



---

*Research article*

## **The informative value of central banks talks: a topic model application to sentiment analysis**

**Maria Paola Priola, Annalisa Molino, Giacomo Tizzanini and Lea Zicchino\***

Prometeia S.p.A. Piazza Trento e Trieste, 3, 40137 Bologna, Italy

\* **Correspondence:** Email: [lea.zicchino@prometeia.com](mailto:lea.zicchino@prometeia.com).

**Abstract:** Central banks communication has lately become an important tool to guide expectations and its impact on the economy has been acknowledged by the literature. Nowadays central banks speeches face an increasing variety of topics, which are not discriminated by text analysis. In this paper we build a topic-weighted central bank sentiment index as a combination of machine learning and text analysis techniques to investigate large datasets. First, we develop a methodological framework to grid search the best Latent Dirichlet Allocation (LDA) model to uncover the latent topics in central banks' speeches and releases published between 2000 and 2021. Then, we build a topic-specific sentiment index based on dictionary techniques. Next, we summarise the results in a topic-weighted Central Bank Sentiment Index (CBSI<sub>w</sub>) for the Bank of Canada (BoC), the Bank of England (BoE), the European Central Bank (ECB) and the Federal Reserve (Fed). We find that the main common driver of the CBSI<sub>w</sub> is the monetary policy topic, followed by macroprudential policy and payments and settlements. We also uncover bank-specific topics and topics related to new challenges, for example innovation and climate change. Moreover, we find that the CBSI<sub>w</sub> decreases after the Great Recession, signalling a worsening in sentiment, as well as during the COVID-19 crisis. Finally, we employ a probit regression to further assess the predictive power of our monetary policy topic-specific index. We find that the indicator helps predicting future changes in policy rate, corroborating the evidence that central banks communication signals future monetary policy decisions.

**Keywords:** central banks communication; monetary policy; Latent Dirichlet Allocation; text analysis; probit model

**JEL classification:** D78, E52, E58, F42

---

## 1. Introduction

Over the last decades central banks have increasingly relied on communication as a tool for monetary policy and, since the financial crisis, there has been a trend towards an increased level of transparency. Information has been disclosed in official documents, during broadcast interviews and on social networks. The impact of this effort has been widely acknowledged in the literature. Similarly, there is an increasing interest of policymakers and market operators in the economic “sentiment” expressed by central banks in their communication. For this reason, a flourish of literature on text analysis applied to central banking has emerged in recent years, since with this technique it is possible to obtain a metric of the qualitative information embedded in communication (Loughran and McDonald, 2011; Picault and Renault, 2017; Correa et al., 2017). The sentiment indicators so obtained are used to investigate the effects of central banks’ communication on macroeconomic and financial variables or to show interconnections among central banks’ policies (Armeliuss et al., 2020; Picault and Renault, 2017). Moreover, these indicators work as a proxy of survey-based indices on economic sentiment (Shapiro et al., 2020). However, one drawback of the lexicon-based technique with large datasets is that sentiment is created without extrapolating the context in which the words are used. Therefore, it provides a sentiment metric of a given speech regardless of the subject.

In this paper we contribute to the literature on central banks communication along several dimensions. We merge topic modelling and text analysis techniques to create a topic weighted Central Bank Sentiment Index (weighted CBSI or CBSIw) for the Bank of Canada (BoC), the Bank of England (BoE), the European Central Bank (ECB) and the Federal Reserve (Fed). To start with, we apply Latent Dirichlet Allocation (LDA), introduced by Blei et al. (2003) to uncover the latent topics in central banks communication and investigate their evolution over time. Given that LDA requires a priori assumptions on the model setup and there is no guidance for the optimal hyperparameters choice, we develop a procedure based on the grid search method. Next, we build a topic sentiment index with a lexicon methodology in line with Priola et al. (2021) to monitor the sentiment relative to each learned topic. Then, we combine the indexes to obtain the CBSIw that informs us not only about the economic sentiment of central banks’ communication, but also it informs us about the relative importance of sub-topics. Furthermore, the sub-indexes can be monitored to investigate the sentiment on specific arguments, e.g. the climate change. The indicator is calculated as weighted monthly average of the topics, where the weights are given by the number of speeches on each topic, each month. Finally, we examine the monetary policy predictability by employing a probit model to regress our monetary policy sentiment indicator on future policy rate decision. Our main findings are the following. First, the main common driver of the weighted CBSI is the monetary policy topic, followed by macroprudential policy and payments and settlements. There are only a few topics that are not relevant for all four central banks. This is the case of the promotion of consumer protection for the Fed and the need of financial integration for the ECB, while international affairs is more relevant for the BoC. Second, the weighted CBSI decreases for all banks after the Great Recession, signalling a worsening in sentiment, and drops during the pandemic crisis. Third, probit estimation shows that the monetary policy topic-indicator is positive and significant for the all Central Banks with the exception of Bank of Canada, giving support to the claim that the monetary policy communication helps predicting future changes in policy rates.

The rest of the paper is organised as follows. Section 2 reviews the empirical literature. Section 3 describes the collection of documents that were analysed and provides descriptive statistics about

the panel of Central Banks. Section 4 discusses the methodology to identify the topics and build the weighted CBSI. Section 5 presents the LDA clustering results and Section 6 provides evidence on the predictive power of our sentiment indicator. Section 7 concludes.

## 2. Literature review

Our paper mainly contributes to the literature on central banks communication and text analysis, or Natural Language Processing (NLP). Text analysis techniques allow to automatically organise, analyse and summarise large quantities of information. Furthermore, they have the advantage to eliminate the subjectivity associated with the human reading approach. The related literature is growing together with the increased amount of central bank communication, which make reading through all the documents challenging.

One approach is to use dictionary-based techniques following Loughran and McDonald (2011). Priola et al. (2021) construct their own Central Bank Sentiment Indexes (CBSIs) applying text analysis to a large dataset of speeches and releases published by for the Federal Reserve (Fed), the European Central Bank (ECB), the Bank of England (BoE) and the Bank of Japan (BoJ) between 2000 and 2020. This method consists in automatically matching a set of words with a pre-given connotation (a dictionary or lexicon), to words in the text of interest. After having determined the frequency of words with their particular connotation, the literature suggests several approaches to build an indicator. The authors proposed the Net Score Index, similar to the approach developed by Birz and Lott Jr (2011) to build their CBSIs.

An appealing feature of dictionary-based techniques is their simplicity. However, a major limitation of text analysis is that it does not capture the context in which the words are used. Other approaches, such as topic modelling, address this issue. A topic model adopts a flexible statistical structure that clusters words together to form “topics”, allowing the same word to belong to more than one topic. Such approach is used to uncover the hidden semantic structure in a text. A common approach is the Latent Dirichlet Allocation (LDA), introduced by Blei et al. (2003). This algorithm has been widely used in social and behavioural science, and lately, to study central bank communication. Machine learning algorithms such as the LDA model can be used to study and uncover topic trends within central banks communication. Given the increase in the amount of speeches and releases and their always increasing importance as a monetary policy tool, the ability to monitor such communications is potentially useful for real-time analysis. This technique can be also used to monitor topic trends over time.

Hansen et al. (2018) employ text analysis to study how greater external communication affects the internal deliberation of monetary policymakers in the Federal Open Market Committee (FOMC). Their database is made of 149 FOMC meeting transcripts from August 1987 to January 2006, under the chairmanship of Alan Greenspan. They construct measures of communication based on both basic text counts and on topic models and use them as a dependent variable in an econometric model. Their LDA model is estimated for all the statements in each meeting in the whole sample, while the econometrics analysis focuses on the statements made in each meeting during the stage of the economic situation discussion and the policy discussion. Tumala and Omotosho (2019) choose LDA to analyse the communication strategy of the Central Bank of Nigeria (CBN) during the period 2004-2019. They base their analysis on 87 policy communiqués released after the meetings of the CBN’s monetary policy committee and published by the Bank on its website. They estimate a LDA model to identify the

key topics driving monetary policy communication during the sample period. Thus, they analyse the evolution of the topics by plotting the topic proportion for the identified topics. Their results show that the topics are influenced by the Bank's announced policy objectives as well as the nature of shocks hitting the economy per period. Keida and Takeda (2018) apply LDA to 70 regular press conference documents of the Bank of Japan's (BOJ) governors Masaaki Shirakawa and Haruhiko Kuroda from July 2012 to July 2018. The authors find that NLP techniques are useful for identifying changes in monetary authorities, such as when Shirakawa succeeded Kuroda. Carboni et al. (2020) investigate the content of the ECB and Fed Governors' speeches for the period 2007-2019 using LDA. Their database consists of 313 speeches of Governors Jean Claude Trichet and Mario Draghi of the ECB and 193 speeches of Governors Ben Bernanke, Janet Yellen and Jerome Powell of the Fed. In the speeches of both central banks they find that topics linked to traditional subjects like monetary policy and financial stability play the main role. They find small tracks of topics related to new challenges, for example environmental sustainability.

To investigate the predictive power of our weighted central bank indicators we use a probit model, which is a standard methodology to assess the significance of central bank communication\*, as shown by recent research. Apel and Grimaldi (2014) examine whether a rate policy decision can be explained by their hawk/dove sentiment indicator for the Swedish Riksbank, after controlling for the lagged dependent variable and macroeconomic variables, and discover that their measure helps to predict policy decisions. Cour-Thimann (2020) estimate the ECB reaction function augmented with their communication indicator, finding that sentiment estimated from minutes significantly improve the reaction function. Bennani et al. (2020) investigate the ECB communication as a signal of monetary policy change proxied by MRO rate or the shadow rate, by controlling for the zero-lower bound (ZLB) and financial market expectations. Their results suggest that communication about conventional measures helps explaining the future ECB rate changes, even controlling for the ZLB period.

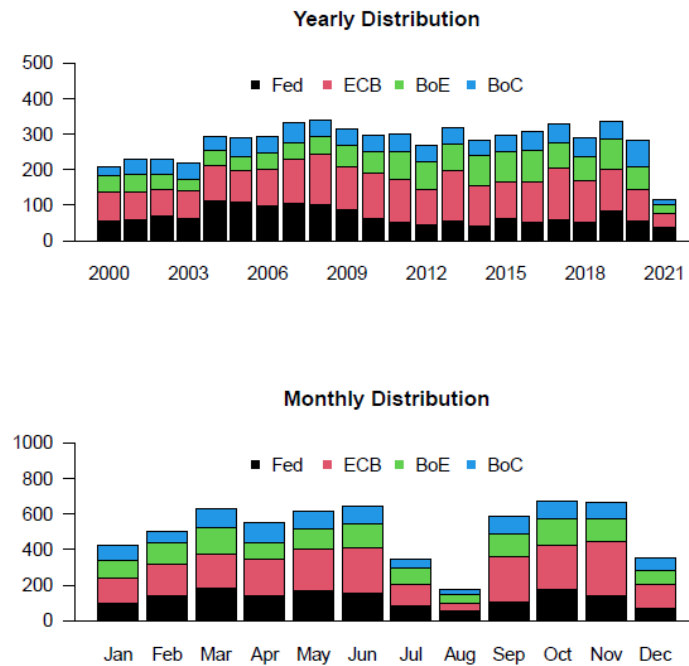
### 3. Data analysis

Our dataset includes speeches and releases available for download from the institutional websites of the European Central Bank, the Bank of England, the Bank of Canada and the Federal Reserve. The choice relies on the fact that these countries are the world's largest advanced economies according to the IMF; therefore, the global economy is largely influenced by the communication of their central banks, which it is worth being investigated. For the Fed, we also include the FOMC meeting transcripts available for download from the website of the Federal Reserve Bank of St. Louis. We only extract the speech transcription, avoiding page headers and footnotes. Speeches can vary in length, depending on the event at which the speech is delivered; releases are typically short and standardized, as they are the summary of monetary policy decisions or press releases with a focus on monetary policy. Overall, our dataset includes 6137 documents spanning the period from January 2000 to April 2021 and distributed as follow: 1007 for the BoC, 1319 for the BoE, 2306 for the ECB and 1505 for the Fed.

The monthly and yearly distributions of our dataset are shown in Figure 1. It is worth mentioning that the number of documents published by the Banks steadily increased from the first half of the 2000s, later stabilizing in the 2010s. This finding is in line with the increase in central banks' transparency documented by Dincer and Eichengreen (2013). We notice a substantial increase in publications between

---

\*See Gerlach-Kristen (2004)



**Figure 1.** Yearly and monthly distribution of speeches and releases for each central bank.

2006 and 2010 for the ECB, the BoC and the BoE, signalling that the financial crisis and the Great Recession had an impact on central banks' transparency. Instead, the largest increase for the Fed is observable during the last years of Greenspan's chairmanship (2000-2006).

We clean and process the documents by applying a set of text mining techniques, along the lines of the literature in the field of NLP (Silge and Robinson, 2017). Such transformations are needed to carry out the empirical analysis. In detail, the documents are converted in text-corpus following a methodology detailed by Grün and Hornik (2011). We clean the corpus by lowercasing the words and removing numbers, extra spaces, punctuation, regular expressions and standard stopwords. After this parsing phase, we tokenize\* the text and extract unigrams and bigrams†. Then, we create a list of bank-specific stopwords including Governors' names and non-English words, leaving out non-English common words with economic relevance (such as COVID-19 and Brexit). Finally, we process the tokens with the Porter stemming algorithm (Porter, 1980) to eliminate common suffixes and prefixes. As a result of this phase, we achieve a document-term matrix.

Next, we prune‡ the dataset using the Terms frequency-Inverse document frequency (Tf-Idf) function (Baeza-Yates and Ribeiro-Neto, 1999) calculated on each central bank's document term matrix. The Tf-Idf is a function that weighs the importance of a term within to the whole dataset (Leskovec et al., 2020). The function is expressed as:

$$(tf - idf)_{i,j} = tf_{i,j} * idf_i \quad (1)$$

\*Tokenization is the process of breaking the text into blocks readable by the software.

†n-grams are series of n adjacent elements from a given string of tokens.

‡Pruning is the process of reducing the dataset size to improving the learning of a machine learning model.

with

$$tf_{i,j} = \frac{n_{i,j}}{|d_j|}; \quad idf_i = \log\left(\frac{|D|}{|\{d : i \in d\}|}\right) \quad (2)$$

where  $tf_{i,j}$  stands for Term frequency of word  $i$  in the document  $j$ ;  $n_{i,j}$  is the number of occurrences of word  $i$  in the document  $j$ , while  $|d_j|$  is the length of document  $j$  as the sum of its terms. In addition,  $idf_i$  stands for Inverse document frequency of word  $i$  in the collection of documents  $D$  and  $|D|$  is the number of the documents in the sample. The denominator measures the number of documents in which  $i$  is present. This function increases the weight of terms that occur very frequently in a given document but that are less frequent in the whole collection of documents (Grün and Hornik, 2011). As a result, common words are filtered out and the resulting word rank is more meaningful than the absolute frequency approach\*. Grün and Hornik (2011) clean the dataset by omitting all terms with a Tf-Idf value below 0.01. We fine-tune their approach by choosing a flexible threshold specific to each document, rather than pruning all the dataset for a fixed value. In fact, for each document, we discard all the terms having a Tf-Idf value below the median of the word distribution in a given document. Our final dataset includes 4267 words for the BoC, 8972 words for the BoE, 8363 words for the ECB, 8075 words for the Fed.

In Figures A1–A4 in Appendix we provide a visual representation of the pruned dataset for each bank using word clouds, including the 100 most important words according the Tf-Idf. In the word clouds, the rank given by the Tf-Idf is represented by the size of the word itself. Many terms are common almost to all banks: notably, “monetary policy”, “inflation”, “risk” and “rate. Nevertheless, there are some things to notice: while “inflation” is among the first words extracted for the Fed and the BoC, this is not the case for the ECB and the BoE. Furthermore, for the Fed, both the words “economy” and “market” are more frequent than for the other banks. Finally, it is also interesting to note the high frequency of the term “treaty/ies” for the ECB, because policy makers need to refer quite often to the legal basis of the institution.

Finally, we investigate each central bank’s text distribution according to Zipf’s Law (Zipf, 1936). Zipf’s law states the inverse relationship between the frequency of a term and its rank, Figures A5–A8 in Appendix show the theoretical Zipf’s curve compared to the empirical ones calculated on the corpus pre and post pruning, finding the Zipf’s law holds when pruning the dataset.

#### 4. Methodology

In this paper we use the Latent Dirichlet Allocation, introduced by Blei et al. (2003). LDA is a generative probabilistic model. The basic idea is that documents are represented as random mixtures over latent topics, and each topic is characterized by a distribution over words. The model takes a corpus as an input and gives as output the distribution of topics for each document in the corpus and the distribution of words for each topic. Figure 2 shows the plate notation for the model. Assuming that the corpus consists of  $M$  document, each of length  $N_i$ , LDA assumes the following generative process.

- For each document, choose  $\theta_i \sim \text{Dirichlet}(\alpha)$  for  $i=1, \dots, M$ , where  $\alpha$  is the concentration parameter of the Dirichlet prior on the per-document topic distribution.
- For each topic, choose  $\varphi_k \sim \text{Dirichlet}(\beta)$  for  $k=1, \dots, K$ , where  $\beta$  is the concentration parameter of the Dirichlet prior on the per-topic words distribution.

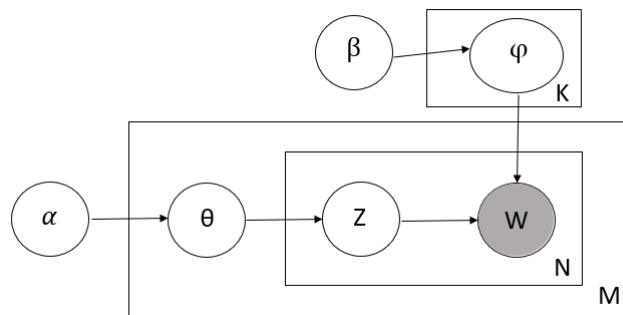
\*The rank is a list of words in order of decreasing frequency, i.e. from the most to the less used word.

- For each word position  $(i, j)$ , where  $i=1, \dots, M$  and  $j=1, \dots, N_i$ 
  - Choose a topic  $z_{i,j} \sim \text{Multinomial}(\theta_i)$
  - Choose a word  $w_{i,j} \sim \text{Multinomial}(\varphi_{z_{i,j}})$

The words  $w_{i,j}$  are the only observable variables, while  $\theta$ ,  $\varphi$  and  $z$  are latent variables. Going through the three steps above, the algorithm infers the posterior distribution of the latent variables:

$$p(\theta, \varphi, z|w, \alpha, \beta) = \frac{p(\theta, \varphi, z|\alpha, \beta)}{p(w|\alpha, \beta)} \quad (3)$$

After the posterior estimates for  $\varphi$  and  $\theta$  are obtained, the algorithm returns the topic representation of each document in the corpus.



**Figure 2.** Plate notation for LDA.

LDA is a complex algorithm which is generally hard to fine-tune and interpret. Besides the text cleaning and pre-processing, getting relevant results with LDA depends heavily on the model setup:  $\alpha$  and  $\beta$  which must be provided as input to the algorithm, as well as the number of topics  $K$ . The hyperparameters  $\alpha$  and  $\beta$  affect the sparsity of the Dirichlet distributions. With a higher  $\alpha$ , documents are assumed to be composed of more topics and this results in a more specific per-document topics distribution. This means that when choosing a low  $\alpha$  it is more likely that a document may contain just a mix of few topics. Likewise, with a higher  $\beta$ , topics are assumed to be composed of more words and this results in a more specific per-topic words distribution. With the last input parameter  $K$ , the LDA model clusters documents into different topics.

#### 4.1. Model estimation

In section three, we described the procedure we adopt to clean and pre-process the data. In this section, we explain how we select the hyperparameters and estimate the LDA for the banks in our sample\*.

In literature there is unfortunately no guidance for the choice of the hyperparameters. Following the general advice of Griffiths and Steyvers (2004) and Navarro et al. (2006), Hansen et al. (2018) set  $\alpha = \frac{50}{K}$  and  $\beta = 0.025$ , and estimate the model for  $K = 50$  and  $K = 70$ . Tumala and Omotosho (2019) and Keida and Takeda (2018) heuristically assume a 6 topic model and a 3 topic model, respectively.

\*To run the model, we use Gensim library in Python.

Carboni et al. (2020) look at a 5-topic and 10-topic model and, based on human judgement, pick the 10-topic model. In this paper we mainly rely on topic coherence measures, which provide a convenient way to assess the quality of the topic extracted. They measure the degree of semantic similarity between the most frequent words in a topic and allow to separate topics that are interpretable from those that are purely statistical. The  $c_v$  measure, introduced by Röder et al. (2015), is a confirmation measure that combines normalized pointwise mutual information (NPMI) with the cosine vector similarity. Given a word  $w_j^k$ , let  $v^j$  be defined as:

$$v_i^j = NPMI(w_i, w_j)^y = \left( \frac{\log \frac{p(w_i, w_j) + \epsilon}{p(w_i)p(w_j)}}{-\log(p(w_i, w_j) + \epsilon)} \right)^y \quad i = 1, \dots, N \quad (4)$$

$c_v$  measure is the arithmetic mean of the top words cosine similarities:

$$\cos(u, v) = \frac{\sum_{i=1}^N u_i v_i}{\sqrt{\sum_{i=1}^N u_i^2 \sum_{i=1}^N v_i^2}} \quad (5)$$

As said, there is no benchmark for the choice of optimal LDA hyperparameters and therefore the selection is up to the researcher. Hence, one of our attempts is to outline a procedure that one could easily follow with less involvement of human judgement. Our approach for finding the optimal hyperparameters is to perform an exhaustive grid search. Going through a specified subset of the hyperparameters space, we train several models and generate candidates  $\alpha$ ,  $\beta$  and  $K$ . As a final step, we pick the combination of  $\alpha$ ,  $\beta$  and  $K$  that optimize the coherence scores. In detail, based on human judgement, we define a range in which the hyperparameters can vary. After investigating each central bank's institutional website, we assume our dataset to be made up of only a few labels and our words to belong to some of the topics. With this in mind,  $\alpha$ ,  $\beta$  are typically set below one. The maximum number of topics  $K$  is set to 70, given that a larger number of topics may hamper interpretability. Thus, for each  $K$ , we train the model for  $\alpha \in [0.01, 1]$  in steps of 0.1. For each  $K$  and  $\alpha$ , we train the model for  $\beta \in [0.01, 0.41]$  in steps of 0.1. We train each model on 90% of the documents and, for each of them, we calculate  $c_v$ . We estimate the posterior distribution of the latent variables using the online variational Bayes inference algorithm developed by Logsdon et al. (2010)\*.

After training all the models, for each  $K$  we pick the pair  $(\alpha, \beta)$  that optimize the coherence scores. Our final choice is to use the pair  $(\alpha^*, \beta^*)$  for which  $c_v$  is optimized more frequently across all  $K$ . Thus, plotting the  $c_v$  curve, we pick  $K^*$  around the optimal values of  $c_v$  that best fit interpretability, based on our sensitivity. Finally, we examine the resulting topics and the associated top ten words. The probabilities associated to each word reflect how important the word is to that topic. We cluster the topics in  $\bar{K}$  bank-specific labels to obtain a list of topics that are interpretable and exhaustive.

Once the LDA is trained, we test the model on unseen documents (10% of the corpus). The testing phase consists in reading the speeches and judge whether the label assigned by the model is coherent with the effective topic. For all the banks in the sample, the percentage of documents that are misclassified by the algorithm is 5% on average.

\*We set the following values for the training parameters: iterations = 300 (maximum number of iterations through the corpus when inferring the topic distribution of the corpus), chunk size = 50% of the collection (number of documents to consider at once during the training), passes = 10 (how many times the algorithm is supposed to pass over the whole corpus during the training).



#### 4.2. Topic weighted CBSI

For each central bank we build a CBSI for each label, with the methodology described in Priola et al. (2021). We choose the Loughran McDonald dictionary\* motivated by the broad consensus in the literature on the validity of this dictionary when applied to finance text analysis (Shapiro et al., 2020; Armelius et al., 2020). Thus, we use the Loughran-McDonald dictionary to assign a value of one if the connotation of the words in a document is positive, minus one if it is negative. We do not consider neutral words. After having determined the frequency of words with the positive and negative connotation, we calculate a normalized score as follows:

$$NetScoreIndex_t = \frac{\#Positive_t - \#Negative_t}{\#Positive_t + \#Negative_t} \quad (6)$$

where the numerator is the difference between the total number of positive and negative words, while the denominator is the total number of positive and negative words. This score (named Net Score from now on) has the advantage of being comparable even among texts with different length without creating distortions in the indicator. The range<sup>†</sup> of the Net Score Index is between -1 and 1. High scores indicate a positive view of the economic context by the institution, whereas low levels denote a negative outlook.

Using the subset of documents belonging to the label  $k$  we build the  $CBSI_t^k$ ,  $k=1, \dots, \bar{K}$ . Thus, we aggregate in monthly averages the indexes. As a final step, we calculate the weighted average:

$$CBSIw_t = \frac{\sum_{k=1}^{\bar{K}} w_{k,t} CBSI_t^k}{\sum_{k=1}^{\bar{K}} w_{k,t}} \quad (7)$$

where  $w_{k,t}$  is the ratio of the number of documents in which the topic  $k$  appears to the total number of documents at month  $t$ .

## 5. Results

In this section we provide first a description of the optimal LDA hyperparameters and the learned topics. Second, we discuss the results of our analysis for each central bank in the sample.

As explained in Section 4, we choose the pairs  $(\alpha^*, \beta^*)$  that most frequently optimize  $c_v$ . Having fixed the hyperparameters, we calculate the coherence measure for each  $K$  and use it as a guidance for choosing  $K^*$ . We look at the models for which the resulting coherence function is relatively higher and opt for the  $K^*$  that delivers a list of exhaustive and interpretable topics. As a result of our experimental procedure, we find that the optimal values are those shown in Table 1. It is worth noting that for the ECB the optimal  $\alpha$  and  $\beta$  are the lowest of the gridsearch (0.01 and 0.11 for  $\alpha$  and  $\beta$ , respectively). This might be explained by the larger number of documents and the multitude of speakers that characterize the ECB dataset.

\*To better reach out the sentiment of the last year, we augmented the dictionary with the words “pandemic” and “coronavirus” among the negative ones; this negative connotation is only temporary and subjected to further research of the authors, with the aim of enriching the standard dictionary.

†The Net Score can show extreme values, i.e. with very short speeches where one or two words are captured in the text. In this case, we opt for NaN values to avoid excessive noise.

**Table 1.** Summary results for LDA training.

Central Bank	K*	$\alpha^*$	$\beta^*$
BOC	55	0.41	0.31
BOE	35	0.61	0.31
ECB	60	0.01	0.11
FED	30	0.61	0.21

For each bank, we cluster the topics resulting from the model training and assign a label to each cluster. Overall, we create eleven labels, described below.

*Climate Change (CLIMATE)*. Documents on climate change issues related to monetary policy and the economy.

*Consumer Protection and Community Development (CONSCOM)*. As stated by the Federal Reserve itself, documents about promoting Consumer Protection and Community Development advances supervision, community reinvestment, and research to increase understanding of the impact of financial services policies and practices on consumers and communities\* with the aim of promoting financial literacy.

*EMU Affairs (EMU\_AFF)*. Documents on European Monetary Union whose contents examine themes such as the affiliation of new member State and/or the adoption of the euro in a country, as well as issues related to the European governance.

*Financial Integration (FINTEGR)*. Documents on developments of a well-integrated financial system<sup>†</sup>.

*Financial Stability and Macprudential Policy (MACPRUD)*. Documents relating to supervising and regulating financial institutions and activities; corporate risk management.

*Fiscal Policy (FIS\_POL)*. Documents on pensions, funds and fiscal issues.

*Innovation (INNOV)*. Documents on technology, digitalization. It may include also survey, reports on data and central banks' research.

*International Affairs (INT\_AFF)*. Documents about international governance and the global economy, international trade.

*Monetary Policy (MON\_POL)*. Documents relating to each central bank's specific monetary policy framework; economic outlook and reports.

*Other Topics (MISCELL)*. Documents on arts, culture and not interpretable topics.

*Payments and Settlements (PAYSECS)*. Documents about coins and banknotes releases, credit cards, payment settlements. More generally, everything involving the payments system and its infrastructure. It includes digital payments when INNOVAT is not clearly clustered.

In Table 2, we show the number of topics clustered under each label for each bank. Ignoring the recipient of irrelevant or intractable results (MISCELL), we find as expected that MON\_POL is one of the most frequent topics, followed by MACPRUD and PAYSECS. Those topics are shared by all institutions, being part of their mandate. For the **Bank of Canada**, we discover 7 labels. After Monetary Policy and Macprudential Policy, International Affairs' weight is in line with the

\*<https://www.federalreserve.gov/aboutthefed.htm>

†<https://www.ecb.europa.eu/home/search/financialintegration/html/index.en.html>

importance of the global trade strategy that Canada has been developing to diversify its economy and improve competitiveness. In fact, Canada heavily relies on trade as it represents more than 60 percent of its output\* - one of the highest ratios among the G7 countries. To sum up, international trade is “the lifeblood of the Canadian economy”† . The **Bank of England** is the one with less topics and the labels MON\_POL and MACPRUD comprehend almost all topics. This result is as expected, since the institutional webpage allow us to explore speeches by filtering only for few topic options. The **European Central Bank** is the target bank whose topics show the highest variety of subjects (9), possibly due to the largest dataset. Again, the MON\_POL and MACPRUD labels include the highest number of topics. The **Federal Reserve** is characterized by 6 labels, where MON\_POL and MACPRUD are almost equivalent in terms of number of topics and CONSCOM is specifically of the Fed, since promoting Consumer Protection and Community Development is one of central bank’s tasks.

**Table 2.** Topics frequency distribution.

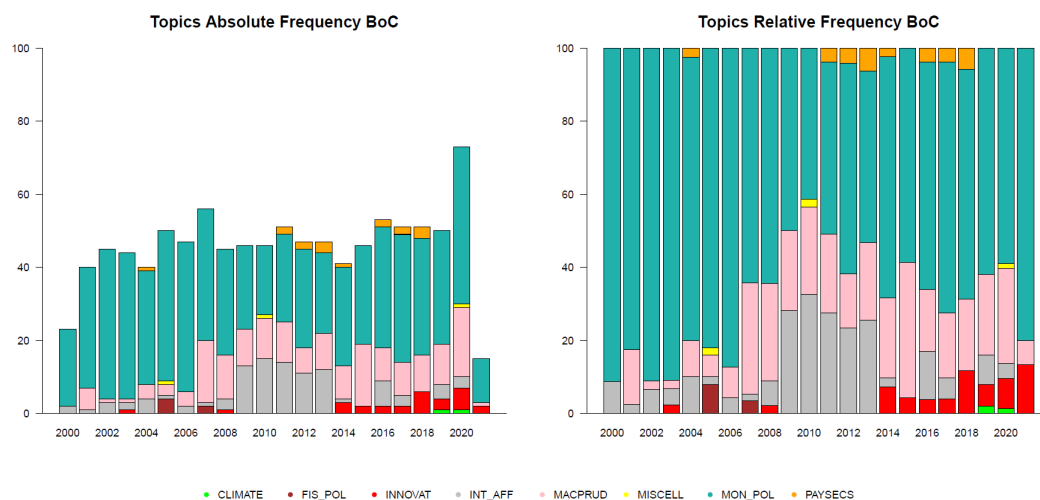
Topics	BoC	BoE	ECB	Fed
CLIMATE	1	-	1	-
CONSCOM	-	-	-	6
EMU_AFF	-	-	6	-
FINTEGR	-	-	1	-
FIS_POL	1	-	2	-
INNOVAT	2	-	3	1
INT_AFF	6	1	3	1
MACPRUD	7	7	10	8
MISCELL	16	16	8	3
MON_POL	18	10	22	9
PAYSECS	4	1	4	2
Num of Topics	55	35	60	30

In Figures 3–6 we show the yearly topic distribution. For the **BoC**, MON\_POL includes 668 speeches and releases, accounting for 65% of the documents. The topic MACPRUD, consisting of 182 documents, weighs on average 18%. The third topic ranked by relative frequency INT\_AFF has 102 documents and on average weighs 10%. For the **BoE**, on average MON\_POL represents 60% of the documents with 788 speeches or releases. The second topic MACPRUD consists of 403 documents and a weight of 30% overall. The third topic ranked by relative frequency PAYSECS has 70 documents. For the **ECB** the main label MON\_POL includes 58% of the documents set with 1345 speeches or releases. The second topic MACPRUD consists of 414 documents and a weight of 18% overall. The third topic ranked by relative frequency EMU\_AFF has 151 documents. FIS\_POL differs from others central banks as its speeches are related to the opportunity of a fiscal union, as a view expressed by the central bank. FINTEGR consists of 140 documents (7% of total) with a focus on the opportunity – as well as the urgency – of a banking union. For the **Fed**, MON\_POL accounts for half the dataset (686 documents), followed by MACPRUD that weighs 30% (446 documents). CONSCOM has 225 speeches in total.

\*Trade is the sum of exports and imports (of goods and services) as a share of gross domestic product. World Bank data, 2019.

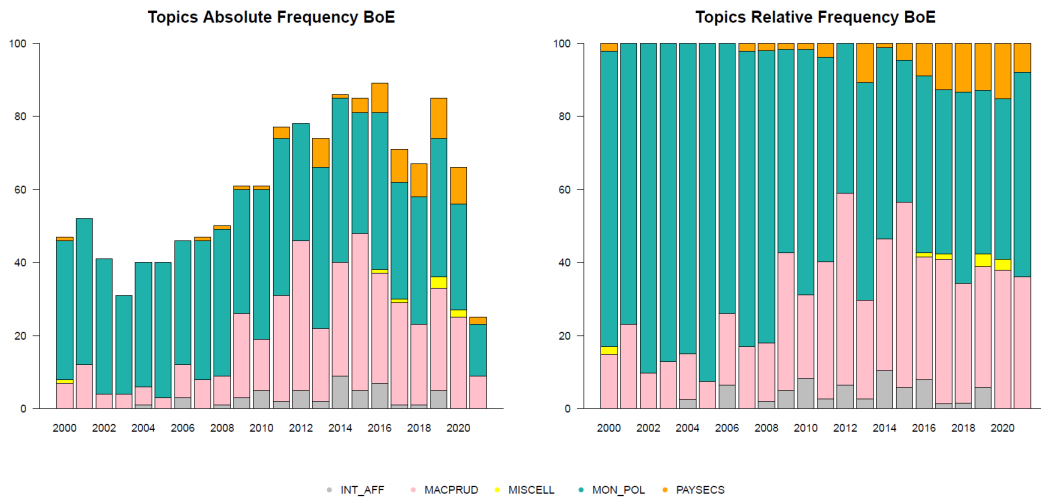
†Remarks by Mr Timothy Lane, Deputy Governor of the Bank of Canada, at the Saskatoon Regional Economic Development Authority, Saskatoon, Saskatchewan, 18 September 2017.

Again, in Figures 3–6 we show the relative yearly frequency and infer the topic trends by looking at the evolution of their frequency over time. Since the subprime crisis (2007–2009) the central banks have given more attention to macroprudential policy issues. In those years, BoC’s focus shifted to macroprudential policy and international affairs. In the last six years, the central bank started talking about digital payments, financial technology and all innovation-related topics. Although the BoE does not share INNOVAT among the identified topics, it focuses on digital payments that are included in PAYSECS segment as shown by the trend of the last past five years. For the ECB, we also notice that PAYSECS and EMU\_AFF were more relevant in 2000–2005 as a consequence of the Euro introduction and new member countries joining the EU. In the following years, central bankers started to talk about fiscal integration (FIS\_POL). Lastly, CLIMAT appears in 2018, following the European Commission Action Plan on Sustainable finance (March 2018). For the Fed, we can see that during the subprime years the focus increases to contents about Consumers and Community, the ones heavily hit by the financial crisis. In fact, we find that financial literacy is a recurrent subject within the documents on this topic. This is because financial distress is not always caused by a lack of financial education, but in many cases such knowledge mitigates the consequences of unaware decisions\*. In the last years, Fed’s Governors touched upon themes such as digital payments and fintech.

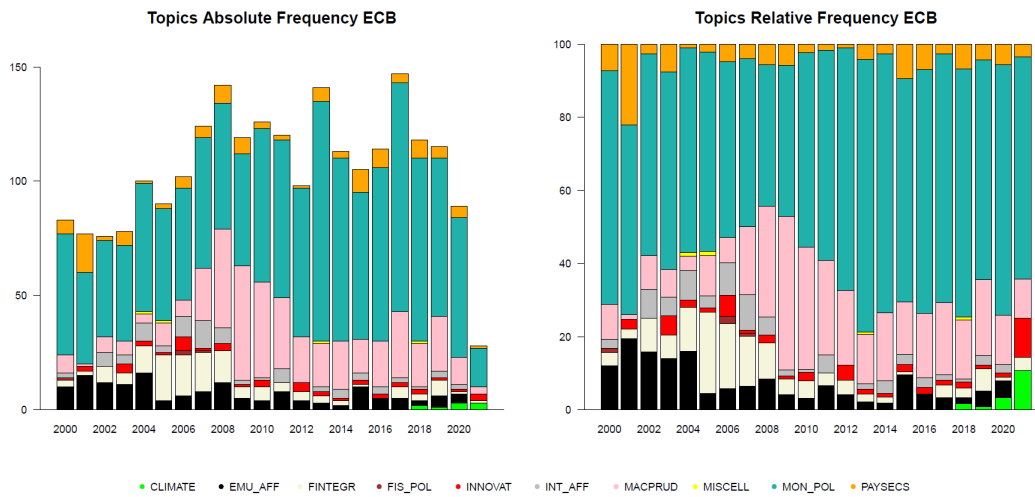


**Figure 3.** BoC Yearly Topic Distribution.

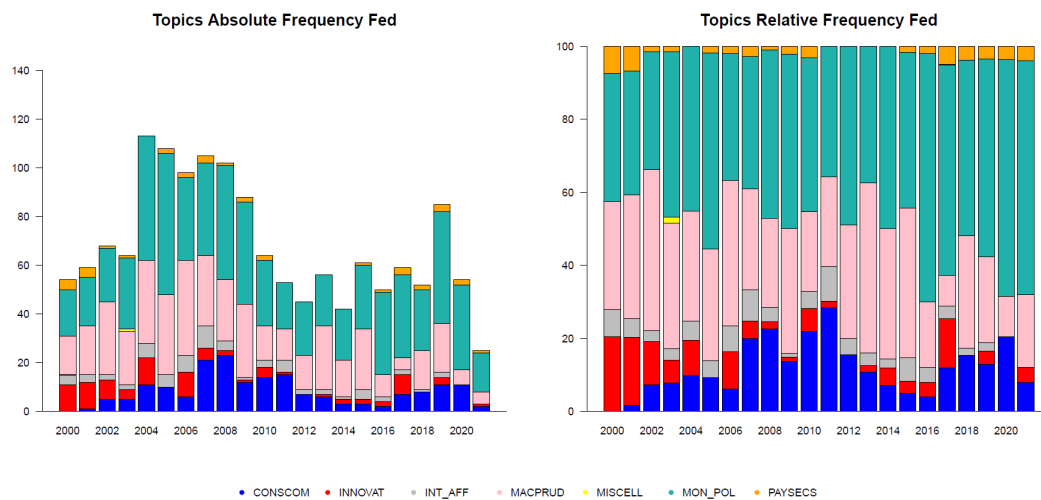
\*Remarks by Alan Greenspan, Chairman of the Board of Governors of the US Federal Reserve System, before the National Council on Economic Education, Chicago, 26 October 2001.



**Figure 4. BoE Yearly Topic Distribution.**



**Figure 5. ECB Yearly Topic Distribution.**

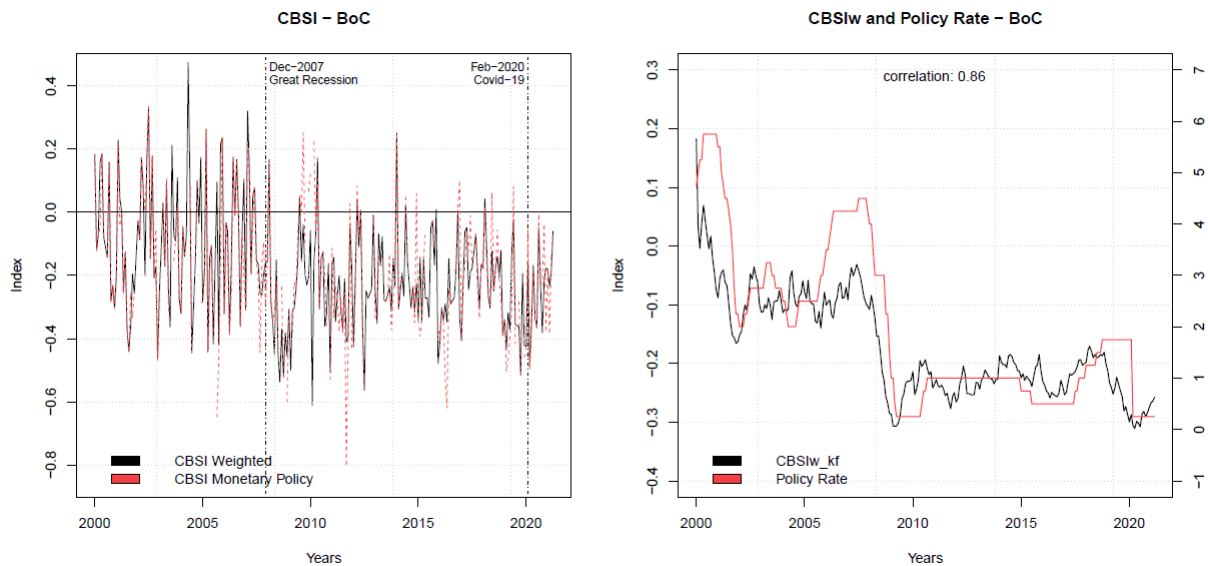


**Figure 6.** Fed Yearly Topic Distribution.

Next, we extract the 10 top words for each K and aggregate the word topic distributions by label. An exhaustive list of top words with assigned probabilities are in Appendix, Tables A5–A8. We provide descriptive statistics of CBSIw and its sub-indicators in Appendix, Tables A1–A4. In Figures 7–10 we show the CBSIw with both global and country-specific events, such as the Great Recession and the beginning of the COVID-19 crisis (dotted lines) and the European Debt Crisis (shaded area). During the years of the subprime crisis (2007–2009) all topic-sentiment indicators moved downwards. To understand better what happened, we plot the weighted CBSI together with the MON\_POL indicator. We see that for all banks the indicator dropped in 2007, when the subprime crisis started. Afterwards, all the CBSIw improved during the period of economic recovery. Furthermore, the ECB indicator goes below zero during the European sovereign debt crisis, in the years from 2010 to 2015. We also find a remarkable drop for all central banks in April 2020 due to the pandemic emergency. In particular, ECB’s sentiment reached its historical minimum in April 2020. These findings are in line with Priola et al. (2020), which build a Central Bank Sentiment Index on the whole dataset of speeches and releases for the BoE, the BoJ, the ECB and the Fed. Next, we look at the correlation between the weighted CBSI and the CBSI sub-indicators, which gives us a measure of which topic is driving the trend of the weighted CBSI. Considering the topics with a consistent number of speeches, we find that MACPRUD, MON\_POL and PAYSECS are the leading topics for the BoC with a Pearson correlation coefficient at least above 0.8; the BoE’s main driver is MON\_POL with 0.84. For the Fed the highest correlation is shown by CONSCOM (0.71). MON\_POL is also the driving cluster for the ECB (0.85). Those statistics are shown in Appendix.

Finally, we investigate the relation between the central bank sentiment, proxied by our weighted CBSI and the relative policy rate. Since CBSIw shows an erratic behavior due to its textual nature, we clean the indicator by extracting the signal with a Kalman filter algorithm (Kalman, 1960), labelled as CBSIw\_kf. In Figures 7–10, we find that the CBSIw\_kf shows evidence of positive correlation for all the policy rates, with a strong correlation for the BoC (0.86), for the ECB (0.76) and a moderate correlation for the BoE (0.52) and the Fed (0.61). We find a misalignment between the two variables for a window of time after the 2008 global financial crisis, where central banks adopted non-standard monetary policies and reached the so-called effective lower bound (ELB). In details, the BoC

determined that the ELB on its key policy rate was 0.25 percent in 2009\* (Witmer and Yang, 2015). The BoE cut the Bank Rate to 0.50 percent in March 2009 and started undertaking a large-scale asset purchases, implying that the current rate was viewed as the effective lower bound for the policy rate (Swanson and Williams, 2014); the Fed kept the federal funds target rate essentially at zero from December 2008 to December 2015, while the ELB has been recently identified between August 2011 and December 2012 (Benamar et al., 2021). For the ECB the break starts in 2014, when the ECB lowered its deposit facility rate below zero while keeping the Main Refinancing Operations (MRO) unchanged. Under the ELB environment the Deposit Facility Rate signalled the monetary policy stance more closely (Cour-Thimann, 2020). This evidence suggests that during the period of quantitative easing key policy rates become less informative, in line with the literature that proposes new indicators to quantify the effects of non-standard policies on the economy (Wu and Xia, 2016).



**Figure 7.** BoC CBSIw and policy rates.

\*In December 2015, the BoC released a new Monetary Policy Framework, estimating that its new ELB could be around -0.50 percent. See the speech given by Poloz (2015) for more details.

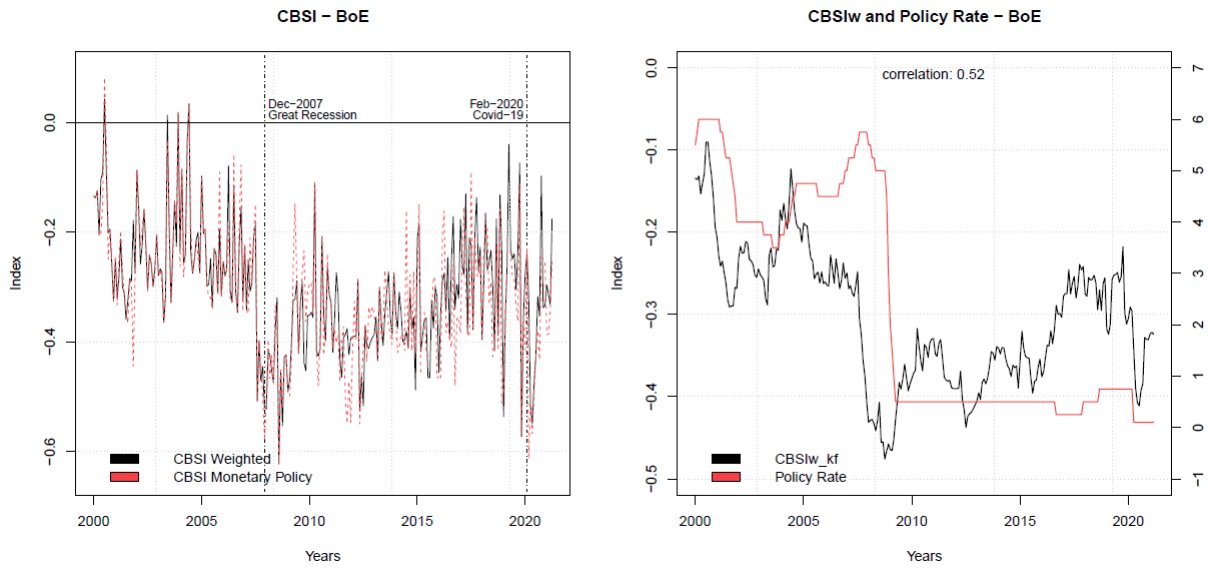


Figure 8. BoE CBSIw and policy rates.

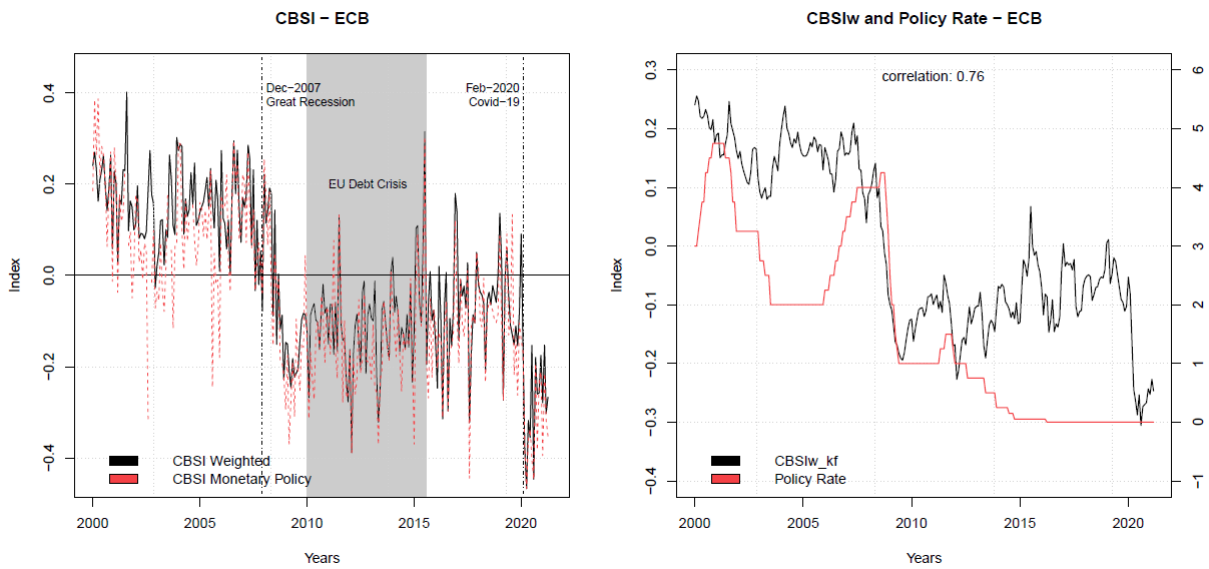
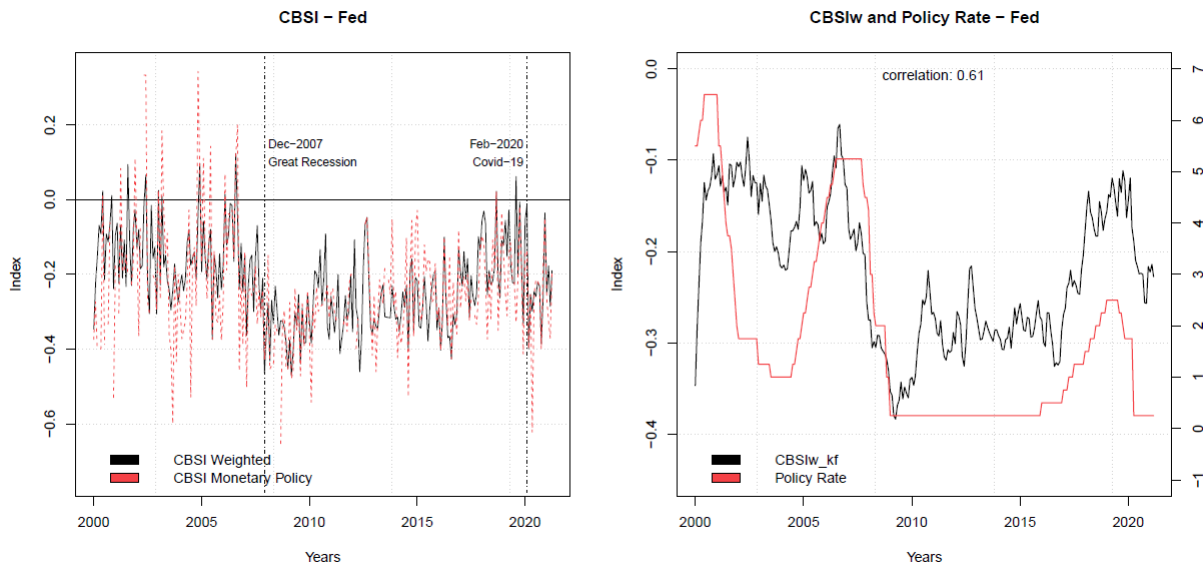


Figure 9. ECB CBSIw and policy rates.





**Figure 10.** Fed CBSIw and policy rates.

## 6. Empirical results

In this section, we assess the quality of information extracted by the LDA algorithm and examine the predictive power of our monetary policy sentiment indicator. We use the following ordered probit model

$$\Delta r_{t+1} = \alpha_1 \Delta r_t + \alpha_2 MON\_POL_t + \epsilon_t \quad (8)$$

Since policy rates typically move by a discrete amount, e.g. by 25 points (Rosa, 2009), we choose the ordered probit method because it takes account of the discrete nature of variables (Hayo and Neuenkirch, 2010). We control for  $\Delta r_{t+1}$  that is the policy rate decision at time  $t+1$ , a categorical variable taking values  $(-1, 0, 1)$ , correspondent to a rate cut, an unchanged rate, a rate hike.  $\Delta r_t$  is the policy rate decision at time  $t$ .  $MON\_POL_t$  is the sub-index related to the monetary policy label, used as a proxy of central bank communication in the baseline specification. Next, we augment the model according to the Taylor rule:

$$\Delta r_{t+1} = \alpha_1 \Delta r_t + \alpha_2 MON\_POL_t + \alpha_3 \Delta GDP_t + \alpha_4 \Delta INF_t + \epsilon_t \quad (9)$$

where  $\Delta GDP_t$  is the GDP growth year-on-year. We interpolate the quarterly GDP via dynamic Chow-Lin developed by Silva and Cardoso (2001) using as leading indicator the industrial production index.  $\Delta INF_t$  is the CPI changes year-on-year. Data is at monthly frequency, covering January 2000 to March 2021. We replace in (2) the sub-index with the  $CBSIw_t$  as a robustness check, given that the  $MON\_POL_t$  sub-index is its major driver, as explained in Section 5. For every estimation we report the McFadden's likelihood ratio index (McFadden, 1974) as a goodness of fit for the ordered probit models, given by:

$$PseudoR = 1 - \frac{LL_{mod}}{LL_0} \quad (10)$$

---

where  $LL_{mod}$  is the log-likelihood value for the fitted model and  $LL_0$  is the log-likelihood value for the baseline model that includes only the intercept as regressor. McFadden's likelihood ratio index ranges from 0 to 1.

Tables 3 shows the results of the estimation. Column I is the baseline model, a regression of the future change in the policy rate on the change in the policy rate in the previous period. Column II is the model with our MON\_POL indicator. The results in column II suggest that in model (1) the monetary policy sentiment index helps predicting the next policy rate decision for all the Banks, with the exception of the Bank of Canada. It worth noting that we do not find a significant result for the BoC specification (II) as it can be affected by collinearity, given the high correlation between the series. Hence, we test the probit specification without the lagged dependent variable (last policy rate decision) finding that the CBSIw is significant for all the banks. In column III, the results obtained controlling for the macroeconomics variables based on the Taylor rule, show that the monetary policy indicator remains informative. The result holds for the Federal Reserve, the European Central Bank and the Bank of England. Next, we run the same regression specification replacing the MON\_POL index with the weighted CBSI. Results are shown in columns V and VI. Our weighted central bank sentiment index is still informative, also after accounting for macroeconomic variables. For the Fed, results are consistent with Hayo and Neuenkirch (2010), who analyse the importance of Federal Reserve communication to predict federal funds target rate decisions for the sample period 1998-2006. For the ECB, results are also in line with Picault and Renault (2017) that show the importance of the ECB communication index for future monetary policy and stock market dynamics.

**Table 3.** Ordered probit estimation results.

variables	(I)	(II)	(III)	(IV)	(V)	(VI)
<b>BANK OF CANADA</b>						
$\Delta r_t$	1.15***, (0.1864)	1.12***, (0.1887)	1.06***, (0.1931)	1.08***, (0.191)	1.1***, (0.1895)	1.05***, (0.1932)
$CBSIw_t$					0.15*, (0.0857)	0.11, (0.0882)
$MON\_POL_t$		0.09, (0.085)	0.05, (0.0873)			
$\Delta GDP_t$		0.28***, (0.1061)	0.29***, (0.1047)		0.27**, (0.1071)	
$\Delta INF_t$		-0.18*, (0.1028)	-0.19*, (0.1014)		-0.18*, (0.102)	
-1/0	-1.41***, (0.1187)	-1.41***, (0.1188)	-1.45***, (0.1227)	-1.45***, (0.1228)	-1.42***, (0.1191)	-1.45***, (0.1228)
0/1	1.43***, (0.1196)	1.44***, (0.1208)	1.47***, (0.1243)	1.47***, (0.1237)	1.45***, (0.1224)	1.48***, (0.1254)
pseudo R <sup>2</sup>	0.1219	0.1254	0.1462	0.145	0.131	0.1497
LLR	-146.49	-145.89	-142.43	-142.63	-144.96	-141.85
<b>BANK OF ENGLAND</b>						
$\Delta r_t$	1.19***, (0.2325)	1.04***, (0.2415)	1.05***, (0.2437)	1.15***, (0.2368)	1.07***, (0.2388)	1.07***, (0.2407)
$CBSIw_t$					0.26***, (0.0979)	0.28**, (0.1089)
$MON\_POL_t$		0.37***, (0.1014)	0.41***, (0.1124)			
$\Delta GDP_t$		-0.03, (0.1008)	0.1, (0.0923)		0.01, (0.0983)	
$\Delta INF_t$		0.09, (0.1034)	-0.02, (0.0957)		0.06, (0.1031)	
-1/0	-1.55***, (0.1288)	-1.63***, (0.1369)	-1.64***, (0.1382)	-1.56***, (0.1292)	-1.59***, (0.1322)	-1.59***, (0.133)
0/1	1.7***, (0.1384)	1.83***, (0.1558)	1.84***, (0.158)	1.71***, (0.1393)	1.77***, (0.1481)	1.78***, (0.1494)
pseudo R <sup>2</sup>	0.106	0.1632	0.1663	0.1104	0.1359	0.1376
LLR	-112.1	-104.93	-104.54	-111.55	-108.35	-108.14
<b>EUROPEAN CENTRAL BANK</b>						
$\Delta r_t$	0.78***, (0.2195)	0.52**, (0.2339)	0.5**, (0.2353)	0.64***, (0.2277)	0.63***, (0.2278)	0.59**, (0.2303)
$CBSIw_t$					0.29***, (0.0956)	0.22*, (0.1168)
$MON\_POL_t$		0.44***, (0.103)	0.42***, (0.1249)			
$\Delta GDP_t$		0.07, (0.11)	0.25***, (0.0929)		0.16, (0.1062)	
$\Delta INF_t$		-0.05, (0.101)	0.03, (0.0955)		-0.02, (0.1003)	
-1/0	-1.44***, (0.1198)	-1.54***, (0.1291)	-1.54***, (0.1293)	-1.48***, (0.1235)	-1.49***, (0.1236)	-1.5***, (0.1247)
0/1	1.55***, (0.1254)	1.71***, (0.146)	1.71***, (0.1452)	1.61***, (0.1314)	1.63***, (0.1352)	1.63***, (0.1353)
pseudo R <sup>2</sup>	0.0475	0.1212	0.1233	0.0788	0.0841	0.0917
LLR	-129.19	-119.19	-118.9	-124.95	-124.23	-123.19
<b>FEDERAL RESERVE</b>						
$\Delta r_t$	1.09***, (0.1856)	1.11***, (0.1873)	1***, (0.193)	0.97***, (0.1913)	1.08***, (0.1868)	0.99***, (0.1922)
$CBSIw_t$					0.2**, (0.0845)	0.15*, (0.0899)
$MON\_POL_t$		0.19**, (0.0838)	0.16*, (0.0858)			
$\Delta GDP_t$		0.25**, (0.0985)	0.28***, (0.0969)		0.25**, (0.0997)	
$\Delta INF_t$		-0.05, (0.0928)	-0.05, (0.0925)		-0.08, (0.0942)	
-1/0	-1.45***, (0.1223)	-1.47***, (0.1237)	-1.52***, (0.1301)	-1.51***, (0.1295)	-1.47***, (0.1232)	-1.52***, (0.13)
0/1	1.36***, (0.1146)	1.4***, (0.1188)	1.43***, (0.1223)	1.41***, (0.1194)	1.4***, (0.1196)	1.43***, (0.1223)
pseudo R <sup>2</sup>	0.11	0.1263	0.1467	0.1367	0.1268	0.1456
LLR	-148.27	-145.55	-142.16	-143.82	-145.46	-142.34

The model is estimated by maximum likelihood. Robust standard errors are in parenthesis. Coefficients labelled with \*\*\*, \*\*, \* denote significance at 1%, 5%, 10% levels, respectively. The number of observations is 255.

Notably, probit coefficients do not offer a direct measure of the contribution of explanatory variables to the dependent one, as it happens in linear models. Thus, we calculate the average marginal effects (hereafter AME) that Gelman and Hill (2006) called “average predicted probability” to assess the magnitude of the effects. We calculate the AME for model (I) and (II) to interpret the results shown in Table 4. For a better interpretation we standardize all continuous variables, therefore results reports the average marginal effects on the probability (of a rate cut, unchanged rate, rate hike) for a one standard deviation increase in the explanatory continuous variable, or for a unit change in discrete explanatory variables. Given the specification in model (I), for the BoC the (not significant) probability of a policy rate hike goes up by 18% after a hike implemented in the last policy rate decision, where the probability

**Table 4.** Average marginal effects

<b>BoC</b>	P(RC)	P(NC)	P(RH)	<b>BoE</b>	P(RC)	P(NC)	P(RH)
	(I)				(I)		
$\Delta r_t$	-0.184***	0.005	0.180***	$\Delta r_t$	-0.160***	0.038	0.121***
	(II)				(II)		
$\Delta r_t$	-0.179***	0.005	0.173***	$\Delta r_t$	-0.134***	0.038	0.096***
<i>MON_POL</i> <sub>t</sub>	-0.015	0.000	0.014	<i>MON_POL</i> <sub>t</sub>	-0.048***	0.014	0.034***
	(III)				(III)		
$\Delta r_t$	-0.164***	0.003	0.161***	$\Delta r_t$	-0.135***	0.039	0.097***
<i>MON_POL</i> <sub>t</sub>	-0.008	0.000	0.008	<i>MON_POL</i> <sub>t</sub>	-0.052***	0.015	0.037***
$\Delta GDP_t$	-0.043***	0.001	0.042**	$\Delta GDP_t$	0.003	-0.001	-0.002
$\Delta INF_t$	0.027*	-0.001	-0.027*	$\Delta INF_t$	-0.011	0.003	0.008
<b>ECB</b>	P(RC)	P(NC)	P(RH)	<b>Fed</b>	P(RC)	P(NC)	P(RH)
	(I)				(I)		
$\Delta r_t$	-0.116***	0.019	0.098***	$\Delta r_t$	-0.161***	-0.027	0.187***
	(II)				(II)		
$\Delta r_t$	-0.073**	0.016	0.058**	$\Delta r_t$	-0.162***	-0.022	0.183***
<i>MON_POL</i> <sub>t</sub>	-0.062***	0.013	0.049***	<i>MON_POL</i> <sub>t</sub>	-0.028**	-0.004	0.032**
	(III)				(III)		
$\Delta r_t$	-0.071**	0.015	0.056**	$\Delta r_t$	-0.138***	-0.025	0.163***
<i>MON_POL</i> <sub>t</sub>	-0.059***	0.012	0.047***	<i>MON_POL</i> <sub>t</sub>	-0.022*	-0.004	0.025*
$\Delta GDP_t$	-0.009	0.002	0.008	$\Delta GDP_t$	-0.035***	-0.006	0.041**
$\Delta INF_t$	0.008	-0.002	-0.006	$\Delta INF_t$	0.007	0.001	-0.008

Coefficients are labelled with \*\*\*, \*\*, \* and denote significance at 1%, 5%, 10% levels, respectively. Probabilities of RC = Rate Cut; NC = No Change, RH = Rate Hike.

of a rate cut increase by 18% after a unit decrease in  $\Delta r_t$ . Given a positive (negative) unit change in last rate decision, a rate hike (cut) is more likely with 12% (16%) of probability for the BoE, 10% (11%) for the ECB, 18% (16%) for the Fed.

In the *MON\_POL* specification (II), we find that the a one-standard deviation increase (decrease) change in monetary policy communication increases on average the probability of a rate hike (rate cut) by 1% for the BoC, 5% (3%) for the BoE, 5% (6%) for the ECB, 3% for the Fed. When adding our communication variable, we find that the marginal contribution of most recent policy rate decision decreases at least for the ECB and the BoE, as partial information is captured by the central banks' monetary policy indicator. Finally, results hold for the specification (III), where controlling for macroeconomic variables results show significant average marginal effects.

As a check, we run the same probit technique with an alternative specification by excluding the period under the Effective Lower Bound (ELB) environment for the BoE, the ECB and the Fed. We cut our samples respectively by March, 2009, by June, 2014 and from August 2011 to December 2012 Benamar et al. (2021). Tables A3–A4 in Appendix show the results of the alternative specification for the probit regression and the AME results, respectively. Results hold when excluding the ELB period, in the model specifications employed. However, we notice that when employing CBSIw as a sentiment indicator, the Taylor-like specification (VI) for the ECB loses significance. Our interpretation is that

since the ECB dataset is the most heterogeneous in terms of topics amount, the MON\_POL index is more likely to capture central bank communication focused on interest rates changes. As expected, we find an increase in the average marginal effects for one-standard deviation increase (decrease) of MON\_POL. Given the model (II), for the BoE the AME doubled from 5% to 10% (rate cut) and from 3% to 7% (rate hike); for the ECB, from 6% to 8% (rate cut) and from 5% to 7% (rate hike), while for the Fed is unchanged (3%) respective of a rate cut and a rate hike. This result might suggest that the adoption of unconventional policies breaks the relationship between key policy rates and the forward guidance given in speeches and releases.

## 7. Conclusions

Central banks communication is strongly associated with a number of critical decisions, therefore the ability to rapidly monitor such information is potentially useful for policymakers and market participants. The literature acknowledges text analysis as a useful technique, as it decodes information embedded in documents into quantitative indicators. However, text analysis does not discriminate among the contexts in which the same words can be used.

In this paper, we develop a combination of machine learning and text analysis techniques to build a topic-weighted Central Bank Sentiment Index for four central banks (BoC, BoE, ECB, Fed). Using a large dataset that includes speeches and press releases, we uncover the latent topics running the LDA algorithm. The main advantage of LDA is that, once the model is trained, it can predict topics for new unseen documents and it provides an intuitive tool for real time analysis. However, the model training requires careful fine-tuning to obtain good results. For this reason, we use a gridsearch approach for the optimal choice of the hyperparameters.

After labelling the dataset according to the identified topics for each bank, we build a topic sentiment index for each bank, according to the methodology proposed in Priola et al. (2021). As a last step, we aggregate the sentiment indexes in a topic weighted Central Bank Sentiment Index. We show that our CBSIw for the four Central Banks share common trends, decreasing during the Great Recession and during the COVID-19 crisis; moreover, we find a strong positive correlation between the indicator and the policy rates. Amongst the sub-indicators that compose the CBSIw, we find that the major driver is monetary policy, followed by macroprudential policy and financial stability. Moreover, all banks draw the attention to common subjects such as digitalization and fintech, but there are also bank-specific topics. For the BoC, International Affairs is the third leading subject that supports the evidence of the increasing importance of global competitiveness after the subprime crisis in 2007. For the BoE, digitalization and fintech innovation for payments is important, as shown by the increasing numbers of related speeches in the last years of the sample. For the first ten years in the sample, the ECB focused on issues related to the development of the European Monetary Union, proxied by European Affairs, and stressed the importance of financial integration. We also find an increasing frequency of climate change and digitalization contents, one of the new ECB commitments stated by Christine Lagarde during the Opening Statement to the Economic and Monetary Affairs Committee, in September 2019. For the Fed, among other topics, we find that the central bank promotes the protection of consumers by publishing speeches about the importance of financial education and community development.

Next, we investigate the predictiveness power of the sentiment indicators by employing a probit model. We test the monetary policy sub-index contents as well as and the weighted index. We find

that coefficients are positive and significant, suggesting that the MON\_POL index helps to predict future changes in policy rates, proxied by a ternary variable taking values of -1,0,1 (respectively a rate cut, unchanged rate, rate hike). The results hold when adding macroeconomic variables with a Taylor rule-like framework. According to the average marginal effects, we find that a rate hike is more likely to happen with a one standard deviation change in monetary policy communication, as well as a rate cut with a downward change of the same index. We add a preliminary investigation on the effective lower bound, with results that suggest that the MON\_POL index increases its explanatory power when excluding the ELB period.

We believe that further research may be relevant. First, the importance of unconventional monetary policies might have caused structural changes in the relationship between central banks' communication and monetary policy decisions. Therefore, an effort to identify this potential break is required with a further investigation on new monetary policy stance indicators that better proxy forward guidance during the period of unconventional monetary policy. Second, with the growing importance of social media for political and financial communication, there is an emerging field of analysis that links central banking news and the sentiment reaction of market participants that might be worth exploring.

### Conflict of interest

All authors declare no conflicts of interest in this paper. Any errors are of the authors. The views expressed in this paper are those of the authors and do not necessarily represent the opinions of Prometeia S.p.A..

### References

- Apel M, Grimaldi MB (2014) How informative are central bank minutes? *Review Econ* 65: 53–76. <https://doi.org/10.1515/roe-2014-0104>
- Armeliu H, Bertsch C, Hull I, et al. (2020) Spread the word: International spillovers from central bank communication. *J Int Money Financ* 103: 102116. <https://doi.org/10.1016/j.jimonfin.2019.102116>
- Baeza-Yates R, Ribeiro-Neto B (1999) *Modern information retrieval*, 463. ACM press New York.
- Benamar H, Foucault T, Vega C (2021) Demand for information, uncertainty, and the response of us treasury securities to news. *Rev Financ Stud* 347: 3403–3455. <https://doi.org/10.1093/rfs/hhaa072>
- Bennani H, Fanta N, Gertler P, et al. (2020) Does central bank communication signal future monetary policy in a (post)-crisis era? the case of the ecb. *J Int Money Financ* 104: 102167. <https://doi.org/10.1016/j.jimonfin.2020.102167>
- Birz G, Lott Jr JR (2011) The effect of macroeconomic news on stock returns: New evidence from newspaper coverage. *J Bank Finance* 35: 2791–2800. <https://doi.org/10.1016/j.jbankfin.2011.03.006>
- Blei DM, Ng AY, Jordan MI (2003) Latent dirichlet allocation. *J Mach Learn Res* 3: 993–1022.
- Carboni M, Farina V, Previati DA (2020) Ecb and fed governors' speeches: A topic modeling analysis (2007–2019). In *Banking and Beyond*, Springer, 9–25.

- Correa R, Garud K, Londono JM, et al. (2017) Constructing a dictionary for financial stability. *Board Governors Fed Reserve System, Washington, DC*. <https://doi.org/10.17016/2573-2129.33>
- Cour-Thimann P (2020) Interest rate setting and communication at the ecb. <http://dx.doi.org/10.2139/ssrn.3654147>
- Dincer NN, Eichengreen B (2013) Central bank transparency and independence: updates and new measures. <http://dx.doi.org/10.2139/ssrn.2579544>
- Gelman A, Hill J (2006) *Data analysis using regression and multilevel/hierarchical models*. Cambridge university press.
- Griffiths TL, Steyvers M (2004) Finding scientific topics. *P NATL A SCI INDIA B* 101: 5228–5235. <https://doi.org/10.1073/pnas.030775210>
- Grün B, Hornik K (2011) topicmodels: An r package for fitting topic models. *J Stat Software* 40: 1–30. <http://www.jstatsoft.org/v40/i13>
- Hansen S, McMahon M, Prat A (2018) Transparency and deliberation within the fomc: a computational linguistics approach. *Q J Econ* 133: 801–870. <https://doi.org/10.1093/qje/qjx045>
- Hayo B, Neuenkirch M (2010) Do federal reserve communications help predict federal funds target rate decisions? *J Macroecon* 32: 1014–1024. <https://doi.org/10.1016/j.jmacro.2010.06.003>
- Kalman RE (1960). A new approach to linear filtering and prediction problems” transaction of the asme journal of basic. <https://doi.org/10.1115/1.3662552>
- Keida M, Takeda Y (2018) The arts of central bank communication: A topic analysis on words of the bank of japan s governors. AEA Annual Conference.
- Leskovec J, Rajaraman A, Ullman JD (2020). *Mining of massive data sets*. Cambridge university press.
- Logsdon BA, Hoffman GE, Mezey JG(2010) A variational bayes algorithm for fast and accurate multiple locus genome-wide association analysis. *BMC Bioinformatics* 11: 1–13. <https://doi.org/10.1186/1471-2105-11-58>
- Loughran T, McDonald B (2011). When is a liability not a liability? textual analysis, dictionaries, and 10-ks. *J Financ* 66: 35–65. <https://doi.org/10.1111/j.1540-6261.2010.01625.x>
- McFadden D (1974) Conditional logit analysis of qualitative choice behavior. *Front Economet* 50: 105–142.
- Navarro DJ, Griffiths TL, Steyvers M, et al. (2006) Modeling individual differences using dirichlet processes. *J Math Psychol* 50: 101–122. <https://doi.org/10.1016/j.jmp.2005.11.006>
- Picault M, Renault T (2017) Words are not all created equal: A new measure of ecb communication. *J Int Money Financ* 79: 136–156. <https://doi.org/10.1016/j.jimonfin.2017.09.005>
- Priola MP, Lorenzini P, Tizzanini G, et al. (2021) Measuring central banks’ sentiment and its spillover effects with a network approach. <http://dx.doi.org/10.2139/ssrn.3764004>

- Röder M, Both A, Hinneburg A (2015) Exploring the space of topic coherence measures. In *Proceedings of the eighth ACM international conference on Web search and data mining* 399–408. <https://doi.org/10.1145/2684822.2685324>
- Rosa C (2009) Forecasting the direction of policy rate changes: The importance of ecb words. *Econ Notes* 38, 39–66. <https://doi.org/10.1111/j.1468-0300.2009.00209.x>
- Shapiro AH, Sudhof M, Wilson DJ (2020). Measuring news sentiment. *J Econom.* <https://doi.org/10.1016/j.jeconom.2020.07.053>
- Silge J, Robinson D (2017) *Text mining with R: A tidy approach.* ” O’Reilly Media, Inc.”.
- Silva JS, Cardoso F (2001) The chow-lin method using dynamic models. *Econ model* 18: 269–280. [https://doi.org/10.1016/S0264-9993\(00\)00039-0](https://doi.org/10.1016/S0264-9993(00)00039-0)
- Swanson ET, Williams JC (2014) Measuring the effect of the zero lower bound on medium-and longer-term interest rates. *Am Econ Rev* 104: 3154–3185. <https://doi.org/10.1257/aer.104.10.3154>
- Tumala MM, Omotosho BS (2019) A text mining analysis of central bank monetary policy communication in nigeria. *CBN J Appl Stat* 10. <http://dx.doi.org/10.2139/ssrn.3545508>
- Witmer J, Yang J (2015) Estimating canada’s effective lower bound. Technical report, Bank of Canada.
- Wu JC, Xia FD (2016). Measuring the macroeconomic impact of monetary policy at the zero lower bound. *J Money, Credit Banking* 48: 253–291. <https://doi.org/10.1111/jmcb.12300>
- Zipf G (1936) *The psycho-biology of language*, routledge, london. <https://doi.org/10.2307/408930>



AIMS Press

©2022 the Author(s), licensee AIMS Press. This is an open access article distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/4.0>)