

A machine learning framework for automated analysis of central bank communication and media discourse.

The case of Narodowy Bank Polski

Krzysztof Rybinski*

Submitted: 2 August 2018. Accepted: 3 December 2018.

Abstract

This paper presents a supervised machine learning framework based on the dictionary and Wordscores models that allows to analyse interactions between the official central bank communication (policy statements) and media discourse (newspaper articles). It was tested on the case of Narodowy Bank Polski in the period 1998–2018, with 70 policy statements accompanying interest rate changes and 21,181 *Rzeczpospolita* daily articles mentioning the central bank. Using a new concept – correlation of policy lexical sentiment – we found that NBP can successfully affect media discourse when the press discusses topics that fall into the NBP mandate. We also documented that the focus of both RPP statements and press articles changed over time. In the validation procedure we showed that the biggest challenge in applying machine learning to detecting monetary policy inclination in press articles is the existence of many competing policy transmission channels.

Keywords: monetary policy, central bank communication, text mining, Wordscores, Narodowy Bank Polski

JEL: C8, E5, G14

1 Introduction: central bank communication and machine learning

Until the 1990s, central banks were very secretive. Before 1994 the Federal Reserve did not even announce its interest rate decision: markets had to guess it on the basis of conducted open market operations. However, in the 1980s and 1990s the academic community started to investigate the issue of central bank transparency and its impact on the effectiveness of the monetary policy, which led to the publication of many papers and influenced the way central banks conducted monetary policy and communicated with the markets.¹ From the 1990s central banks became more open and started to use communication as one of the monetary policy tools.

Until recently, the majority of central bank communication studies were conducted according to the following simplified pattern. Central bank press releases, official documents, or central bank officials' quotes in the media were collected and manually coded by researchers.² Then an econometric model was used to quantify the relationship between the coded characteristics of central bank communication and financial markets or economic variables, such as short-term interest rates. Such a human-intensive approach has its obvious limits. It would be next to impossible or very expensive (even using the graduate student hourly wage) to read and manually code tens of thousands, or millions of texts. However, the emergence of machine learning algorithms has allowed monetary policy researchers to analyse larger text corpora or use machine learning algorithms to extract new relationships between central bank communication and economic and financial variables. Such approaches can be classified into two broad groups. The first group belongs to the so-called unsupervised learning class of models, where we do not know a priori whether a given analysed text is associated with monetary easing or monetary tightening. In the majority of cases the authors of these papers used latent semantic analysis and latent Dirichlet allocation models to extract the topics of the central bank communication.³ These methods provided useful insights about monetary policy so that even the analysts of commercial banks started to use them to predict future interest rate changes; see for example Wan (2017) and Ortiz, Rodrigo and Turina (2017).

The second group of models belongs to the supervised learning category, when we know the outcome or the feature of interest of the analysed objects. Such data is called in machine learning the "labelled data". In the case of central bank communication, the "label" is the stance or inclination of the monetary policy, i.e. we know whether a given text is associated with an increase, no change, or decrease of the interest rate or another central bank instrument, depending on the monetary policy regime in the given country. Because nowadays the central bank policy instrument decisions are always accompanied by a release of the central bank statement, we can collect a corpus of such statements and assign to each statement the proper label corresponding to the size or direction of the policy instrument change. For example, Jansen and de Haan (2010) used a supervised learning Wordscores

¹ See for example Canzoneri (1985), Faust and Svensson (1998), Eijffinger, Hoeberichts and Schaling (2000), Guthrie, Wright (2000), Jensen (2002), Geraats (2002), Hahn (2002), Carpenter (2004), Connolly and Kohler (2004), Ehrmann and Fratzscher (2005a), Kohn and Sack (2003), Gürkaynak, Sack and Swanson (2005), Gerlach (2005), Heinemann and Ulrich (2005), Ehrmann and Fratzscher (2005b), Woodford (2005), Blinder et al. (2008) and Siklos, Strum (2013). The first research on central bank communication in CEE was done by Rozkrut et al. (2007).

² For example in Rozkrut et al. (2007) Czech, Hungarian and Polish central bank policy statements (169 texts) and media quotes (367 texts) were manually coded in three categories: inclination to tighten/ease monetary policy, improving/deteriorating economic outlook and overvalued/undervalued exchange rate.

³ See for example Boukus and Rosenberg (2006), Hendry (2012), Moniz and de Jong (2014), Masawi, Bhattacharya and Boulter (2014), Correa et al. (2017), Luangaram and Wongwachara (2017), Lündering and Tillman (2016), Hansen, McMahon and Prat (2018).

model to analyse the consistency of ECB communication between 1999 and 2009. They trained the model on ECB introductory statements from years 1999–2001 and made predictions about the policy stance by applying the trained model to texts from 2002–2009. They found a high correlation between the estimated text scores and actual policy decisions and concluded that ECB communication was consistent in the analysed period. Bennani (2015) applied Wordscores models to 2-grams to estimate the forward guidance of major central banks. Both papers used policy decisions to label the data, which means that the text of the policy statement received a label equal to the size of the rate move, or to +1/-1 for a tightening/easing decision. In the case of Poland, there has been only one paper analysing central bank communication with machine learning tools. Doryń (2018) applied sentiment analysis to Poland's Monetary Policy Council minutes and found relatively high correlations of such sentiment and various business and consumer confidence surveys in Germany and in Poland, and to a lesser extent with other economic variables.

In the cited unsupervised learning literature one can find examples of applications to central bank policy statements, press articles or social media, such as Twitter posts. However to date there has been no attempt to apply supervised learning models to analyse interactions between central bank communication and media discourse. Media discourse can yield important insights about the future direction of the monetary policy, can verify whether and how central bank official communication affects (leads) the policy expectations held by the general public, or on the contrary, whether central bank responds to (lags) monetary policy discourse in the media. This paper fills in the gap in the literature and for the first time applies supervised machine learning models to central bank communication to analyse the relationship between the official policy statements and the media discourse. The machine learning framework developed in this paper can be used to automatically analyse large text corpora without human involvement. We test this new methodology on the case of Narodowy Bank Polski⁴ communication, but it is easily scalable and can be used to analyse any central bank.

2 Research question and data description

The goal of this paper is to analyse to what extent official RPP statements affect the media perception and discourse after the policy decision and whether RPP statements are influenced by media discussion before the policy meeting. It is a common practice that central bank officials give interviews, are quoted in the press or meet with analysts and journalists who later publicly present their opinions about the economy, inflation or monetary policy. The broad perception of the monetary policy or economic situation is formed by both: official central bank statements and media discourse that consists of central bankers' quotes or interviews and general commentary making reference to the central bank and its policy. This research will allow to understand how both central bank communication and media discourse are related. For example, we analyse whether newspapers articles mentioning the central bank after a hawkish policy statement (with high text score calculated by the Wordscores algorithm) have high or low text scores. If the majority of newspaper articles mentioning the central bank after a hawkish policy statement have low text scores, it would indicate that the general public (analysts

⁴ In Poland monetary policy decisions are taken by the Monetary Policy Council (in Polish: Rada Polityki Pieniężnej – RPP), its members are: the President of NBP and nine external members. By law NBP Vice Presidents also participate in the RPP meetings, but they do not vote on the policy decisions. RPP usually meets once per month. The author of this paper participated in the RPP meetings in years 2004–2008 as an NBP Vice President.

and journalists) has a different opinion about the economy, inflation or the appropriateness of the policy stance than the RPP. In other words it would indicate that the central bank has little influence on the general public opinion and that its communication policy is ineffective.

Our research is designed as follows. Each RPP statement released after the policy meeting receives a label derived from the interest rate decision taken during that meeting. There are three possibilities, interest rates were hiked, cut or kept unchanged. Of course, depending on the type of monetary regime and range of instruments used, there are also other possibilities regarding exchange rate, monetary targets or policy bias etc. In this paper we will use only rate decisions for labelling. In the analysed period between January 1998 and February 2018 there were 70 decisions that changed the main interest rate: 20 decisions increased the interest rate and 50 lowered it. Each policy decision was accompanied by a policy statement that we automatically scraped from the NBP website. We used Relenium package to automatically collect articles in the major Polish daily *Rzeczpospolita*⁵ published in this period that mentioned the central bank or the RPP.⁶

After filtering out short newspaper texts informing about current NBP exchange rates or monthly money supply data, we retained 21,181 *Rzeczpospolita* articles.

In the first step we conducted a standard text cleaning applied in the “bag of words” models.⁷ We removed stop words⁸, punctuation, capital letters, white spaces and digits. We also used large Polish language lexicon containing more than 3.3 million word-lemma pairs to replace words with lemmas.⁹ For a human, words (in Polish) “bank”, “banki”, “banków”, “bankiem” refer to the bank as the institution, however for a computer they are different words. The lemmatization process replaces words with their lemmas, so all the above words would be replaced with the word “bank”, which makes the texts harder to read for a human, but they become much more readable for a machine learning algorithm. However despite the large size of the used lemmatizer, not all words can be found and replaced by their lemmas. In such cases, the words are retained in their original form.

After cleaning and lemmatization, the text corpus was transformed into a document-term matrix (DTM). Rows in DTM represent texts, while each column corresponds to a token (word or lemma) that appears in the text corpus. Each DTM cell contains a number equal to the number of times a given token (column) appears in a given text (row). For the final analysis we removed tokens (columns) that appeared less than three times in the text corpus, as such rare tokens are not helpful in text mining and they simply introduce noise into the analysis.

⁵ We scraped all pages of each issue available in the *Rzeczpospolita* archive.

⁶ We use regular expressions to identify all texts that mention RPP and NBP, both abbreviations and full names, in all possible declination forms, for regular expressions were able to find (in Polish) “NBP”, “Narodowy Bank Polski”, “Narodowym Bankiem Polskim” etc., “RPP”, “Rada Polityki Pieniężnej”, “Radzie Polityki Pieniężnej” etc.

⁷ Bag of words models analyse texts as a set of words, ignoring their position in the text.

⁸ We used the list of stopwords in the Polish language available in the lsa library in R (277 words). The library is described here: <https://cran.r-project.org/web/packages/lsa/lsa.pdf>.

⁹ Many text mining papers use stemming to prepare texts for automatic analysis. Because for the Polish language we have a freely available large lemmatizer, instead of stemming we applied lemmatization. The lemmatizer at the time of writing this paper was available at www.lexiconista.com/datasets/lemmatization/. This lemmatizer has been used in the following papers published in the refereed journals Rybinski (2017) and Rybinski (2018). The author can provide the lemmatizer to interested researchers.

3 Methodology: Wordscores model and policy lexical sentiment correlation

Finally, the analysed DTM had 21,251 documents¹⁰ and 39,441 tokens and this matrix was used for the Wordscores algorithm.¹¹ 70 RPP policy statements included in DTM were coded as +1/-1 (+1 when the interest rate was increased, and -1 when the interest rate was reduced) and were used as reference texts for the Wordscores algorithm described below. Each of the 70 policy statements is treated as a “bag of words”, so initially each word (lemma) in the given statement receives the word score of +1 or -1, depending on the score of a given policy statement. So in the initial phase of the Wordscores algorithm, words inherit the policy score from the policy statement they appeared in. But the same word (lemma) can appear in many policy statements, and some of them can accompany interest rate hikes and some interest rate cuts. So the next stage of the Wordscores algorithm works as follows. If a given word appears in both hawkish and dovish policy statements, its score is calculated using relative frequency of appearances. For example if a word appeared once in a hawkish policy statement and three times in a dovish statement its score would be $(1*(+1) + 3*(-1)) / 4 = -0.5$. Such an exercise is conducted for all words (lemmas) that appeared at least once in any of the 70 policy statements and we obtain a set of words with their scores. Then this set of scored words forms a dictionary that is applied to all *Rzeczpospolita* articles in our corpus. Each article score is calculated as a weighted average of scores of words (lemmas) that appear in the text and can be found in the dictionary. This way we were able to calculate labels for newspaper articles. If a given *Rzeczpospolita* text contains many words that were typical for policy statements accompanying interest rate hikes, its label will be higher and will reflect monetary tightening. If the text contains many words typical for policy easing statements, its label will be lower and will reflect monetary easing. There are many words in newspaper texts that did not appear in any of the policy statements so they do not have any score. Such words are not used in the text score calculations. Finally, as a result of the application of the Wordscores algorithm we received a set of 21,181 *Rzeczpospolita* articles with labels representing the monetary policy stance.

The Wordscores model has several drawbacks typical for all “bag of words” models. For example it cannot understand negation, irony, sarcasm or multiple word meanings. For example the phrase (in Polish) “na bank” (eng. for sure) will be interpreted as an occurrence of a word indicating bank as an institution. Despite the described deficiencies, dictionary based methods are often used in the literature and produce insightful results.

In this paper we introduce a new concept that will be used to analyse the central bank communication and media discourse. When text scores of policy statements and newspaper articles preceding or following these statements show similar word patterns (have similar text scores) we refer to this situation as “positive lexical sentiment correlation”. In cases when text scores of policy statements and newspaper texts are very different, we would refer to this situation as a “negative lexical sentiment correlation”. We use the term correlation as our measure involves calculating the correlation coefficient between the policy statement text score and the average text score of newspaper articles preceding/following the statement in a given time window, across all RPP rate changing meetings. Before implementing the policy lexical sentiment correlation analysis we identify the key topics of newspaper articles by applying topic dictionaries, as described in the next section.

¹⁰ 21,181 articles and 70 RPP statements.

¹¹ The Wordscores model was introduced in Laver, Benoit and Garry (2003). A detailed description of the Wordscores algorithm and other text mining techniques can be found in Grimmer and Stewart (2013). All calculations were conducted in R programming language, using NLP libraries tm and quantda.

4 Understanding the relationship between the official central bank communication and media discourse

Inspection of a random selection of *Rzeczpospolita* articles revealed that many articles mentioning the central bank did not discuss topics relevant from the monetary policy perspective. They referred to the central bank in a different context and they constitute noise from our analysis perspective. In order to improve the signal-to-noise ratio, we defined ten topics that are usually communicated by the central banks: general economy, exchange rate, financial markets, financial stability, inflation, interest rates, the labour market, monetary policy, money and public finances. Each topic is characterized by several key words typical for a given topic.¹² The English translation of key words (lemmas)¹³ representing these topics is presented in Table 1.

Table 1
Topic dictionaries

Topic	Key words
Economy	GDP, income, production, economic, economy, consumption, investment, exports, imports, trade, deficit, surplus, balance
Exchange rate	exchange rate, currency, devaluation, depreciation, appreciation, zloty, euro, dollar, franc, exchange, ERM2
Financial markets	share, bond, instrument, derivative, stock exchange, volume, turnover, market, currency, euro, dollar, commodity, oil
Financial stability	stability, stable, capital, reserves, sector, prudential, supervision, requirement, non-performing, financial, measure (plus others, Polish language specific)
Inflation	inflation, deflation, price(s), purchasing power, expensive, cheap, falling (rising) price, sales, inflationary, deflationary, pressure
Interest rate	rate, percent (age), interest, reference, base
Labour market	work, labour, employment, employ, unemployment, unemployed, shed, wage, pay, salary, unemployment benefit, employer, employee, worker
Monetary policy	monetary, goal (target), strategy, statement, projection, guidelines, policy, report, conference
Money	money, supply, deposit, loan, account, liability, asset
Public finances	budget, debt, tax(es), income, expenditure, revenue, deficit, surplus, finance, public

Of course, the selection of key words is subjective and in dictionary-based models there are two main methods to verify whether the selection was proper: (1) a robustness analysis can be conducted to check whether altering the list of keywords changes the results and (2) validate the results by

¹² Sometimes one keyword can appear in more than one topic, for example “deficit” belongs to two topics: general economy and public finances, because it can refer to trade deficit and to budget deficit. A list of keywords for each topic was formed on the basis of the author’s experience of working in financial markets in various roles for more than 20 years of, which involved close contacts with the media.

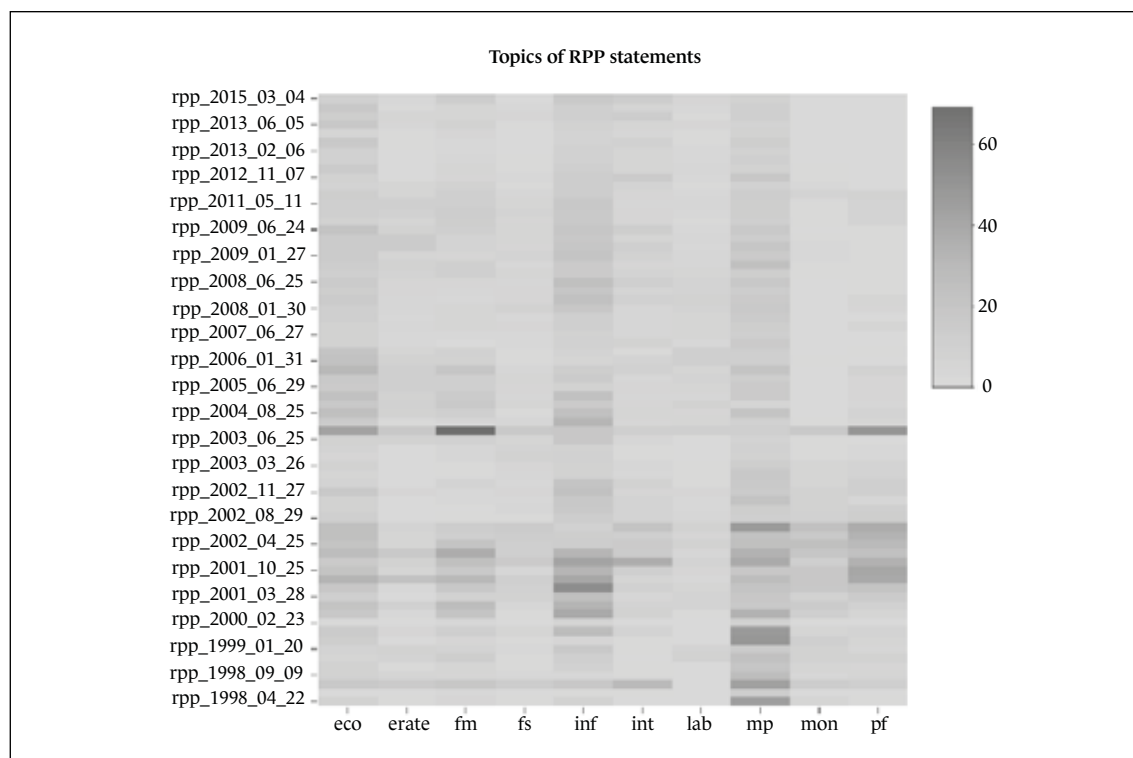
¹³ Keywords are in lemma form, because the words in the text corpus were lemmatized. Sometimes the English language translation of one Polish word takes two words.

comparing computer and human coding of the same texts. We adopted the validation procedure and showed in the latter part of the paper that these simple dictionaries do a surprisingly good job in properly describing the main topics of the newspaper articles. We applied these dictionaries to the text corpus to identify the key themes of both RPP statements and newspaper articles.

Figure 1 below presents all 70 RPP statements accompanying rate changes and the number of words (lemmas) detected in each statement that belongs to each of the ten topics.

Figure 1

Application of topic dictionaries to RPP statements



Notes:

Grey colour intensity on the heatmap represents how many words (lemmas) belonging to a given topic were found in the RPP statement.

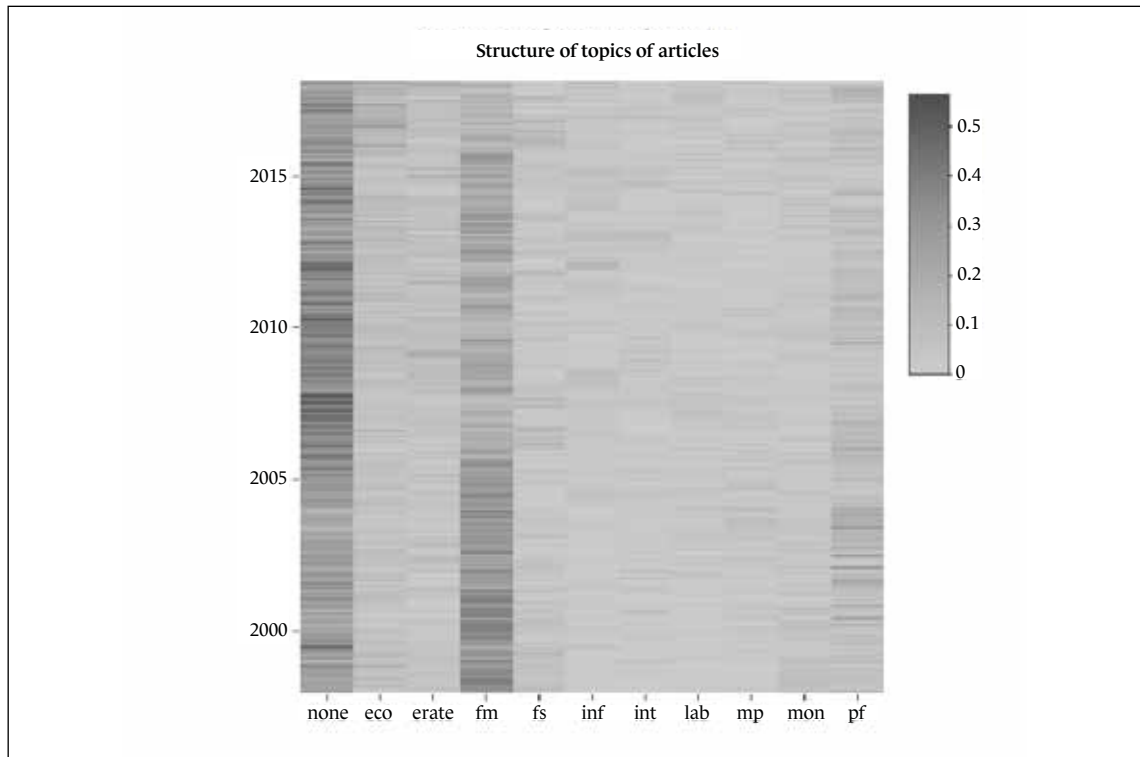
Abbreviations' meaning: eco – economics, erate – exchange rate, fm – financial markets, fs – financial stability, inf – inflation, int – interest rates, lab – labour market, mp – monetary policy, mon – money, pf – public finance.

With the exception of one RPP statement in June 2004, which was exceptionally long, we found that two topics are most frequent in RPP communication: monetary policy and inflation, which is natural for the inflation targeting central bank. General economy is the third most frequent topic. However, there was a period in 2001–2002 when the topic of public finances was featured very prominently in RPP statements, after Leszek Balcerowicz was nominated to the post of President of NBP in January 2001. Recent RPP statements accompanying rate change decisions after 2011 became much more balanced in terms of discussed topics, which coincides with Marek Belka assuming the post of President of NBP and with the change of the composition of the RPP that took place in 2010.

Then we looked at the key topics discussed by press articles. We chose the key topic for a given text if this topic dictionary score is higher by at least two than the scores of all the other topics. Then we calculated what share of articles published in a given month belong to each topic category. The results are presented in Figure 2 below.

Figure 2

Application of topic dictionaries to *Rzeczpospolita* articles. Structure of articles featuring given topic for each month



Notes:

Grey colour intensity on the heatmap represents what percentage of articles in a given month featured a given topic, determined as a key topic.

Abbreviations' meaning: none – dictionary method was not able to determine the key topic, eco – economics, erate – exchange rate, fm – financial markets, fs – financial stability, inf – inflation, int – interest rates, lab – labour market, mp – monetary policy, mon – money, pf – public finance.

In most months we were not able to determine the key topic for between 30 and 50 per cent of the articles in the month. Financial markets is the most frequent topic across the data set; however, it is less dominant in the last few years. The second most important topic is public finances; however, in recent years general economy and financial stability have been discussed more often.

Then we combined both data sets and looked for patterns in the topics of RPP statements and *Rzeczpospolita* articles. For press articles we applied a 14-day time window before and after the RPP meeting that resulted in the rate change (see Table 2 below).

Table 2

Number of texts discussing specific topics

Topic	Articles				RPP statements			
	total per topic	average no. before RPP meeting when rate hike	average no. before RPP meeting when rate cut	average no. after RPP meeting when rate hike	average no. after RPP meeting when rate cut	total per topic	when rate cut	when rate hike
Economy	1 592	3.4	3.8	3.6	3.4	9	9	–
Exchange rate	1 372	2.7	3.9	3.3	3.6	–	–	–
Financial markets	5 103	12.2	13.6	12.0	12.4	2	1	1
Financial stability	1 002	3.1	2.2	2.7	3.1	–	–	–
Inflation	739	3.4	1.9	2.1	2.8	24	14	10
Interest rate	652	2.3	2.3	2.1	2.6	–	–	–
Labour market	632	2.1	2.1	1.8	2.3	–	–	–
Monetary policy	644	2.3	2.1	1.7	2.1	19	12	7
Money	831	2.1	2.5	2.4	2.6	–	–	–
Public finances	1 861	3.3	4.5	4.3	3.9	3	3	–
Not determined	6 753	–	–	–	–	13	11	2
Total	21 181	–	–	–	–	70	50	20

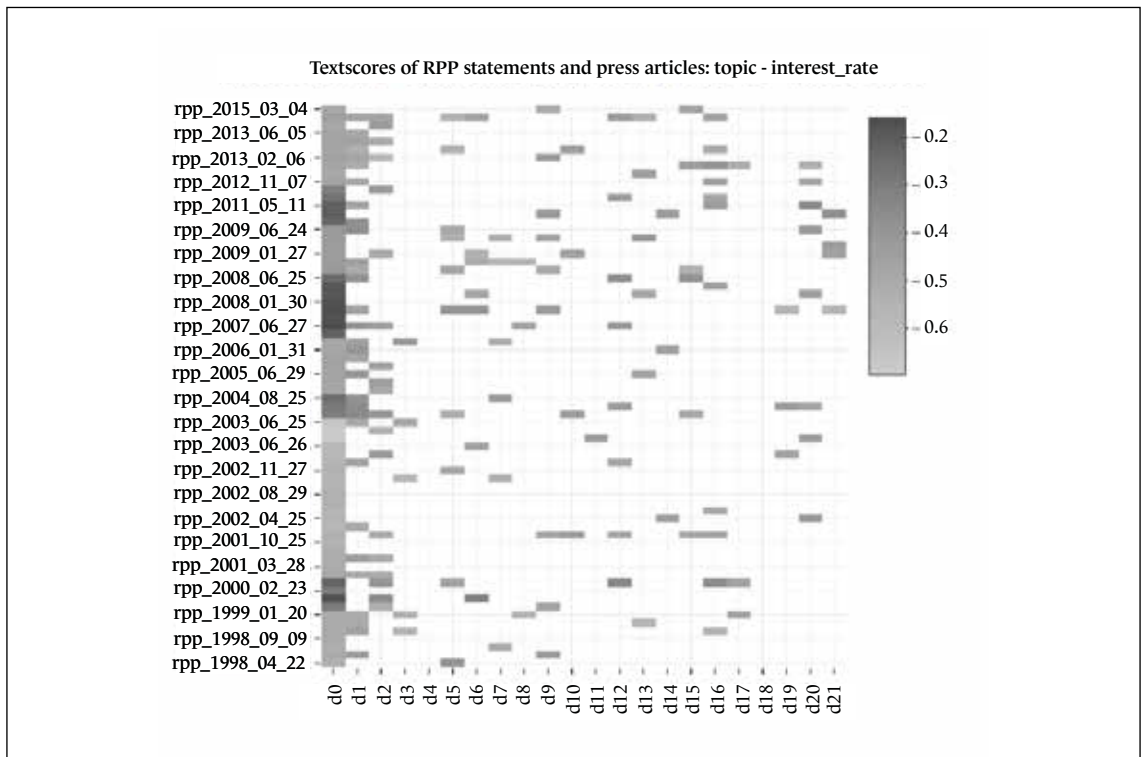
Note: for calculating average number of articles per topic before/after policy decision we use a 14-day window.

In the press articles in a 14-day time window financial markets is the dominant topic, followed by public finances, economy and the exchange rate. In RPP statements, as shown earlier, the most frequently discussed topic is inflation, then monetary policy and general economy. Table 2 documents that there is a certain asymmetry. General economy as a dominant topic appears only in policy statements following a rate cut. Inflation is a dominant topic in 50 per cent of policy statements following a rate hike, and only in less than a third of policy statements following a rate cut. Interestingly, in three policy statements following the rate cut public finance emerges as a key topic. It confirms that there were periods when public finance was a much more important factor in monetary policy decisions in Poland than one would typically observe in other countries following an inflation targeting strategy. All three policy statements with public finances as a key topic are from the period 2001–2002, where the post of NBP President was held by Leszek Balcerowicz, former (twice) finance minister who was very vocal about public finances and the need to curb public spending and advance structural reforms.

Inflation topic asymmetry found in RPP statements is also present in press articles. Within two weeks before a policy meeting that hiked interest rates there were on average 3.4 press articles discussing inflation, while this number fell to 1.9 before RPP rate cut meetings. However, after the decision the situation reverses, and there were more articles discussing inflation as a main topic after a rate cut (2.8) than after a rate hike (2.1).

We then apply the Wordscores model to understand the monetary policy inclination (hawkish/dovish) as described in RPP statements and press articles. Statements released after a rate hike have higher text scores (-0.23 on average) than statements released when the rate was reduced (-0.51 on average). However, even hawkish statements have negative text scores. There were more rate reductions (50) than rate hikes (20), so there are more words with negative than with positive score, which affects the text scores. The vast majority of newspaper articles have text scores between -0.7 and -0.3, but it includes many articles that mention NBP but do not discuss topics relevant from the monetary policy perspective. After applying topic dictionaries, we can split articles into ten topics and check the relationships between the text scores of monetary policy statements and those of newspaper articles discussing specific topics. Let us first illustrate this relationship by plotting a few examples. In Figure 3 we show a heatmap with text scores for articles discussing the interest rate that appeared within three weeks days after the policy meeting.

Figure 3
Heatmap presenting average text score of articles that appeared in a given day, where the article(s) main topic was the interest rate. Articles published within a 21-day period **after** the RPP meeting



Notes:

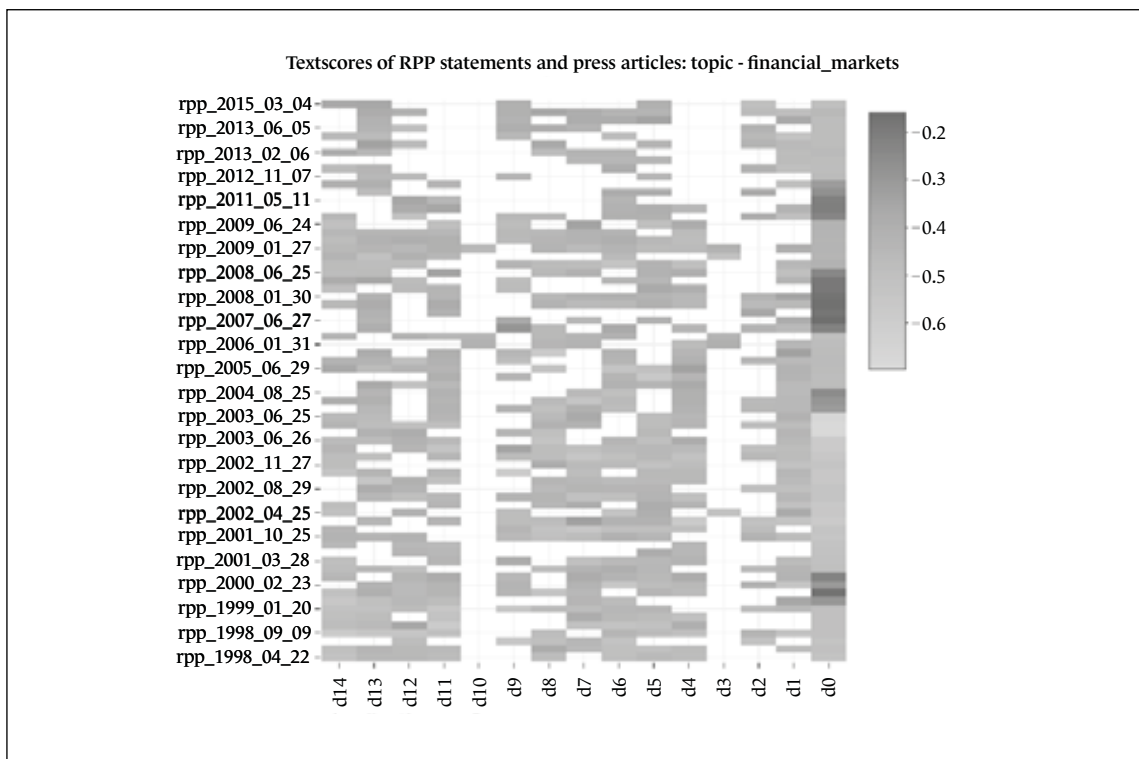
Only every second RPP meeting appears as the vertical axis label. RPP statements are shown in column d0. Grey colour intensity corresponds to the text score of each text (or average of articles published on the same day), as indicated by the scale. Darker shades indicate hawkish, brighter dovish policy inclination.

Several interesting observations can be made. Firstly, the interest rate is discussed during two days after the policy decision, and then appears rarely as a key topic. Secondly, articles following hawkish policy statements seem to have a darker grey colour than ones published after dovish statements. So it seems that RPP policy statements (official communication) are able to influence lexically the press articles (media discourse), or that the RPP policy stance is often shared by the general public (analysts, journalists). We later test these propositions more formally.

Another example compares text scores of articles discussing as a main topic financial markets that appeared within two weeks before the policy meeting.

Figure 4

Heatmap presenting average text score of articles that appeared in a given day, where the article(s) main topic was financial markets. Articles published within a 14-day period **before** the RPP meeting



Note: column d0 presents RPP statements.

This chart allows to examine whether words that were defined as “hawkish” or “dovish” in RPP policy statements appear in press articles before the policy meeting. The topic “financial markets” is especially relevant because often before the policy meeting the press discusses what decision is anticipated by the financial markets (futures and forward contracts) or the press often polls financial market analysts to present their rate decision expectations. As in the previous heatmap, here also the intensity of the articles’ grey colour on the heatmap tends to match – on average – the grey colour of the corresponding RPP statement.

So how we should interpret positive lexical sentiment correlation (co-occurrence of high and low text scores, respectively) between articles published before the policy meeting and the RPP policy statement released at that meeting? One should note that often words or memes created by central bankers become widely used in the media, as illustrated by two probably most famous examples: Greenspan's "irrational exuberance" and Draghi's "whatever it takes". Opposite situations, when press-invented memes become part of central bank official statements are possible, but rare.¹⁴ So it is more likely that positive lexical sentiment correlation shown in Figure 4 indicates successful central bank policy guidance rather than financial market tyranny.

Table 3 below summarizes these calculations and presents correlations between average text scores of newspaper articles published before and after a RPP policy meeting, and the text score of the RPP statement released after that meeting, grouped by the main topic of the articles. To verify the robustness of the obtained results, the calculations are conducted in several time windows, 21, 14 and 7 days after the rate decision, and 14 and 7 days before the rate decision.

Table 3

Lexical sentiment correlations between average text score of newspaper articles published before and after the RPP meeting, and text score of RPP policy statement released after that meeting. Articles grouped by their main topic

14 days before	7 days before	Topic	7 days after	14 days after	21 days after
0.22*	0.10	economy	0.43***	0.44***	0.43***
0.18	0.32**	exchange rate	0.37***	0.42***	0.42***
0.48***	0.41***	financial markets	0.06	0.26**	0.27**
0.14	0.06	financial stability	-0.09	0.06	0.09
0.36***	0.40**	inflation	0.50***	0.39***	0.34***
0.47***	0.47***	interest rate	0.52***	0.53***	0.52***
0.28**	0.43***	labour market	0.08	0.31**	0.32**
0.29**	0.56***	monetary policy	0.40**	0.39***	0.38***
0.26*	0.23	money	0.33**	0.26**	0.26**
-0.09	-0.18	public finances	0.16	0.05	0.13

Notes:

Correlation is calculated for all 70 RPP policy statements and corresponding averages of articles with a given main topic and in a given time window.

*** denotes statistical significance at 0.01 level,

** at 0.05 level,

* at 0.1 level.

¹⁴ One such rare situation took place when the author of this paper, working as ING Barings chief Polish economist asked a question during the RPP press conference: "what is the policy bias?". Such a concept was not officially used by the central bank. Boguslaw Grabowski answered the question stating that, "RPP maintained the tightening policy bias" and later the word "bias" appeared in the official RPP statements.

There are three groups of topics. Group one, where lexical sentiment correlations of articles and RPP statements are low or even negative and statistically insignificant: financial stability and public finances. Group two, when lexical sentiment correlations are high: interest rate, inflation and monetary policy. And group three, where lexical sentiment correlations vary depending on the analysed period, or exhibit asymmetries: economy, exchange rate, financial markets, labour market and to a lesser extent money.

Obtained lexical sentiment correlations results match very well the formal role of the central bank in Poland. While RPP publishes annually the opinion about the budget and financial stability report, it is the Ministry of Finance and Polish Financial Supervision Authority that are the most important institutions in these areas.¹⁵ Hence the limited role of the RPP, or NBP in general, in shaping the public discourse. On the contrary, in areas where RPP is the key decision-maker it is able to influence the media discourse: interest rate, inflation and monetary policy. Lexical sentiment correlation for the topic of exchange rate is somewhat lower before the meeting and high and significant after the meeting. On the one hand, the central bank adopted the floating exchange rate policy and – with rare exceptions – did not intervene through transactions or verbally on the forex market. On the other hand, the exchange rate transmission channel was important for many RPP members and the zloty exchange rate swings were often very large. So the lexical sentiment correlation results probably reflect this balance, when NBP tried not to comment too often about the specific levels of the exchange rate, and did acknowledge that zloty appreciation or depreciation had an important impact on economic activity and on inflation.

But probably the most interesting are two topics in group three. There is a weak lexical sentiment correlation for articles with economy as the main topic before the RPP meeting, and much stronger after the RPP statement was released. And the situation is opposite for the topic financial markets. It shows that a wide range of views on the economic situation is present before the RPP meeting, but afterwards the tone set by the central bank seems to dominate. The asymmetry in the financial markets topic, as already argued above, shows that the central bank has an efficient policy guidance practice.

To verify the consistency between official and unofficial central bank communication, we created a dictionary aimed at detecting articles that quote or refer to senior central bank officials. This dictionary contains the words: president, vice president, chairman, member, director, department. Figure 5 presents the text scores of articles published within two weeks after the RPP policy meeting, where the reference to a senior central bank official was detected.

We calculated lexical sentiment correlations between RPP statements and press articles quoting senior officials for different time windows before and after RPP meeting:

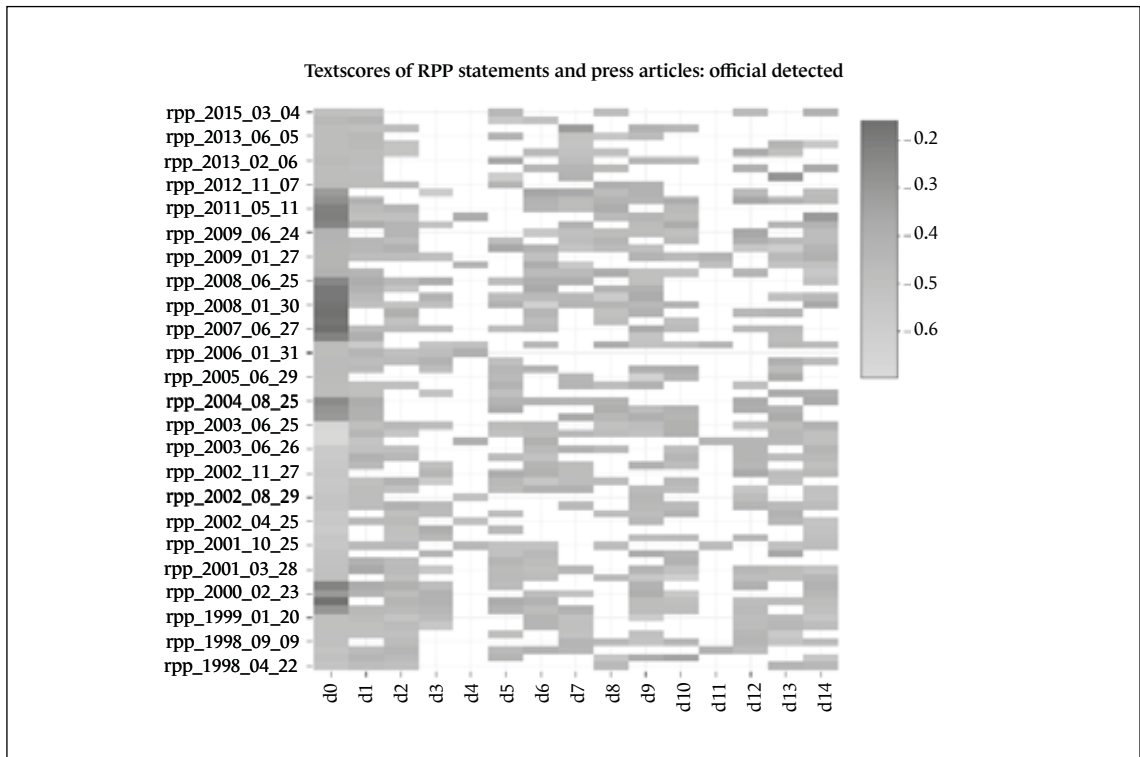
- after: one week (0.44***), two weeks (0.38***) and three weeks (0.36***);
- before: one week (0.26) and two weeks (0.29).

Articles published shortly after the RPP meeting are closest to the RPP statement in terms of lexical hawkish/dovish sentiment. It is not surprising, as some of such articles contain quotes from the RPP press conference. This correlation slowly decays over time. Interestingly articles published before the RPP meeting exhibit lower policy lexical sentiment correlation, and the correlation even falls closer to the meeting. This is consistent with the practice of communication with media adopted by the Polish

¹⁵ From January 1998 till December 2007 the banking sector was supervised by the Banking Supervisory Commission chaired by the President of NBP, and from January 2008 by the Financial Supervisory Authority, headed by its chairman appointed by the Prime Minister.

Figure 5

Heatmap presenting average text score of articles that appeared in a given day, where the article(s) quoted a senior central bank official. Articles published within a 14-day period **after** the RPP meeting



Note: column d0 presents RPP statements.

central bank. While any staff member has to obtain the NBP President's permission to officially appear in the media, RPP members are free to speak whenever they want.¹⁶ As the voting patterns reveal, RPP is often divided regarding members' views on what the appropriate level of interest rate should be. In line with the adopted communication model, RPP members use media channels to communicate their personal views, which are quoted by the media, especially shortly before the RPP meetings. It results in low and statistically insignificant lexical sentiment correlations before the policy meeting reported above.

5 Validation of the machine learning model results

The key challenge in applying supervised machine learning models to text analysis is whether it makes sense at all. To validate the automated analysis, usually a random sample of texts is selected, a human reader codes the texts, and later this coding is compared with the computer model results. We randomly selected one per cent of newspaper articles and eliminated those for which the main topic of the

¹⁶ With one exception: there were periods when the RPP internally decided that they will not publicly comment on interest rates one week prior to the policy meeting, but this moral commitment was not strictly obeyed.

text was not determined by the computer. There were 145 articles coded by the author. Among them the author could not determine the main topic of the text for 45 of them, so finally 100 texts were used for validation. The perceived policy inclination was coded in the range -1/easing, +1/tightening.

The Pearson correlation between computer and human assessment of the policy stance was 0.34 (p-value < 0.001). Agreement between computer and human on what is the key topic of the article was 71 per cent, which is remarkable given that there are ten topics to choose from and topic dictionaries contain relatively few words. The Pearson correlation between the number of senior official quotes found by the computer and human assessment if the central bank official spoke was 0.54 (p-value < 0.001). This validation exercise shows that very simple dictionaries can correctly classify the topic of the newspaper article and identify whether the central bank official was quoted in the text. The Pearson correlation for the monetary policy stance was somewhat lower, 0.34, but highly significant. We note that this lower correlation stems from the fact that it is difficult for the human reader to determine the policy stance described in the text. It depends on the relative importance of various monetary policy transmission channels, as judged by the reader. For example, what should be the policy stance assessment of the article that describes inflation being above the central bank target and a strong currency appreciation at the same time? Should we expect interest rates to be increased due to higher than desired inflation, or should interest rates be reduced if the central bank believes that strong currency appreciation can significantly slow economic activity in the future? Possibly different economists would manually code the same texts differently, depending on their background or views on what the key factors affecting the central bank decision process, inflation expectations channel or exchange rate channel are.

Also it is possible that this lower correlation is a result of changing communication of the central bank over time. For example there were periods where public finances were discussed in the policy statements, and many words related to public finances received a non-neutral score. Later public finances were not an important factor in policy decisions and did not appear in policy statements, but newspaper articles that mentioned public finances received a non-neutral policy inclination from the Wordscores model, which was trained on such data. In order to verify the changing communication hypothesis we trained the Wordscores model and conducted the validation exercise for a single RPP term: 1998–2004, which exhibits many interest rate decisions. Based on human coding of 2 per cent of the articles, the Pearson correlation coefficient for the monetary policy inclination assessed by human and by computer was 0.38 (p-value < 0.001), slightly higher than for the full period (0.34). Topics were correctly identified in 72 per cent of cases, similar to the accuracy achieved in the full data set.¹⁷ Because the correlation for the 1998–2004 subsample was only slightly higher than for the full sample, and the difference is not statistically significant, we would conclude that changing central bank communication content over the long term cannot be held responsible for lowering the correlation between human and machine understanding of the analysed texts.

Finally we checked whether human and computer agree on monetary policy inclination in cases where the computer had a “strong opinion”, i.e. the article score was away from the mean score for all articles. We selected articles with a computer policy score above and below one standard deviation

¹⁷ The majority of topic errors were when the human identified “exchange rate” and the computer “financial markets” as the main topic, or the other way around. In such cases, the articles often discussed both currency and bond market developments, so it was difficult to determine the main topic for a human reader.

from the mean, which left around 26 per cent of articles (5660).¹⁸ A validation exercise was performed on a random sample of 3 per cent drawn from these articles.¹⁹ The Pearson correlation between human and computer assessment was 0.51 (p-value < 0.001) and 90 per cent of topics were correctly identified. So when the computer has a strong opinion about the monetary policy inclination described in the press article, the human also tends to agree with such an opinion.

All three validation exercises together showed that the topics of the articles are correctly identified in the vast majority of cases and that computer and human tend to agree on the policy stance described in the text, especially in cases when the computer had a strong opinion about the policy stance.

6 Conclusions

In this paper we developed a machine learning framework based on simple dictionaries and the Wordscores model that allow to analyse the central bank communication, the monetary policy related discourse in the media, and the interactions between the two. We documented that the topics of both RPP statements and press articles mentioning the central bank gradually changed over time, with RPP statements becoming more balanced and newspaper articles focusing more on general economy and less on financial markets. We also showed that general economy as a dominant topic appeared only in policy statements following a rate cut. Inflation is a dominant topic much more often in RPP statements explaining a rate hike than ones accompanying a rate cut. Interestingly, in three policy statements following a rate cut public finance emerges as a key topic. Inflation topic asymmetry found in RPP statements is also present in newspaper articles. Inflation is a much more frequent topic before meetings that hiked the interest rate than before meetings that resulted in an interest rate cut.

We found that in areas where the central bank is not the key decision maker (public finances and financial stability) it has a very limited power to affect the media discourse. On the contrary, when the media discuss topics where the central bank is a key player or has a formal strong mandate (inflation, interest rates and monetary policy) we found high lexical sentiment correlation between policy statements and newspaper articles. To some extent it also applies to articles discussing the exchange rate. There is a weak lexical sentiment correlation for articles with economy as the main topic before the RPP meeting, and much stronger after the RPP statement was released. This shows that a wide range of views on the economic situation is present before the RPP meeting, but afterwards the tone set by the central bank seems to dominate. The opposite asymmetry in the financial markets topic shows that the central bank has an efficient policy guidance practice. When analysing press articles quoting central bank officials we found that the official and unofficial communication is internally consistent in the short run. However, it also leaves some room for presenting different opinions by RPP members. Results presented in the paper were positively human-validated by three validation exercises.

This research suffers from weaknesses that are typical for all “bag of words” models applied to unigrams. They cannot understand negation, irony or sarcasm. Therefore applying the machine learning framework described here to 2-grams or using the word2vec approach seems a promising future research opportunity. Also applying this framework to several central banks, especially ones

¹⁸ For normal distribution more than one standard deviation from the mean leaves 32 per cent of observations. It shows that in our case there are much more articles with neutral monetary policy stance determined by the computer than predicted by normal distribution.

¹⁹ After removing articles where the key topic could not be determined, the validation was applied to 104 articles.

pursuing different monetary policy targets, could yield important insights into the different styles of communication.

Despite the described weaknesses, this framework would allow central banks to effectively monitor how their official communication interacts with the media discourse on monetary policy.

References

- Bennani H. (2015), *The art of central banks' forward guidance at the zero lower bound*, SSRN Working Paper, June.
- Boukous E., Rosenberg J. (2006), *The information content of FOMC minutes*, Federal Reserve Bank of New York Working Paper, July.
- Blinder A.S., Ehrmann M., Fratzscher M., De Haan J., Jansen, D.-J. (2008), Central bank communication and monetary policy: a survey of theory and evidence, *Journal of Economic Literature*, 46(4), 910–945.
- Canzoneri M. (1985), Monetary policy games and the role of private information, *American Economic Review*, 75(5), 1056–1070.
- Carpenter S.B. (2004), *Transparency and monetary policy: What does the academic literature tell policymakers?*, Federal Reserve Board Finance and Economics Discussion Series, 35.
- Connolly E., Kohler M. (2004), News and interest rate expectations: a study of six central banks, Federal Reserve Bank of Australia Working Paper, 10.
- Correa R., Garud K., Londono J.M., Mislang N. (2017), *Sentiment in central banks financial stability reports*, International Finance Discussion Papers, Board of Governors of the Federal Reserve System, 1203.
- Doryń W. (2018), Analiza wydźwięku dyskusji na posiedzeniach decyzyjnych Rady Polityki Pieniężnej, *Wiadomości Statystyczne*, 63, 7(686), 72–85.
- Eijffinger S., Hoerberichts M., Schaling E. (2000), Why money talks and wealth whispers: monetary uncertainty and mystique, *Journal of Money, Credit, and Banking*, 32(2), 218–235.
- Ehrmann M., Fratzcher M. (2005a), *Communication and decision-making by central bank committees: different strategies, same effectiveness?*, European Central Bank Working Paper, 488.
- Ehrmann M., Fratzcher M. (2005b), *How should central banks communicate?*, European Central Bank Working Paper, 577.
- Faust J., Svensson L. (1998), *Transparency and credibility: monetary policy with unobservable goals*, National Bureau of Economic Research Working Paper, 6452.
- Geraats P. (2002), Central bank transparency, *Economic Journal*, 112(483), F532–F565.
- Gerlach S. (2004), *Interest rate setting by the ECB: words and deeds*, Centre for Economic Policy Research Discussion Paper, 4775.
- Guthrie G., Wright J. (2000), Open mouth operations, *Journal of Monetary Economics*, 46(2), 489–516.
- Grimmer J., Stewart B. (2013), Text as data: the promise and pitfalls of automatic content analysis methods for political texts, *Political Analysis*, 1–31.
- Gürkaynak R., Sack B., Swanson E. (2005), Do actions speak louder than words? The response of asset prices to monetary policy actions and statements, *International Journal of Central Banking*, 1(1), 55–93.

- Hahn V. (2002), Transparency in monetary policy: a survey, *Ifo Studien Zeitschrift für empirische Wirtschaftsforschung*, 3, 429–455.
- Hansen S., McMahon M., Prat A. (2018), Transparency and deliberation within the FOMC: a computational linguistics approach, *The Quarterly Journal of Economics*, 133(2), 801–870.
- Hendry S. (2012), *Central bank communication or the media's interpretation: What moves markets?*, Bank of Canada Working Paper, 2012-9.
- Jansen D.-J., de Haan J. (2010), *An assessment of the consistency of ECB communication using Wordscores*, DNB Working Paper, 259.
- Jensen H. (2002), Optimal degrees of transparency in monetary policy-making, *Scandinavian Journal of Economics*, 104(3), 399–422.
- Kohn D., Sack B. (2003), *Central bank talk: Does it matter and why?*, The Federal Reserve Board Finance and Economics Discussion Series, 55.
- Laver M., Benoit K., Garry J. (2003), Extracting policy positions from political texts using words as data, *American Political Science Review*, 97(02), 311–331.
- Luangaram P., Wongwachara W. (2017), *More than words: a textual analysis of monetary policy communication*, PIER Discussion Papers, 54, Puey Ungphakorn Institute for Economic Research.
- Lüdering J., Tillmann P. (2016), *Monetary policy on Twitter and its effect on asset prices: evidence from computational text analysis*, MAGKS Papers on Economics, 201612, Philipps-Universität Marburg, Faculty of Business Administration and Economics, Department of Economics (Volkswirtschaftliche Abteilung).
- Masawi B., Bhattacharya S., Boulter T. (2016), The power of words: a content analytical approach examining whether central bank speeches become financial news, *Journal of Information Science*, 40(2), 198–210.
- Moniz A., de Jong F. (2014), Predicting the impact of central bank communications on financial market investors' interest rate expectations, in: V. Presutti, E. Blomqvist, R. Troncy, H. Sack, I. Papadakis, A. Tordai (eds), *The Semantic Web: ESWC 2014 Satellite Events. ESWC 2014. Lecture Notes in Computer Science*, Springer.
- Ortiz A., Rodrigo T., Turina J. (2017), *How do the central banks talk? A big data approach to Turkey*, BBVA Research, 17 October.
- Rozkrut M., Rybinski K., Sztaba L., Szwaja R. (2007), Quest for central bank communication: Does it pay to be “talkative”?, *European Journal of Political Economy*, 23(1), 176–206.
- Rybinski K. (2017), Sentiment analysis of Polish politicians, *e-Politikon*, 24, 162–195.
- Rybinski K. (2018), Political sentiment analysis of press freedom, *Studia Medioznawcze*, 3(74), 31–48.
- Siklos P.J., Sturm J.-E., eds (2013), *Central bank communication, decision making, and governance*, CESifo Seminar Series, The MIT Press.
- Wan M. (2017), *Malaysia: reading between the lines – text mining the central bank*, Credit Suisse Fixed Income Research, Emerging Markets, 27 November.
- Woodford M. (2005), *Central bank communication and policy effectiveness*, paper presented at the Federal Reserve Bank of Kansas City Conference: “The Greenspan Era: Lessons for the Future”, Jackson Hole, August 25–27.

Acknowledgements

Author would like to thank two anonymous referees for careful reading of the manuscript and for giving many constructive remarks that substantially helped to improve the quality of the paper.

