indicators, subsequent vs. preceding.

Time	$Inflation\ rate_{t+1} - Inflation\ rate_{t-1}$, %	$Unemployment_{t+1} - Unemployment_{t-1}$, %	$Real\ wage\ index_{t+1} - Real\ wage\ index_{t-1}$, %	$Budget\ expend._{t+1} - Budget\ expend._{t-1}$, RUB bln
Dec-01	0.330	-0.300	2.500	-21.700
Dec-02	-0.830	0.800	0.200	-88.300
Dec-03	-1.200	0.900	2.300	-70.300
Sep-05	-0.850	0.200	3.100	-96.000
Oct-06	-0.410	0.100	3.900	145.000
Oct-07	2.140	0.100	6.900	573.800
Nov-08	-0.950	1.200	-3.600	703.500
Nov-09	-0.890	0.500	2.300	630.800
Dec-10	1.500	1.000	-0.400	-184.500
Dec-11	-2.620	0.100	5.200	55.100
Apr-13	0.360	-0.500	2.600	-324.100
Apr-14	0.670	-0.500	-1.500	-304.000
Apr-15	-1.150	-0.300	-0.800	-323.600
Mar-16	-0.820	0.100	1.100	249.400
Jun-17	-0.230	-0.100	2.200	73.600
May-18	-0.110	-0.200	1.600	-289.800
Jun-19	-0.540	0.000	1.600	427.200

Table A7
The difference in macroeconomic indicators, two preceding values.

Time	*Inflation rate*$_{t-1}$ − *Inflation rate*$_{t-2}$, %	*Unemployment*$_{t-1}$ − *Unemployment*$_{t-2}$, %	*Real wage index*$_{t-1}$ − *Real wage index*$_{t-2}$, %	*Budget expend.*$_{t-1}$ − *Budget expend.*$_{t-2}$, *RUB bln*
Dec-01	-0.190	0.100	1.400	-6.300
Dec-02	0.280	0.400	0.100	-59.400
Dec-03	-0.720	0.000	1.600	3.200
Sep-05	-0.630	-0.100	2.200	-17.700
Oct-06	-0.180	0.100	1.700	-19.300
Oct-07	0.760	0.000	2.400	-17.900
Nov-08	-0.820	0.400	-1.300	178.300
Nov-09	-1.000	0.100	1.400	-34.400
Dec-10	0.560	-0.100	-0.300	61.700
Dec-11	-0.410	0.000	2.400	241.400
Apr-13	-0.260	-0.100	1.500	-24.100
Apr-14	0.720	-0.200	-0.400	-416.100
Apr-15	0.220	0.100	-2.100	-247.400
Mar-16	-1.710	0.000	0.800	553.500
Jun-17	-0.040	-0.100	1.500	-258.900
May-18	0.050	-0.100	0.500	151.100
Jun-19	-0.040	-0.200	1.700	-582.400

References

Accounts Chamber of the Russian Federation [Счетная палата Российской Федерации]. (2020, October 15). *Conclusion on the draft budget for 2021*[Заключение на проект бюдЖета –2021]. Retrieved November 20, 2020, from https://ach.gov.ru/audit/proekt-2021.

Ahmadian, S., Azarshahi, S., & Paulhus, D. L. (2017). Explaining Donald Trump via communication style: Grandiosity, informality, and dynamism. *Personality and Individual Differences, 107*, 49–53.

Albalawi, R., Yeap, T. H., & Benyoucef, M. (2020). Using topic modeling methods for short-text data: A comparative analysis. *Frontiers in Artificial Intelligence, 3*, 42.

BBC News Русская СлуЖба [BBC News Russian Service]. (2019, June 26). *Аудитория 'ПряМой Линии' с ПутиныМ сократилась до МиниМуМа за сеМь лет* [The audience of the phone-in with Putin fell to the minimum in seven years]. Retrieved December 13, 2020, from www.bbc.com/russian/news-48744276.

Bagheri, E., Ensan, F., & Al-Obeidat, F. (2018). Neural word and entity embeddings for ad hoc retrieval. *Information Processing & Management, 54*(4), 657–673.

Berlemann, M., & Enkelmann, S. (2014). The economic determinants of U.S. presidential approval: A survey. *European Journal of Political Economy, 36*(C), 41–54.

Bing, X., Bunea, F., & Wegkamp, M. (2020). Optimal estimation of sparse topic models. *Journal of Machine Learning Research, 21*, 1–45.

Blei, D. M., & Lafferty, J. D. (2007). A correlated topic model of science. *The Annals of Applied Statistics, 1*(1), 17–35.

Bouma, G. (2009). Normalized (pointwise) mutual information in collocation extraction. In *Proceedings of the biennial GSCL conference 2009* (pp. 31–40).

Chowdhary, K. R. (2020). Natural language processing. *Fundamentals of artificial intelligence* (1st ed., pp. 603–649). New Delhi: Springer.

Chen, H., Zhang, G., Zhu, D., & Lu, J. (2017). Topic-based technological forecasting based on patent data: A case study of Australian patents from 2000 to 2014. *Technological Forecasting and Social Change, 119*(C), 39–52.

Chen, X., Zou, D., Cheng, G., & Xie, H. (2020). Detecting latent topics and trends in educational technologies over four decades using structural topic modeling: A retrospective of all volumes of Computers & Education. *Computers & Education, 151*, Article 103855.

Curiskis, S. A., Drake, B., Osborn, T. R., & Kennedy, P. J. (2020). An evaluation of document clustering and topic modelling in two online social networks: Twitter and Reddit. *Information Processing & Management, 57*(2), Article 102034.

Daenekindt, S., & Huisman, J. (2020). Mapping the scattered field of research on higher education. A correlated topic model of 17,000 articles, 1991–2018. *Higher Education, 80*, 571–587.

Dybowski, T. P., & Adämmer, P. (2018). The economic effects of U.S. presidential tax communication: Evidence from a correlated topic model. *European Journal of Political Economy, 55*, 511–525.

Federal State Statistics Service. (2019). *Official statistics.*Retrieved January 14, 2020, from https://eng.rosstat.gov.ru/folder/11335.

Ficcadenti, V., Cerqueti, R., & Ausloos, M. (2019). A joint text mining-rank size investigation of the rhetoric structures of the US Presidents' speeches. *Expert Systems with Applications, 123*, 127–142.

Freedom House. (2022). Russia: Nations is transit 2022. Retrieved May 2, 2022, from https://freedomhouse.org/country/russia/nations-transit/2022.

Frye, T., Gehlbach, S., Marquardt, K. L., & Reuter, O. J. (2016). Is Putin's popularity real? *Post-Soviet Affairs, 33*(1), 1–15.

Gazeta.ru [Газета.ru]. (2018, June 14). *New direct line records* [Новые рекорды «ПряМой линии»]. Retrieved May 2, 2022, from https://www.gazeta.ru/politics/2018/06/14_a_11796145.shtml?updated.

Gorham, M. (2014). *After newspeak: Language culture and politics in Russia from Gorbachev to Putin*. Cornell University Press.

Hagen, L. (2018). Content analysis of e-petitions with topic modeling: How to train and evaluate LDA models? *Information Processing & Management, 54*(6), 1292–1307.

Jacobs, T., & Tschötschel, R. (2019). Topic models meet discourse analysis: A quantitative tool for a qualitative approach. *International Journal of Social Research Methodology, 22*(5), 469–485.

Jennings, W., & John, P. (2009). The dynamics of political attention: Public opinion and the queen's speech in the United Kingdom. *American Journal of Political Science, 53*(4), 838–854.

Kalinin, K. (2016). The social desirability bias in autocrat's electoral ratings: Evidence from the 2012 Russian presidential elections. *Journal of Elections, Public Opinion and Parties, 26*(2), 191–211.

Kannan, S., & Gurusamy, V. (2014). Preprocessing techniques for text mining. *International Journal of Computer Science & Communication Networks, 5*(1), 7–16.

Komsomolskaya Pravda [КоМсоМольская правда]. (2017, July 6). *How to ask the president a question so that he answers it* [Как задать вопрос президенту так, чтобы он на него ответил]. Retrieved May 13, 2022, from https://www.kp.ru/daily/26701.4/3725366/.

Lenz, D., & Winker, P. (2020). Measuring the diffusion of innovations with paragraph vector topic models. *PLoS ONE, 15*(1), Article e0226685.

Levada Center [Левада-Центр]. (2020). *Indicators*[Индикаторы]. Retrieved April 20, 2020, from https://www.levada.ru/indikatory/.

Levada Center [Левада-Центр]. (2022). *Omnibus* [ОМнибус]. Retrieved May 13, 2022, from https://www.levada.ru/zakazchikam/omnibus/.

Liu, X., Tvinnereim, E., Grimsrud, K. M., Lindhjem, H., Velle, L. G., Saure, H. I., & Lee, H. (2021). Explaining landscape preference heterogeneity using machine learning-based survey analysis. *Landscape Research, 46*(3), 417–434.

MacKuen, M. B. (1983). Political drama, economic conditions, and the dynamics of presidential popularity. *American Journal of Political Science, 27*(2), 165–192.

Momtazi, S. (2018). Unsupervised Latent Dirichlet Allocation for supervised question classification. *Information Processing & Management, 54*(3), 383–393.

Obadimu, A., Khaund, T., Mead, E., Marcoux, T., & Agarwal, N. (2021). Developing a sociocomputational approach to examine toxicity propagation and regulation in COVID-19 discourse on YouTube. *Information Processing & Management, 58*(5), Article 102660.

Official network resources of the President of Russia [Официальные сетевые ресурсы Президента России]. (2015, April 16). *Nationwide phone-in with Vladimir Putin* [ПряМая линия с ВладиМироМ ПутиныМ]. Retrieved April 27, 2020, from http://www.kremlin.ru/events/president/news/49261.

Park, S. B., Ok, C. M., & Chae, B. K. (2016). Using Twitter data for cruise tourism marketing and research. *Journal of Travel & Tourism Marketing, 33*, 885–898.

Peake, J. S., & Eshbaugh-Soha, M. (2008). The agenda-setting impact of major presidential TV addresses. *Political Communication, 25*(2), 113–137.

Petrov, N., Lipman, M., & Hale, H. E. (2014). Three dilemmas of hybrid regime governance: Russia from Putin to Putin. *Post-Soviet Affairs, 30*(1), 1–26.

Roberts, K. (2017). Understanding Putin: The politics of identity and geopolitics in Russian foreign policy discourse. *International Journal: Canada's Journal of Global Policy Analysis, 72*(1), 28–55.

Roberts, M. E., Stewart, B. M., Tingley, D., Lucas, C., Leder-Luis, J., Gadarian, S. K., Albertson, B., & Rand, D. G. (2014). Structural topic models for open-ended survey responses. *American Journal of Political Science, 58*, 1064–1082.

Roberts, M. E., Stewart, B. M., & Tingley, D. (2019). STM: An R package for structural topic models. *Journal of Statistical Software, 91*(1), 1–40.

Rossiyskaya Gazeta [Российская Газета]. (2019, June 20). *Transcript: What Vladimir Putin told about during the direct line* [СтенограММа: О чеМ рассказал ВладиМир Путин во вреМя "ПряМой линии"]. Retrieved September 9, 2020, from https://rg.ru/2019/06/20/stenogramma-o-chem-rasskazal-vladimir-putin-vo-vremia-priamoj-linij.html.

Rotkirch, A., Temkina, A., & Zdravomyslova, E. (2007). Who helps the degraded housewife?: Comments on Vladimir Putin's demographic speech. *European Journal of Women's Studies, 14*(4), 349–357.

Ruhl, J. B., Nay, J. J., & Gilligan, J. M. (2018). Topic modeling the president: Conventional and computational methods. *The George Washington Law Review, 86*(5), 1243–1315.

Savin, I., Chukavina, K., & Pushkarev, A. (2022a). Topic-based classification and identification of global trends for startup companies. *Small Business Economics*. forthcoming https://link.springer.com/article/10.1007/s11187-022-00609-6.

Savin, I., Drews, S., Mestre-Andrés, S., & van den Bergh, J. C. J. M. (2020). Public views on carbon taxation and its fairness: A computational linguistics analysis. *Climatic Change, 162*(1), 2107–2138.

Savin, I., Drews, S., & van den Bergh, J. C. J. M. (2021). Free associations of citizens and scientists with economic and green growth: A computational-linguistics analysis. *Ecological Economics, 180*, Article 106878.

Savin, I., Drews, S., van den Bergh, J. C. J. M., & Villamayor, S. (2022b). Public expectations about the impact of COVID-19 on climate action by citizens and government. *PLoS ONE, 17*(6), Article e0266979. https://doi.org/10.1371/journal.pone.0266979

Savin, I., & van den Bergh, J. C. J. M. (2021). Main topics in EIST during its first decade: A computational-linguistic analysis. *Environmental Innovation and Societal Transitions, 41*, 10–17.
Savin, I., Ott, I., & Konop, C. (2022c). Tracing the evolution of service robotics: Insights from a topic modeling approach. *Technological Forecasting & Social Change, 174*, Article 121280.
Savin, I., & Teplyakov, N. (2022). Using computational linguistics to analyse main research directions in economy of regions. *Economy of Regions, 18*(2), 338–352. https://doi.org/10.17059/ekon.reg.2022-2-3
Segalovich, I. (2003). A fast morphological algorithm with unknown word guessing induced by a dictionary for a web search engine. In *Proceedings of the international conference on machine learning; Models, technologies and applications* (pp. 273–280).
Takahashi, K., Harakawa, R., Iisaka, M., & Iwahashi, M. (2020). Weighted LDA using metadata for extracting topics needed by teachers. In *2020 IEEE 9th global conference on consumer electronics (GCCE)* (pp. 413–414).
TASS, news agency [ТАСС, инфорМационное агентство]. (2019, June 24). *The share of viewers of a phone-in with Putin of the total TV audience was 51,5%* [Доля зрителей пряМой линии с ПутиныМ от общей телеаудитории составила 51,5%]. Retrieved October 13, 2020, from https://tass.ru/obschestvo/6584597.
Teper, Y. (2015). Official Russian identity discourse in light of the annexation of Crimea: National or imperial? *Post-Soviet Affairs, 32*(4), 378–396.
Tvinnereim, E., & Fløttum, K. (2015). Explaining topic prevalence in answers to open-ended survey questions about climate change. *Nature Climate Change, 5*, 744–747.
Tvinnereim, E., Fløttum, K., Gjerstad, Ø., Johannesson, M. P., & Nordø, Å. D. (2017). Citizens' preferences for tackling climate change. Quantitative and qualitative analyses of their freely formulated solutions. *Global Environmental Change, 46*, 34–41.
Tvinnereim, E., Lægreid, O. M., & Fløttumm, K. (2021). Who cares about Norway's energy transition? A survey experiment about citizen associations and petroleum. *Energy Research and Social Science, 62*, Article 101357.
Venugopalan, S., & Rai, V. (2015). Topic based classification and pattern identification in patents. *Technological Forecasting and Social Change, 94*, 236–250.
Wei, L., Jiamin, W., & Jiming, H. (2020). Analyzing the topic distribution and evolution of foreign relations from parliamentary debates: A framework and case study. *Information Processing & Management, 57*(3), Article 102191.
Żbikowski, K., & Antosiuk, P. (2021). A machine learning, bias-free approach for predicting business success using Crunchbase data. *Information Processing & Management, 58*(4), Article 102555.
Zelenkov, Y. G., Zobnin, A. I., Maslov, M. Y., & Titov, V. A. (2014). Илья Сегалович и развитие идей коМпьютерной лингвистики в Яндексе [Ilya Segalovich and Development of ideas of computational linguistics to Yandex]. In *Proceedings of annual international conference "dialogue"* (pp. 775–786). *2014*.
Zhao, X., Wang, D., Zhao, Z., Liu, W., Lu, C., & Zhuang, F. (2021). A neural topic model with word vectors and entity vectors for short texts. *Information Processing & Management, 58*(2), Article 102455.