Media Sentiment Analysis for Measuring Perceived Trust in Government

Vladislav Petkevič

Amsterdam University College, Science Park 113, 1098 XG Amsterdam, petkevich.vladislav@gmail.com

Anotation. This study employs sentiment analysis of articles published in popular Lithuanian media outlets in an attempt to determine the citizens’ trust in the Lithuanian government at a given time. Based on the classical political theory by Thomas Hobbes and contemporary political theories of divisionary politics and the “rally ‘round the flag” effect, it is hypothesized that with an increased sense of insecurity, the population puts more trust into their government. The hypothesis is tested by examining the sentiment expressed in articles concerning governmental affairs published in Lithuanian media and the user-generated comments published on Facebook for the period of six months prior to and after March 2014 – Russia’s annexation of Crimea, an event that created a sense of fear among the Lithuanian public. The analysis of the gathered data found that the decrease in sentiment towards Russia correlated with an increase in the sentiment towards the government of Lithuania in conventional media. The analysis of Facebook comments suggested the same result, but overall was inconclusive.

Keywords: sentiment analysis, media effects, Lithuania-Russia relations, government trust.

Introduction

This paper attempts to answer the question of whether an external threat increases trust in the governmental institutions in a democratic setting. The Republic of Lithuania is taken as an exemplar due to its relatively new democratic rule and the recent worsening of diplomatic relations with Russia. To answer the question, an analysis of opinions about the government of Lithuania expressed in social and mass media is conducted for the period of six months before and six months after March, 2014 – the time when Russia annexed Crimea, a part of Ukraine, resulting in significantly heightened fear among the
Lithuanian citizens. This is achieved through the use of semantic analysis – a method of determining a text’s emotional charge through automatic examination of the words used in it. The method is applied to the articles extracted from three of the most read (online) Lithuanian newspapers and from commentary left by the general population on social media portal Facebook to evaluate the attitude towards the Lithuanian government expressed in them. It is then examined whether the trust in government is influenced by the amount of threat perceived by the citizens of Lithuania, which is hypothesized based on Russia’s actions in Ukraine and Crimea and measured using semantic analysis of the negativity expressed towards Russia. The effect is hypothesized to take place based on Thomas Hobbes’ political theory and more recent concurrent developments in the field of political science such as divisionary politics and the “rally ‘round the flag” effect which draw a direct link between a state’s security issues and the citizens’ opinion of it. To measure the proposed effects, Bayesian time-series and vector autoregression models are employed. Finally, the qualitative analysis of the analyzed data is presented and conclusions are drawn.

Political Theory Background

This paper adopts a classical political theory postulated by Thomas Hobbes in “The Leviathan” (Hobbes & Gaskin, 1998) in combination with a contemporary congruent theory of divisionary politics (Smith, 1996) and the “rally ‘round the flag” effect” (Muller, 1970). According to Hobbes, the lives of individuals in the state of nature were “solitary, poor, nasty, brutish and short” (Hobbes & Gaskin, 1998), as everyone lived in constant danger due to lack of any laws and peoples’ ability to inflict suffer on each other at will. To escape this misery, people would come together to form a “social contract” – an occurrence during which individuals chose to cede some of their individual rights so that others would cede theirs (e.g. person A gives up his/her right to kill person B if person B does the same). This resulted in the establishment of the state, a sovereign entity which would create laws to regulate social interactions. Human life was thus no longer “a war of all against all”. What is of interest here is that Hobbes infers that people would give some of their individual rights up only if their protection was ensured. Hence, the state’s legitimacy is completely dependent on its ability to guarantee the security of its citizens. Then, assuming that people join the sovereign (and thus give certain rights up) in order to ensure their safety, the absence or presence of a perceived threat would (partially) determine their view on usefulness/legitimacy of the state. This assumption is supported by more recent developments in political science and international relations: diversionary foreign policy theory and the adjacent ‘rally ‘round the flag’ effect (Mueller, 1970). The theory suggests that the state leader(s) divert citizens’ attention to foreign policy issues to create a distraction from domestic problems. Its effectiveness is perhaps best demonstrated by
the rally ‘round the flag effect: an increase on popularity of political leaders at times of conflict. A study that analyzed presidential popularity from 1945 to 1969 in relation to real-world political events suggested that spikes in presidential popularity follow international events that can be perceived as a kind of threat to the country. Contemporary research, too, suggests a similar effect. In their book “The persuadable voter” Hillygus and Shields (2014) argue that rhetoric about an alleged threat from the outside (i.e. not from the state itself) can serve as an effective measure for politicians running for office to gain popularity. Using data from a large number of political campaigns, the authors showed how an issue of, for instance, immigration can become a pivotal point in a politician’s campaign, drawing the public’s attention to the candidate. Thus, the political framework suggests that in case of an external threat, or a perception of such, citizens put more trust in their government as compared to the absence of such perception.

H1: The perception of threat increases citizen’s trust in their government.

(Conventional) Media and Public Opinion

It is known that media has the capacity to shape opinions of the public through the processes of agenda-setting, priming, and mediatization (Pan & Kosicki, 1993; Lawrence, 2000; Strömbäck & Esser, 2014). Considering that the majority of people rely on media for information about local and international affairs, an explanation of its effects is in order. There is a strong scholarly case built demonstrating that through the processes of mediatization and agenda-setting, mass media can have strong influence on the public opinion on a given issue. With its growing importance in politics, media has become the main source of communication between the public and the political arena (Strömbäck & Esser, 2014). The epiphenomenon of such a development, is that the public is subjected to media logic and agenda (Pan & Kosicki, 1993; Lawrence, 2000): people receive most of their information about the developments in the political sphere through media outlets, which unavoidably distort the entireness of available information in choosing what to present to the public.

The theory of agenda-setting postulates that the mass media have a significant impact on the public by inevitably selecting what they cover. The theory suggests that the news is not just a reflection of reality, but rather a constructed, edited, and shaped part of it. This happens due to the fact that before reaching the consumer, raw information is processed by journalists, editors, publishers and owners of mass media - that is all those participants in the information delivery process that stand between the event and the final consumer of news. Thus, the media have the ability to highlight certain events (problems, topics, phenomena) and focus on them, forcing the audience to perceive these events as extremely important. In so doing, the mass media form the agenda of the society.
The priming effect, similarly to agenda-setting, is a process in which the media draws attention of the public to certain issues and problems, thereby silencing others and changing the criteria according to which the public evaluates the current affairs. A number of studies have demonstrated the existence of unique effects of priming in media, which differ from those of agenda-setting (Scheufele & Tewksbury, 2007). The theory of political priming is based on the assumption that people do not have sufficient knowledge about the political processes taking place in society and do not take into account all the information available to them when making political decisions. Instead, people tend to consider only the information that lies on the surface of the information field. Unlike agenda setting, priming defines not only what issues the public is concerned with, but also which particular aspects of a given issue it deems relevant (Iyengar, Peters, and Kinder, 1982). In the context of this paper, the priming effect suggests that if news on Russia are mostly concerned with the Ukrainian/Crimean conflict, the public will associate new information on Russia with that particular issue, thus primarily deeming the country aggressive.

Another significant effect exerted on the public by the media is issue framing (Goffman, 1974). In its fundamental form, the media framing theory holds that framing of a certain issue (i.e. the manner in which a piece of information is presented) influences the way audience process the received information. This effect affects the audience differently from agenda-setting or priming by suggesting how to evaluate a certain issue, rather than merely telling what issue to think about (Scheufele & Tewksbury, 2007). This happens when certain aspects of an issue discussed in the news are accentuated (at the expense of others) through metaphors, spin, word choice, jargon, etc. (Fairhurst & Sarr, 1996), thereby making the audience unintentionally view the issue in particular light.

Considering these effects, the following hypothesis is postulated.

H2: The security issues voiced in media result in increased trust in their government by the general population.

Methodological Background

Extraction of Public Opinion from Social Media

One of the applications of sentiment analysis is determining the public’s political sentiment. Sentiment analysis has been applied successfully to extract political preferences and opinions from social media. Sentiment analysis of ‘tweets’, short publications on Twitter, accurately portrayed political landscape, reflecting real-world coalitions, political positions, and election results (Timasjan et al., 2010). It was also shown that sentiment analysis of ‘tweets’ can predict national election results in favor of traditional polls (Bermingham & Smeaton, 2011). Moreover, social media’s content can serve as an
indicator of inter-state relations, accurately reflecting a country’s ties with another state (Chambers et al., 2015).

Research has shown that correlation between opinions extracted from social media and political opinion surveys to be considerably high. O’Connor et al. (2010) used measures of public opinion from polls with sentiment measured from text. Analysis of several surveys on consumer confidence and political opinion showed that they correlate to sentiment word frequencies in contemporaneous Twitter messages. Across various datasets, while high overall, the correlation reached 80% in some instances, indicating that text streams can serve as a substitute to the traditional polling.

Tumasjan et al. (2010) developed a system for real-time analysis of public sentiment toward presidential candidates in the 2012 U.S. on Twitter, following an assumption that the web-site has become a hub for expression of people’s opinions and views on political parties and candidates. The research notices that emerging events or news are often instantly reflected by a burst in Twitter posts’ volume. This coupled with an ability to semantically analyze the said posts provides an opportunity to determine the relation between expressed public opinion and relevant electoral and political events. Not only that, sentiment analysis can also serve as a toll for exploring how these events affect public opinion. While traditional content analysis on such a topic takes days or weeks to complete, the system demonstrated by the researchers analyzed sentiment in the entire Twitter traffic about the 2012 U.S. election, delivering results instantly and continuously.

Thus, the research suggests that online opinion mining can be an effective tool in evaluating current and future real-world political landscape. In manner exemplified above, sentiment analysis is used in this paper to measure the dependent variable – citizens’ trust in government.

**Sentiment Analysis**

Sentiment analysis is a class of methods of content analysis designed for an automated detection of the emotive vocabulary in a text and its author’s emotional evaluation (or opinion) with respect to topics in this text (Pang & Lee, 2008). Sentiment analysis aims to determine the emotional attitude of the author towards an object (real-world objects, events, processes, or their properties / attributes) expressed in the text. The main objective of sentiment analysis is evaluation of polarity of a (written) object, that is, determining whether the opinion expressed in the document is either positive, negative or neutral. Modern methods of automated detection of sentiment in a text most often employ a binary emotional measure: positive or negative. However, there are successful cases of use of other methods (Pang & Lee, 2008; Liu, 2010).

The polarity of the document can be determined on a binary scale. In this case, to determine the polarity of a document, two classes of assessment are used: positive or negative. One of the downsides of this approach is that the emotional content of a document cannot always be clearly defined (i.e. the document may contain signs of both
positive and negative opinion) (Pang & Lee, 2008). A document’s polarity can also be determined on a multi-polarity scale (Pang & Lee, 2005; Snyder & Barzilay, 2007). In this case, the usual classification of “positive vs. negative” is extended to a 3- or 4-point scale. This method of evaluating polarity requires a more intricate dictionary than the binary method.

H3: Sentiment analysis can be used to determine the public opinion about the government.

Methodology

To answer the research question, articles from popular Lithuanian media outlets and user-generated comments from Facebook were scraped. Then, the sentiment expressed towards the Lithuanian government in them was measured and its change over time was analyzed. The sentiment from media articles and social media comments was checked for correlation and causation. Finally, the change in expressed sentiment towards Russia was measured and compared to the change in sentiment towards the Lithuanian government.

Acquiring the Data

The article and commentary text for the further analysis were extracted from two of the most popular online media sources in Lithuania: DELFI (www.delfi.lt) and 15 Minučių (www.15min.lt) including the Facebook page of www.delfi.lt (given the regularity of posts with the news articles and the abundance of user commentary on them). In order to scrape the necessary information (article/comment text and the date published), a code was built using R programming language. External packages were employed when building the code to expand the language’s functionality: RVest (Wickham, 2016) package which allows for extraction of relevant information from selected HTML and XLM pages; Stringr (Wickham, 2017) package that introduces expanded manipulation of character strings; the Rfacebook(Barbera, Piccirilli, Geisler, and van Atteveldt, 2017) package that simplifies the scraping process of Facebook posts and discussions; the AmCatR (van Atteveldt&Welbers, 2014) package that allows for the in-R interaction with the AmCat API used in the research to store and query the articles; the SemNet (Welbers & Atteveldt, 2016) package for creation and analysis of semantic networks; and the Causal Impact package (Brodersen et al., 2015) designed for simplified applications of Bayesian time series model.

The articles and comments were collected for the period of September 1, 2013 to August 31, 2014, which created a half-year period prior to and after the event of annexation of Crimea. During this period, the same government was in place (both the parliament
and the president), creating favorable conditions for the analysis by excluding a crucial confounding variable.

In total, 20837 articles and 81969 Facebook comments were downloaded. The data was then uploaded to Amsterdam Content Analysis Toolkit (AmCAT), a website that facilitates large-scale content analysis, allowing to filter out the articles and comments that are not on topic. Out of all the articles downloaded, only the ones related to the Lithuanian government, its representatives or parts were chosen. This was accomplished by using the following query:

```
prezident* OR ministr* OR pirminink* OR valdž* OR (seim* AND nar*) OR valstyb* OR Grybauskait* OR Kubili* OR ministerij*
```

This resulted in 8162 articles and 1448 comments being selected for the analysis.

**Operationalization: Building the Dictionary**

Due to unavailability of sentiment dictionaries in Lithuanian, a new one was constructed from scratch by (1) manual reading of the articles and location of the words used in relation to the government, and (2) by querying the articles and locating the words in proximity with the key words. The dictionary’s validity was assessed through a peer review: peers were asked to independently evaluate negativity of a sample of articles which was then compared to the results of an automated analysis. The aim of the peer review was not to establish whether the absolute values of negativity are the same between the manual coding and the automated one, but rather to see whether relation in negativity across the articles stays the same in both cases. If, for instance, two articles that were coded manually had three and one negative mentions of the government respectively, a successful automated analysis would produce two values with the ratio of 3:1 between them, regardless of what these absolute values are.

First, 200 on-topic news articles from Delfi and LietuvosRytas were analyzed to determine the sentiment-carrying words used in relation to the government, its parts and members. The selected words were divided into positive and negative categories and stored as separate files. Only the words that carry a sentiment (nearly) universally were chosen, resulting in 113 negative terms and 37 positive ones. To expand the dictionaries, a semantic network of all the queried articles was constructed to monitor the word co-occurrence and thus to determine what sentiment-carrying words had been mentioned in proximity with the query words. A range of 20 words was chosen, with 10 words before and after a query word, based on the average sentence length in Lithuanian media (Leonavičienė, 2010). Providing, that the average sentence length is 10 words, all the words in the query-word-containing sentence were captured regardless of the query word’s position in it. In total, 96798 words were captured; by eliminating the duplicates
and the words that co-occurred with a query word only once, the list was shortened to 12043 individual words. As before, the sentiment carrying words were selected and sorted by positive and negative meaning. The expanded dictionary contained 511 negative terms and 143 positive ones.

Once the dictionary was completed, sentiment in a random sample of 8 articles was determined (using the method described below) and compared to that of peer coders’ who also had to evaluate the sentiment of these articles. In 6 of these articles, the automated method detected higher negativity than the human coders; in the other 2 the negativity was moderately lower compared to human coders. Overall, the accuracy of the automated sentiment analysis was 72% compared to human coding.

**Operationalization: Determining the Sentiment**

Due to the scarcity of the materials available, a basic form of sentiment analysis was conducted. It was performed by means of keywords (query words) and distance-based measurement of the co-occurrence of words in texts (van Atteveldt, 2008). Thus, the analysis was a common application of collocation sentiment analysis, where the co-occurrence of words is regarded as an indicator of sentiment within a given piece of text. This measure was regarded as indicative of tone. The distance within which the words were considered to co-occur was determined by the average sentence length in the language of Lithuanian media (Leonavičienė, 2010).

Each keyword was measured against (1) the words in negative dictionary and (2) the words in the positive dictionary, occurring within the range of 10 words from the key word. The matches for each of the keywords were summed up per article resulting in the total negativity/positivity score per article. The results for individual articles were then aggregated by the day of the year, with the daily relative negativity value being calculated using the following formula:

\[ \sum (\text{article negativity}) \div (\text{article count}) = \text{relative negativity} \]

The daily sentiment was calculated in a similar way:

\[ \sum (\text{article positivity}) \div (\text{article count}) - \sum (\text{article negativity}) \div (\text{article count}) = \text{article sentiment} \]

Thus, all articles published on a given day were included in the calculation regardless of their sentiment score, with the days when sentiment could not be located receiving a value of 0.
Modelling: Evaluating Impact of Russia’s Actions

The data was fitted into Bayesian structural time-series model (Gelman, Carlin, Stern & Rubin, 2014) to determine whether Russia’s aggressive foreign policy had an impact on Lithuanian citizens’ trust in their government. The model is a “diffusion-regression state-space model that predicts the counterfactual response that would have occurred had no intervention taken place” (Brodersen et al., 2015): it creates predicted values for the post-intervention period based on the pre-intervention part of the time-series and compares them to the actual post-intervention values to assess the difference between the two, allowing for determination of the temporal impact of the intervention. The relative negativity and sentiment time-series for both media articles and Facebook comments were assessed using this model. For all instances of the model usage, the intervention period of 1 March – 22 March, 2014 was chosen based on Russia’s actions in Crimea outlined above plus the lag of one day to allow for reaction to the developments to become visible in (social) media.

To check the model’s credibility, a simulated intervention was constructed. The values in the post intervention period were made identical to the ones in the pre-intervention part of the time-series. Then, using the same intervention period (March 1 - March 22, 2014), the values were fitted into the Bayesian structural time-series model to determine the presence or absence of a causal impact and in doing so to assess the model’s validity for the data of this study.

To further determine the validity of the aforementioned model, an auxiliary ARIMA model for the pre-intervention values was constructed to forecast the post-intervention values for their comparison to the actual ones. Based in the ADF and KPSS test values, and the ACF and partial ACF plots, an ARIMA model with autoregression component of order 3 and the moving average component of order 2 was constructed (i.e. 3,0,2). The model was then used to determine the predicted values for the post-intervention period (after March 22, 2014). Using a simple t-test and VAR analysis (with the lag of 1 calculated using the AIC, SC, HQ, and FPE parameters), the predicted values were compared to the observed ones.

Modelling: Relation between Media Articles and Facebook Comments

A VAR analysis was performed to establish whether there is an immediate relationship between the sentiment expressed in comments and articles that could (1) serve as a demonstration of the fact that media reflects opinions of the public or (2) indicate a kind of priming effect media exerts on the public opinion (Scheufele & Tewksbury, 2007).
Before conducting the analysis, both time series (from media articles and comments) were checked for stationarity with ADF and KPSS tests. In both cases the data appeared to be stationary with the KPSS test producing a $p$-value greater than 0.1 and the ADF test producing a $p$-value below 0.05. The ACF and partial ACF plots also suggested the data is stationary (Fig. 1). For the analysis, a lag of 1 was chosen based on all (AIC, HQ, SC, and FPE) relevant parameters.

Fig. 1. ACF plots (top) for articles (left) and Facebook comments (right) and corresponding partial ACF plots (bottom)

Modelling: Relation between Sentiment towards Russia and the Lithuanian Government

A VAR analysis was conducted to check for presence of priming effect the negative exposure of Russia could have on the Lithuanian public on the meso scale. To conduct this analysis on the macro scale, a Bayesian model was used.

First, the VAR analysis for relation between sentiment expressed in media articles towards Russia and sentiment towards the government of Lithuania in Facebook comments was conducted. The lag chosen was 6 based on AIC and FPE parameters which are considered to be more predictive than the HQ and SC parameters (Venus Khim-Sen, 2004). Before running the analysis, the data was checked for stationarity. Based on KPSS test values ($p<0.05$), the ‘Russia’s’ sentiment time series were not stationary. Hence, the value differences were used instead. After applying this treatment, both KPSS and ADF tests suggested stationarity, with ACF and pACF plots showing the same outcome (Fig. 2).
Then, the VAR analysis for relation between sentiment expressed in media articles towards Russia and towards the government of Lithuania was performed. Again, the lag of 6 was chosen, based on all but HQ parameters.

Finally, the sentiment towards Russia was fitted into the Bayesian time-series model. The intervention period, following the previous models, was March 1 – March 22, 2014.

**Results**

Overall, the results show that: (1) the negativity towards Russia in Lithuanian media has increased dramatically after end of February, 2014 – following its actions in Crimea; (2) at the same time, the sentiment towards the Lithuanian government in both media articles and Facebook comments became more positive; (3) no correlation on meso scale between the sentiment expressed towards Russia and towards the Lithuanian government was observed; (4) no correlation was observed between media articles and Facebook comments on micro and meso scales; and (5) the model used to estimate the effects of Russia’s actions was effective.

**Change in Negativity towards Russia**

The results show that while the daily count of articles in Lithuanian media concerning Russia has remained similar throughout the examined period, the negative attitude towards it has increased dramatically following its actions in Crimea. The graph below visualizes the relationship between article count and article negativity towards the country. At the end of February 2014, as Russia began overt military operations in Crimea, there was a rapid and dramatic increase in negativity towards in the Lithuania media. By 3 March the increase had peaked – right after the announcement of a new pro-Russian minister of Crimea. The negativity remained as high until March 17 (when reportedly, a
Russian spy attempted to recruit a Lithuanian official) after which it started to decrease slowly, remaining at levels higher than the pre-Crimean period.

At the beginning of August (Fig. 3) there is another sharp increase in negativity, that could be explained by Russia’s ban on agricultural, raw, and food products from Lithuania. Moreover, the mentions of Russia in the context of the Ukrainian conflict closely match the negativity expressed towards it.

**Impact of Russia’s Foreign Policy**

To test whether Russian Federation’s aggression against Crimea had an impact on the sentiment in media articles and social media comments, a Bayesian structural time-series model was constructed with the intervention period of 1 March – 22 March, 2014 specified. First, the by-day time series of average sentiment were fitted into the model (Fig. 4).

![Mentions and Negativity of Russia in News](image1)

![Negativity and 'Crimea' mentions of Russia in News](image2)

*Fig. 3. Mentions of Russia in media articles and the negativity towards it over time (left); and negativity and mentions of Russia and Crimea over time (right)*

![Sentiment growth after the intervention date with the 3rd graph showing the estimated effect of the intervention](image3)

*Fig. 4. Sentiment growth after the intervention date with the 3rd graph showing the estimated effect of the intervention*
The effect of the intervention produces a positive 20% difference as compared to values predicted from a non-intervention model with probability of it occurring by chance $p = 0.037$ (probability of causal effect being 96.29%). In absolute values, it translates into 0.39 difference a day on average after the intervention (predicted $-2.0$, compared to actual $-1.6$), with the 95% confidence interval (CI) being $[-2.4, -1.5]$. Similar effect could be observed when the intervention was modeled using the relative negativity only (as opposed to the sentiment that also takes into account the positivity of an article). The result is presented in the figure below (Fig. 5).

![Image](image.png)

*Fig. 5. Negativity after the intervention date with the 3rd graph showing the estimated effect of the intervention*

When calculating the effect employing the relative negativity, the actual values deviate from the ones predicted by the non-intervention model by 13% ($p = 0.048$). While the average estimated daily value of the non-intervention model was 4.9, the actual model produced an average value of 4.3 in the post-intervention period. The 95% confidence interval was $[4.2, 5.7]$. It could be concluded from these results that, as hypothesized, the media started viewing the Lithuanian government more positively as the negativity towards Russia increased.

This could also be observed when analyzing the sentiment extracted from the Facebook article comments. The Figure 6 below presents a visualization of the change in the sentiment / negativity in the post-intervention period.
The probability of the intervention (Russia's aggression) having an effect on expressed sentiment as derived from comments was significantly lower when compared to the change in sentiment derived from media articles. The p-value of the effect taking place calculated based on sentiment was 0.2 (with the effect producing a relative change of 49%), and the p-value estimated based on relative negativity was 0.17 (with the intervention creating a relative effect of 19%). The predicted (non-intervention) average daily value for the sentiment-based estimation was 0.26 against the actual value of 0.39. Comparably, the predicted daily value for the calculations based on relative negativity was 1.5 against 1.4. The 95% CI’s were [–0.031, 0.55] and [0.79, 2.3] for the sentiment- and relative negativity-based calculations respectively. Thus, the results do point towards the increase in positive sentiment towards the Lithuanian state also observed in media articles, but they are not quite statistically significant.

Correlation Between Sentiment Towards Russia and the Government of Lithuania

A formal analysis of the relationship between the sentiment expressed towards Russia and the Lithuanian government was conducted to assess the possibility of priming and/or effects taking place. The VAR analysis of the sentiment in relation to Russia predicting the values of the sentiment expressed towards the Lithuanian government produced the p-value of 0.46 at the lag 6 (chosen based on AIC, SC, and FPE parameters; the HQ parameter was optimal at the lag of 4, but was not considered based on the weight of other parameters). Similarly, other lag values (1–10) also produced no significant result. Below, in Figure 7, the visualization of this VAR is presented.
The absence of statistically significant results in VAR analysis of the sentiment expressed towards the government of Lithuania and towards Russia could indicate the absence of priming effect between the two variables. This suggests that the negative sentiment towards Russia in news articles and social media comments did not cause a predisposition among journalists and the general population of Lithuania to view their government in a more positive light when considering priming as a continuous phenomenon, a regular but short-term effect.

The Bayesian time-series model to assess the long-term effect of the Russia’s aggression in Crimea on the sentiment expressed towards it in Lithuanian media and Facebook comments showed a relative increase in negativity of 145% after March 22 (with the intervention period being March 1 – March 22, as previously). The reported $p$ value was $0.001$ with the confidence interval of $[114\%, 175\%]$. A visualization of the effect is shown above (Fig. 8). On the larger scale, one can observe a sharp rise in negativity towards Russia in news media between the end of February, 2014 and end of March 2014. Im-
Immediatly after that, there is another spike in negativity in the middle of April. After the spikes even out, the negativity remains significantly higher, rising by 141% on average compared to it before. This change also coincides with developments in Lithuania-Russia relations. Such a sharp rise points towards a more long-term priming effect where the rise in negativity towards Russia and its subsequent heightened levels cause the fall in negativity towards the Lithuanian government.

**Correlation between Sentiment Expressed in Articles and Comments**

Next, it was examined whether the sentiment expressed in media articles and Facebook comments followed the same pattern and could serve as each other’s predictor. A VAR analysis (Fig. 9) produced a correlation of 0.08 with the lag of 1 chosen based on all (AIC, HQ, SC, and FPE) relevant parameters. The probability of article sentiment predicting the sentiment in comments was $p = 0.84$, whereas the opposite relation produced the $p$ value of 0.44. So, no evident correlation or predictive ability was observed, suggesting that daily sentiment values in media articles cannot predict the daily sentiment in Facebook comments.

![Diagram of fit and residuals for articles](image)

*Fig. 9. VAR analysis for predictive ability between articles and comments*

**Bayesian Model Validity**

To assess the Bayesian structural time-series model’s validity when applied to this study’s data, two additional tests were conducted. They suggested that the results produced by the model are credible. First, A model with the post-intervention values swapped for the pre-intervention values (so that the pre- and post-intervention intervals are identical) produced a $p$ value of 0.5 with the intervention causing a relative effect of 0.05%, showing, as expected, that the supposed intervention caused no significant effect (Fig. 10).
Then, an ARIMA model based on the pre-intervention values with the post-intervention values predicted from it suggested the same outcome. A simple t-test comparing the predicted values to the actual ones in the post-intervention period showed that the means of the two groups are different ($p < 0.05$). A VAR analysis, too, produced a $p$ value of $>0.05$ (Fig. 11).

![Diagram of fit and residuals for predicted](image)

*Fig. 10. Synthetic intervention model for validity test*

*Fig. 11. VAR output for predicted vs. actual values*
Discussion

The analysis of media articles demonstrates that there is a cotemporal change in the sentiment expressed towards Russia and Lithuania: as the former decreases the latter becomes more positive. Assuming that the media reflects the attitudes of the general public, such a relationship means that the heightened sense of insecurity from the outside results in a more positive view of the state.

The mechanism of how this happens could be explained by the media effects on the audience. The fact that the number of mentions of Russia in media articles over time changed only slightly indicates that little to none of the first-level agenda setting effect was taking place, as its intensity is highly dependent on the frequency of mentions of the issue in question. However, the significant rise in both the negativity towards Russia and its often mentions in the context of Crimea suggest that framing and priming effects could be taking place. The fact that after its invasion of Crimea Russia was mostly mentioned in relation to the Crimean conflict in the Lithuanian media, shows that the Crimean conflict could have become the salient issue when it comes to Russia for the public of Lithuania. A survey conducted by 15 Minučių (one of the media portals used in this study) on 17 December, 2014 also hints towards that conclusion: 61% of the respondents said that the most memorable event of 2014 was the war in Ukraine and the occupation of Crimea. Furthermore, after the events in Crimea, Russia has been portrayed with much more negativity than before (141% on average) hinting towards the probability of a framing effect taking place, since the audience is susceptible to adopting the attitude with which an issue is presented in the media.

However, no direct relationship between the sentiment expressed in articles and the definite public opinion (as expressed in social media comments) could not be established: the trends observed in both data sets suggested the same conclusion of increasing sentiment towards the governmental instituitons of Lithuania at the time of the annexation of Crimea, but the analysis of Facebook comments produced no statistically significant results.

Overall, what is certain is that the sentiment in media articles can serve as a reliable reflection of the developments in international relations between Lithuania and Russia: the sentiment expressed in articles coincides with the contemporaneous events on the international arena, spiking in negativity immediately after Russia’s actions detrimental to Lithuania. The analysis of sentiment towards the Lithuanian government too produced definite results showing that the timeline of the rise in sentiment matches the rise in negativity towards Russia. If we were to consider the effects of priming and framing exerted by the media and rely on the inconclusive results acquired from the Facebook comments, we could reason that the trust in the Lithuanian government has increased following Russia’s annexation of Crimea.
Relevance

This paper contributes to the field of political communication by combining and comparing the sentiment extracted from both conventional media and social media for attainment of more rigorous results, something that has been hardly done before. It demonstrates that the data acquired from the both mediums is complementary, revealing and suggesting more than these mediums would in isolation. Furthermore, the article adds to the field of political theory by hinting towards the relationship between an outside threat, international conflicts, and confidence in the state institutions, thereby supporting the validity of the “rally 'round the flag” effect. Finally, it finds support, though indirect, for priming and framing effects which are essential in communication theory.

Limitations and further research

Their lower p-values could be explained by several factors. Out of 81969 comments that were scraped from the Facebook webpage, only 1448 could be identified as relevant to the topic based on the aforementioned query. With sentiment measured through co-occurrence of keywords, the sample further decreased to 615 individual comments. Additionally, in 70.7% of these comments only 1 or 2 counts of expressed sentiment were detected.

When measuring relative negativity and sentiment, the number of all comments published on a given date was taken into account. This could have resulted in under-reported and overly flat measures of expressed sentiment. More importantly, the sample of the acquired comments was too small (especially compared to the sample of media articles) to produce any definite results.

The small size of the on-topic comments sample and that of the sample with detected sentiment could be explained by two factors. First, the stylistic structure of comments could have played a role. The subject of a comment is often implied rather than overtly stated, as it is usually written as a short reaction to a certain topic. This poses a difficulty in automatically identifying the comments that are on a certain topic. Secondly, the dictionary used in this study could have been suboptimal: it was created based on an analysis of media articles which can have a vastly different vocabulary from journalistic style of writing. No slang words and only some (commonly accepted) abbreviations were included in this study’s dictionary, whereas social media comments tend to be abundant in both. Consequently, it is likely that some of the sentiment-carrying words in the scraped comments were missed.

Another limitation encountered in this paper is its treatment of the governmental trust only from the perspective of security. Undoubtedly, the perceived legitimacy of the state is dependent on other factors such as its economy, corruption and transparency levels,
societal involvement in the institutions, legislative effectiveness, etc. These variables were not controlled for, whereas they could serve as potential explanations for the hypotheses postulated in the paper. Thus, further research with rigorous control for these factors is necessary to improve the credibility of this study’s findings.

To the author’s knowledge, this paper was the first example of sentiments analysis in the Lithuanian language. Consequently, there is a lot to be done to bring the analysis tools to the state of the art level. A tool should be developed that goes beyond identifying the polarity of a piece of text via: (1) improving the dictionary not only by adding more words and phrases to it, but also by creating a scale, assigning affinity and using grammar rules, and (2) implementation of machine learning algorithms (latent semantic analysis, support vector machines), etc. that can provide more contextual results.

References


Žiniasklaidos požiūrio analizė pasitikėjimo vyriausybe
patikimumui matuoti

Vladislav Petkevič

Amsterdamo universiteto kolegija, Science Park 113, 1098 XG Amsterdam, petkevich.vladislav@gmail.com

Santrauka


Esminiai žodžiai: nuotaikos analizė, žiniasklaidos efektais, Lietuvos ir Rusijos santykiai, pasitikėjimas vyriausybe.

Gauta 2018 04 09 / Received 09 04 2018
Priimta 2018 04 19 / Accepted 19 08 2018