What do press conferences tell us about central banker's sentiment?

Renaud BEAUPAIN[1] and Minh Ly LIEU[1]

[1]IESEG School of Management, Univ. Lille, CNRS, UMR 9221 - LEM - Lille Economie Management, F-59000 Lille, France

July 25, 2024

Abstract

This research study contributes to understanding the tone in press conferences held by the European Central Bank (ECB). The study applies the Large Language model for sentiment analysis, specifically focusing on the financial context using the finBERT model and most popular lexicon-based methods in finance, the Loughan and McDonald (LM) model. This study is the first to compare these two models in sentiment analysis for the central bank context, specifically for the ECB. Additionally, the results suggest several interesting conclusions. Firstly, the tone in the introductory part of the conferences is related to macro events such as crises, the COVID-19, and the Ukraine war, while the Q&A portion is connected to both shocks and presidential periods. Secondly, there are different communication styles observed in the speeches of the ECB's presidents. These results contribute to our understanding of the tone in the ECB's communications and encourage further research in this area.

Keywords: Central bank communication, Sentiment Analysis, Lexicon-based approach, Transformer models.

1 Introduction

With its function of protecting the stability of the monetary system, inflation, and currency, the decisions and announcements of the central bank always raise attention and serve as guideline actions for both experts and markets. Furthermore, the announcements of central banks in major countries have an influence not only within those countries but also spread widely to many other countries. For instance, the announcements of the European central banks can lead to significant movement worldwide. Among the various communication channels utilized by the ECB, press conferences hold a significant place as they provide direct insights into the central bank's policy decisions, economic outlook, and strategic considerations.

Understanding the sentiment conveyed in these press conferences is essential for several reasons. Firstly, sentiment analysis can reveal the underlying tone and emotional context of the ECB's communication, which can influence market behavior and public perception. Secondly, it provides a quantitative measure to analyze the ECB's communication strategy over time, identifying trends, shifts in policy emphasis, and responses to economic challenges. Thirdly, by comparing sentiment across different time periods and economic contexts, researchers can assess the consistency and effectiveness of the ECB's messaging.

In this study, we contribute to the understanding of the ECB's press conference sentiment with several novel points. First, we investigate both the introductory and Q&A parts of the press conferences - parts that have not received deserved attention in the literature. Second, we apply the finBERT model, a large language model that allows us to analyze the sentiment in the financial context. Third, we compare the finBERT model with the lexicon model (Loughan and McDonald model - LM model) to investigate the diversity in sentiment analysis methods in central bank communication. Finally, we answer the question of whether the change in the president of the ECB could impact the sentiment of their press conferences and how they could change over the presidencies. Specifically, we find that the structure of the introductory sections of press conferences aligns with shocks and crises, including the global crisis, COVID-19, and the Ukraine war. On the other hand, the tone of the Q&A parts is closely connected with the personality styles of the ECB's presidents. Interestingly, in addition to the change in communication style by presidents, we find a persistent decline in the fluctuation of sentiment indicators in the Q&A parts. This phenomenon could be explained by the adoption and adjustments made by presidents to stabilize the market.

The subsequent sections of this paper are structured as follows: Section 2 provides essential background discussions, Section 3 summarizes the methods applied and outlines the data structure in this study. Empirical results are presented in Section 4, and Section 5 engages in a comprehensive conclusion.

2 Literature review

The topic of central bank communications encompasses various directions and types of communication. Theoretically, this study classifies the literature into two main branches: (i) the impact of communications on other indices or markets, and (ii) the structure and characteristics of communications.

Economic research has demonstrated the wide-ranging influence of central bank communications on markets and indices such as the stock market Gorodnichenko et al. (2023), bond market (Ehrmann & Talmi, 2020), future policy rates (Niţoi et al., 2023), investor sentiment (Bennani, 2020), return volatility (Apergis & Pragidis, 2019), and inflation expectations (Bennani & Neuenkirch, 2017). Generally, central bank communications play a crucial role in steering the economic system and markets. Specifically, positive signals from central bank communications tend to lead markets in a positive direction, while uncertainty or inconsistency in communication often results in market volatility.

In the second branch, which focuses on the structure and characteristics of communications, various approaches are used, such as readability (Ferrara & Angino, 2022), consistency in topics (Klejdysz & Lumsdaine, 2023), consistency in lexicon (Rosa & Verga, 2007), field-specific weighted lexicon (Picault & Renault, 2017), sentiment variance (Hayo & Zahner, 2023), and incoherence (Jansen, 2011). This branch investigates the structure and patterns in central bank communications and also yields very interesting results. In particular, readability and coherence in central bank statements not only create media engagement (Ferrara & Angino, 2022) but also lead to market reactions (Rosa & Verga, 2007; Hayo & Zahner, 2023).

While the literature on central bank communications covers various types, this paper focuses specifically on the European Central Bank (ECB) press conferences. The ECB's press conferences provide significant additional information to financial markets beyond what is included in the monetary policy decisions. The value of this information is strongly connected to the nature of the decisions themselves (Ehrmann & Fratzscher, 2009), making press conferences one of the most influential communications of the ECB. Studies have also identified a link between the ECB's press conferences and various financial indices. This connection is stronger compared to the other ECB's communications, such as meeting accounts, Executive Board speeches (Kaminskas & Jurkšas, 2023), and inter-meeting speech communications (Kanelis & Siklos, 2024).

From a technical perspective, the literature on the impact of press conferences can be divided into two periods: (1) before the development of textual analysis techniques, and (2) after the development of textual and sentiment analysis techniques.

Before the advent of textual and sentiment analysis techniques, the literature on ECB statements primarily focused on announcement times and the market reactions around these timestamps (Ehrmann & Fratzscher, 2009; Brand et al., 2010). Specifically, Ehrmann & Fratzscher (2009) analyzed the reaction of 3-month Euribor futures to ECB press conferences from 2001 to 2006. This study compared market developments on days with press conferences to those on days without. Similarly, Brand et al. (2010) focused on the timing of ECB announcements and their impact on the euro area money market yield curve. Notably, Brand et al. (2010) argued that monetary policy news from the ECB had a more substantial and longer-lasting impact on the euro area yield curve than the interest rate policy decisions themselves. This finding was later corroborated by Leombroni et al. (2021) concerning long-term interest rates.

Beyond examining the timing of communications, scholars have increasingly focused on the content of ECB communications. For example, Rosa & Verga (2007) used hand-coding to classify statements on a scale from hawkish to dovish, finding that market expectations respond to unexpected elements in ECB information, even after accounting for monetary policy shocks. Recently, understanding of ECB communication has advanced with the support of textual analysis tools. Instead of focusing solely on the timing of announcements, recent studies delve into the content and sentiment of ECB communications. Various techniques are applied to extract and measure the sentiment or tone of communications, including hand-coding (Rosa & Verga, 2007; Apergis & Pragidis, 2019; Hayo & Neuenkirch, 2013), lexicon-based methods (Schmeling & Wagner, 2016; Picault & Renault, 2017; Anastasiou & Katsafados, 2023), and large language models (Niţoi et al., 2023).

For instance, Schmeling & Wagner (2016) analyzed the transcripts of the ECB president's opening statements to determine the tone of the communications using a lexicon-based method, specifically the Loughran-McDonald sentiment dictionary (Loughran & McDonald, 2011), and compared these findings with various market indices. The study found that the tone of the ECB's statements affects stock returns, volatility risk premia, policy rates, upward revisions of real GDP growth, recent higher stock market returns, and government bond yields. Then, Klejdysz & Lumsdaine (2023) applied Latent Dirichlet Allocation (LDA) to analyze the ECB's press conferences. Based on topic modeling, they found that while the introductory statements concentrate on a consistent topic and reflect changes in the monetary policy regime, the Q&A sections contain diverse content. They also investigated the correlation between tone and topics. Kanelis & Siklos (2024) analyzed the introductory statements of the ECB's press conferences using finBERT, a large language model, to derive new sentiment indicators. By focusing on monetary policy and financial stability topics through Structural Topic Modeling, they found a significant positive relationship between the average sentiment of inter-meeting monetary policy-related speeches and the sentiment expressed in the introductory statements, but not in the financial stability topic.

Notably, ECB press conferences have a specific structure comprising two parts: (1) introductory statements and (2) questions and answers. While most current studies focus on the introductory statements (Rosa & Verga, 2007; Dybowski & Kempa, 2020; Kanelis &

Siklos, 2024), some emphasize the role and complexity of the Q&A sections (Ehrmann & Fratzscher, 2009; Klejdysz & Lumsdaine, 2023). In this study, we consider both the introductory statements and the Q&A sections so that we can observe the evolution of the introductory statements and the Q&A parts and contribute more understanding of the press conferences.

3 Data and Methodology

The time series in this studies were obtained from the sentiment of the European Central Bank Press conferences by two different approaches the finBERT and the LM model from the first date announcement (9th June 1998) till June 2024. The ECB organized the press conferences on a monthly format started from the 9th June 1998. Then from 2015, the frequency of the meetings was changed to a six-week format. Thus, we collect and analyze totally 280 press conferences of ECB's press conferences. The press conferences have a specific structure compared to the other communications with rich contents and structures. In general, the press conference includes (1) the monetary policy decisions (2) economic analysis, and (3) question and answer. This structure of press conferences provides an overview of policies and markets. Also, at the same time press conferences offer an interaction environment with the participants. Thus, press conferences are still considered the most influential communication in all of the official communication of ECB.

3.1 Data Processing

Our sentiment analysis method consists of four steps: (1) text extracting, (2) text preprocessing, (3) sentiment information extraction and classification, and (4) sentiment indicators calculation. A detailed description of each step is provided below.

(1) Text Extracting

To extract the communications from the European Central Bank, we based on the media part on the website of the European Central Bank. On the offical website of the European Central Bank, there are six main types of communications including press conferences, press releases, monetary policy decisions, the ECB blog, speeches, and interviews. In this study, we collect and extract all communications from the first-day announcement of that communication to September 2023.

(2) Text Preprocessing

To analyse the sentiment of a text, text pre-processing is necessary. In general, text is a combination of many paragraphs with specific structure; thus, we need to tokenize the text into the components of paragraphs. In this study, each communication is a text including many paragraphs. We store these texts by published date in our database. Then we tokenize these texts into paragraph units based on some symbols such as ".", "!", "?". We use paragraph-unit as the mean unit for the whole of our study. By storing paragraph-unit data, we decomposed unstructured text into multiple structured parts while keeping the meaning of texts in paragraphs. In compliance with the finBERT model's tokenization requirement, the maximum token limit for the finBERT model is set at 512 tokens. Therefore, during the tokenization process, we carefully count the number of tokens and ensure that each paragraph contains fewer than 400 tokens, up to the most recent sentence-ending symbols, which include ".", "!", and "?". We have selected this 400-token limit to accommodate situations where sentences are exceptionally long. This ensures that even after any necessary truncation, the content remains within the 512-token limit of the finBERT model. This technique applies to all communications.

In general, the dataset includes 280 conferences considered with up to 1,834,439 tokens [1] and 18,507 paragraphs. In other words, there are approximately 6552 words per talk, each talk normally has around 66 paragraphs with a length as 99 tokens per paragraph. These numbers show the structure of press conferences as one of the longest and most well-structured talks including long paragraphs and abundant information content.

Sentiment analysis is a natural language processing (NLP) technique used to determine the emotional tone, attitude, or sentiment expressed in text data. It involves analyzing written or spoken language to discern whether the text conveys a positive, negative, or neutral sentiment, as well as the intensity of that sentiment. In this study, we introduce and apply two most popular methods to extract sentiment analysis including the Lexicon-based method with Loughan & Mcdonald (LM dictionary) (Loughran & McDonald, 2011) and the Large language model (LLM) with finBERT (Huang et al., 2023).

3.2 Sentiment analysis

Lexicon based method The lexicon-based method generally employs lexicons or dictionaries, consisting of lists of words and phrases associated with various emotions, to classify words (e.g., positive, negative, or neutral) and identify sentiment. There are various dictionaries such as LM dictionary (Loughan & Mcdonald dictionary), AFINN lexicon, SentiWordNet, and VADER (Valence Aware Dictionary for sEntiment Reasoning), etc. Among those dictionaries, the LM dictionary is the most widely utilized lexicon in the finance domain (Leippold, 2023). Since this study focuses on the ECB's communications closely related to the financial context, we apply the LM model to our sentiment analysis.

Notably, keyword-based methods are still susceptible to sentiment adversarial attacks despite the advanced capabilities of the well-known language model. This technique can alter the sentiment of sentences without significantly changing their semantic meaning, a method often employed in human communication (Leippold, 2023). Thus, besides using a lexicon-based model, we also apply a large language model in this study.

[1] Token is the output of the tokenization procedure that splits texts into smaller units. Token can be either words, characters, or subwords.

Large Language Models - FinBERT Common Large Language Models (LLMs) based on the transformer architecture include BERT and GPT models. Even though both leverage the transformer architecture, they are designed with different purposes and function differently. BERT, short for Bidirectional Encoder Representations from Transformers, is designed to learn contextual representations of input sequences by considering both the left and right context. Thanks to its bidirectional approach, BERT excels at tasks that require a deep understanding of context such as Named Entity Recognition (NER) and Question Answering (QA).

To extract the sentiment, we apply finBERT (Huang et al., 2023). The finBERT is a large languages model (LLM) in financial context based on the BERT algorithm. There are three reasons for choosing finBert. First, the central bank communications are related to financial context such as inflation, monetary policy, economic prospects. Thus, the BERT for the general context could lead to misclassify the sentiment. Indeed, Huang et al. (2023) indicate that finance vocabulary helps FinBERT retain its performance (i.e., reduce its accuracy deterioration) when the training sample becomes smaller. Second, finBERT's performance outperforms both the Loughran-McDonald (LM) dictionary and other machine learning models (NB, SVM, RF, CNN, and LSTM) (Huang et al., 2023). Third, FinBERT's could detect the negative sentiment more accurately (Huang et al., 2023) than non-BERT models.

Sentiment indicators calculation After the paragraphs are classified as positive, negative, and neutral. We also calculated the polarity, subjectivity, negativity, and positivity of each text as following formula.

$$Polarity = \frac{N_{\text{Positive}} - N_{\text{Negative}}}{N_{\text{Positive}} + N_{\text{Negative}}} \tag{1}$$

$$Subjectivity = \frac{N_{\text{Positive}} + N_{\text{Negative}}}{N_{\text{Total}}} \tag{2}$$

$$Negativity = \frac{N_{\text{Negative}}}{N_{\text{Total}}} \tag{3}$$

$$Positivity = \frac{N_{\text{Positive}}}{N_{\text{Total}}} \tag{4}$$

Where:

- N_{Positive} is the number of positive paragraphs or sentiments in the text.
- N_{Negative} is the number of negative paragraphs or sentiments in the text.
- N_{Total} is the total number of paragraphs or sentiments in the text.

The polarity value ranges from -1 (indicating negative tones) to 1 (indicating positive tones). The subjectivity value ranges from 0 (representing a neutral position) to 1 (representing an emotional position). Additionally, the negativity (positivity) scale ranges from 0 to 1, with 1 denoting completely negative (positive) text.

In the literature review, there is a wide range of studies using polarity to summarize the sentiment result e.g (Gorodnichenko et al., 2023; Hayo & Zahner, 2023; Mullings, 2023; Apergis & Pragidis, 2019; Hubert & Fabien, 2017; Picault et al., 2022). Whereas, Correa et al. (2021) used a slightly different numerator as negative words - positive words. Gorodnichenko et al. (2023) also applied the polarity for dovish and hawkish tones. Whereas, Bennani (2020) uses a similar definition as polarity but for two groups "confident" and "cautious" terms. In this study, we apply four indicators including polarity, subjectivity, negativity, and positivity.

3.3 Kalman Filter

The Kalman filter estimates a process through feedback control. It predicts the process state at a specific time and then adjusts it based on incoming (noisy) measurements. This process involves two types of equations: time update equations and measurement update equations. Time update equations predict the next state and error estimates in advance (a priori) for the next time step, acting as predictor equations. Measurement update equations refine these predictions based on new measurements, functioning as corrector equations (Welch et al., 1995). Thus, the overall estimation method operates like a predictor-corrector algorithm used for numerical problem-solving.

The Kalman filter provides two outputs: filtering and smoothing. Filtering refers to the general process of extracting valuable information from a noisy signal. Smoothing, a specific type of filtering known as a "low-pass filter," passes low-frequency components while reducing high-frequency components. In some cases, filtering high-frequency data to predict future states by extracting relevant information from a noisy signal is the best use of the Kalman filter. Conversely, smoothing often relies more on past data, as averaging recent measurements can sometimes yield more accurate results than using only the latest measurement. Thus, in this study, we combine both output to capture the trend and the smoothness of the series.

Notably, in the Figure 1 we shows the plot between the raw series and the Kalman filter series. Then the filter series are presented in all of our analyses later.

4 Results

4.1 Summary statistics

Table 1 presents the summary statistics of sentiment outputs from two models, finBERT and the LM model, across three sections: (1) full talk, (2) introductory part, and (3) Q&A part. While there were 280 press conferences from June 1998 to June 2024, some of the earliest press conferences did not include Q&A sessions. Therefore, the number of Q&A observations is limited to 268.Consequently, the number of Q&A parts are smaller than the number of the whole texts and introductory parts.

Overall, finBERT, with its specialized capability to determine sentiment in a financial context, likely identifies more sentiment labels than the LM model, which is based on the financial-specific dictionary of Loughran and McDonald, across all indicators (subjectivity, polarity, positivity, and negativity) for the full talk, the introductory part, and the Q&A section.

The LM model indicates negative polarity for both the full text and the Q&A part. While the polarity of the full text differs between the models (positive in the finBERT and negative in the LM model), both models exhibit negative polarity in the Q&A part. Thus, although the ECB maintains a positive sentiment in the introductory sections, the Q&A sections reflect a more negative sentiment. Additionally, considering the standard deviation, polarity is the most varied indicator, regardless of the model and the section of the press conferences.

Comparing the consistency between the two models, finBERT and the LM model, they produce similar outputs with small differences in the Q&A part (subjectivity 0.105 0.153, polarity $-0.359\ -0.205$, positivity 0.033 0.061, negativity 0.072 0.092). However, there are significant differences between the two models in both the full talk and the introductory part. We argue that introductory parts are often prepared in advance and utilize professional lexicon to maintain neutral statements, making it difficult to detect tone without context. Conversely, the Q&A parts depend on the situation and are challenging to prepare in advance, thus preserving the structure of natural language, which can be effectively captured by both models.

Table 1: Summary statistics

		Mean	St. Dev.	Min	Pctl(25)	Median	Pctl(75)	Max	N
Full talk lexicon model	Subjectivity	0.099	0.015	0.030	0.089	0.099	0.108	0.141	280
	Polarity	−0.256	0.150	−0.600	−0.355	−0.274	−0.177	0.562	280
	Positivity	0.037	0.009	0.013	0.030	0.036	0.042	0.065	280
	Negativity	0.062	0.012	0.017	0.054	0.062	0.070	0.100	280
Full talk finBERT model	Subjectivity	0.274	0.082	0.000	0.222	0.268	0.323	0.615	280
	Polarity	0.140	0.366	−1.000	−0.093	0.122	0.400	1.000	280
	Positivity	0.155	0.071	0.000	0.113	0.145	0.188	0.615	280
	Negativity	0.118	0.061	0.000	0.077	0.113	0.158	0.345	280
Introductory part lexicon model	Subjectivity	0.086	0.019	0.030	0.073	0.085	0.097	0.135	280
	Polarity	0.014	0.224	−0.613	−0.132	0.019	0.154	0.600	280
	Positivity	0.043	0.013	0.013	0.034	0.044	0.052	0.077	280
	Negativity	0.042	0.014	0.012	0.032	0.041	0.051	0.090	280
Introductory part finBERT model	Subjectivity	0.655	0.197	0.000	0.553	0.714	0.789	0.944	280
	Polarity	0.353	0.411	−0.667	0.077	0.385	0.667	1.000	280
	Positivity	0.448	0.198	0.000	0.300	0.462	0.600	0.882	280
	Negativity	0.207	0.137	0.000	0.091	0.200	0.300	0.600	280
Q&A part lexicon model	Subjectivity	0.105	0.017	0.063	0.093	0.104	0.115	0.160	268
	Polarity	−0.359	0.140	−0.628	−0.463	−0.363	−0.271	0.107	268
	Positivity	0.033	0.008	0.014	0.028	0.033	0.038	0.064	268
	Negativity	0.072	0.015	0.040	0.061	0.071	0.081	0.123	268
Q&A part finBERT model	Subjectivity	0.153	0.071	0.000	0.097	0.149	0.203	0.500	268
	Polarity	−0.205	0.469	−1.000	−0.500	−0.261	0.091	1.000	268
	Positivity	0.061	0.044	0.000	0.030	0.056	0.083	0.228	268
	Negativity	0.092	0.057	0.000	0.054	0.089	0.121	0.438	268

Notes:.

4.2 Stationarity testing results

The stationarity results are shown in Table 2 and 3. Particularly, Table 2 shows the Dickey-Fuller test for the full talk section, while Table 3 shows the test for the introductory part and Q&A part. In general, the time series are stationarity in most cases except for the non-zero mean and the trend case in the LM model and the non-zero mean of subjectivity and positivity of the finBERT model.

Table 2: Dickey-Fuller Test Results of Full conferences

Full Conferences		**FinBERT**		**Lexicon**	
Variable	**ADF Type**	**ADF**	**p value**	**ADF**	**p value**
Subjectivity	Zero Mean	-6.469	0.01***	-10.267	0.01***
	Non-Zero Mean	-2.746	0.262	-4.348	0.01***
	Trend	-4.937	0.01***	-3.898	0.015**
Polarity	Zero Mean	-5.149	0.01***	-3.890	0.015**
	Non-Zero Mean	-5.903	0.01***	-4.995	0.01***
	Trend	-4.787	0.01***	-3.644	0.029**
Positivity	Zero Mean	-4.686	0.01***	-6.076	0.01***
	Non-Zero Mean	-3.189	0.090	-5.454	0.01***
	Trend	-3.627	0.031*	-4.547	0.01***
Negativity	Zero Mean	-3.727	0.023**	-6.787	0.01***
	Non-Zero Mean	-4.249	0.01***	-3.278	0.075
	Trend	-3.765	0.021**	-3.103	0.111

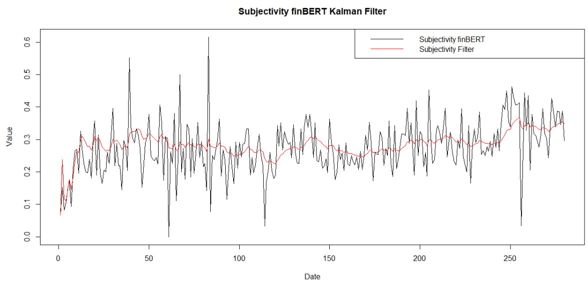
*Notes: *** p<0.01, ** p<0.05, * p<0.1.*

Table 3: Dickey-Fuller Test Results of Introductory Part

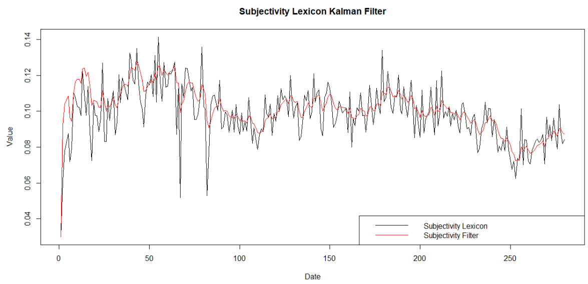
		Introductory Part				**Q&A Part**			
		FinBERT		**Lexicon**		**FinBERT**		**Lexicon**	
Variable	**ADF Type**	**ADF**	**p value**	**ADF**	**p value**	**ADF**	**p value**	**ADF**	**p value**
Subjectivity	Zero Mean	-3.852	0.017**	-5.715	0.010***	-3.325	0.067	-3.669	0.027**
	Non-Zero Mean	-2.641	0.306	-2.774	0.250	-3.065	0.127	-2.545	0.347
	Trend	-3.489	0.044*	-2.369	0.420	-2.716	0.275	-3.250	0.080
Polarity	Zero Mean	-4.178	0.010***	-4.054	0.010***	-4.260	0.01***	-3.780	0.020**
	Non-Zero Mean	-4.570	0.010***	-5.385	0.010***	-4.559	0.01***	-3.801	0.019**
	Trend	-4.346	0.010***	-4.694	0.010***	-4.133	0.01***	-3.573	0.036**
Positivity	Zero Mean	-2.413	0.402	-5.266	0.010***	-4.718	0.01***	-5.405	0.01***
	Non-Zero Mean	-1.833	0.646	-4.734	0.010***	-4.330	0.01***	-4.421	0.01***
	Trend	-2.226	0.481	-4.275	0.010***	-3.047	0.135	-5.105	0.01***
Negativity	Zero Mean	-3.178	0.092	-4.261	0.010***	-3.140	0.098	-2.900	0.197
	Non-Zero Mean	-3.307	0.070	-4.003	0.010***	-3.457	0.047*	-2.597	0.324
	Trend	-3.296	0.072	-3.372	0.059	-3.772	0.021**	-2.719	0.273

*Notes: *** p<0.01, ** p<0.05, * p<0.1.*

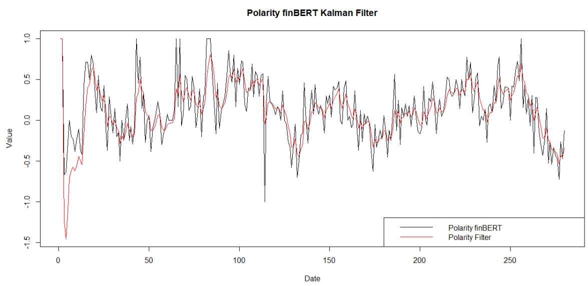
In Table 3, the outputs of the finBERT model are more non-stationarity compare to the LM model. However, both two models have consensus on the stationarity of polarity and non-stationarity of negativity in all sections.



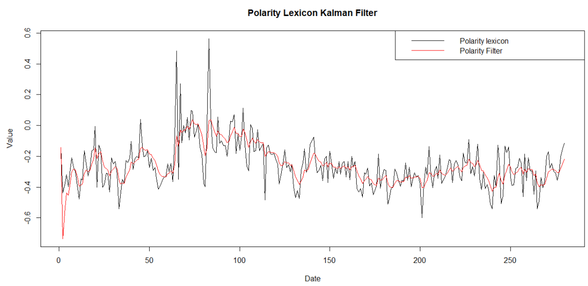
(a) Subjectivity finBERT



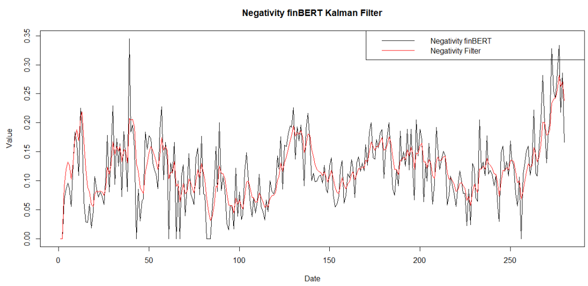
(b) Subjectivity lexicon



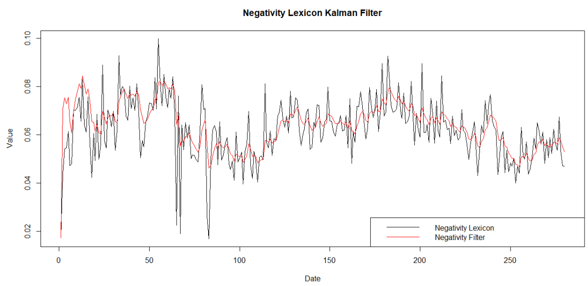
(c) Polarity finBERT



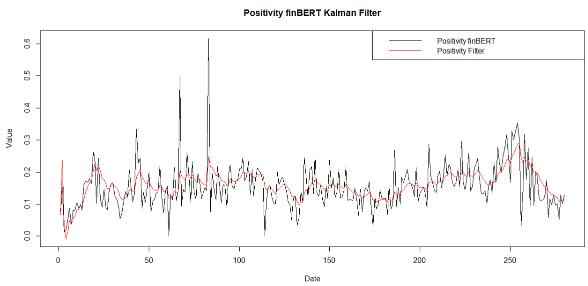
(d) Polarity lexicon



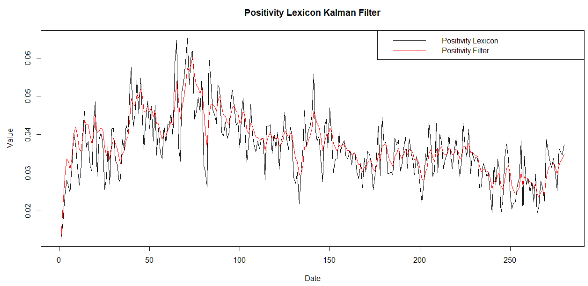
(e) Negativity finBERT



(f) Negativity lexicon



(g) Positivity finBERT



(h) Positivity lexicon

Figure 1: A comprehensive view of full press conferences

4.3 Do different presidents have different communication strategies?

In this section, we focus on a subset of data representing four periods corresponding to the tenures of four presidents of the ECB, as shown in Table 4.

Table 4: Number of Presidents of ECB

No	President	Start Date	End date
Period 1	Wim Duisenberg	1 June 1998	31 October 2003
Period 2	Jean-Claude Trichet	1 November 2003	31 October 2011
Period 3	Mario Draghi	1 November 2011	31 October 2019
Period 4	Christine Lagarde	1 November 2019	Incumbent

The result of the stationarity testing along periods was shown in Table 5. On the contrary with the full periods series with many series are stationarity, these four sub periods are non-stationarity.

Table 5: Dickey-Fuller Test Results of Full conferences of 4 periods

		Duisenberg period				**Draghi period**			
		FinBERT		**Lexicon**		**FinBERT**		**Lexicon**	
Variable	**ADF Type**	**ADF**	**p value**	**ADF**	**p value**	**ADF**	**p value**	**ADF**	**p value**
Subjectivity	Zero Mean	-2.098	0.535	-0.536	0.978	-2.263	0.468	-2.041	0.559
	Non-Zero Mean	-1.217	0.893	-0.479	0.980	-2.022	0.566	-1.777	0.667
	Trend	-0.801	0.957	-0.637	0.971	-2.00	0.576	-1.115	0.914
Polarity	Zero Mean	-2.816	0.244	-4.197	0.01***	-2.640	0.314	-2.820	0.241
	Non-Zero Mean	-3.141	0.112	-3.369	0.068	-2.487	0.377	-3.313	0.076
	Trend	-3.238	0.089	-3.263	0.085	-2.741	0.273	-4.033	0.013**
Positivity	Zero Mean	-3.844	0.022*	-1.214	0.894	-2.594	0.333	-3.202	0.094
	Non-Zero Mean	-2.346	0.435	-1.488	0.783	-2.460	0.387	-3.339	0.072
	Trend	-2.095	0.536	-1.369	0.831	-2.896	0.210	-3.888	0.019
Negativity	Zero Mean	-2.530	0.360	-2.063	0.549	-2.505	0.369	-1.964	0.590
	Non-Zero Mean	-2.335	0.439	-1.910	0.612	-2.438	0.397	-2.043	0.558
	Trend	-2.937	0.195	-2.151	0.514	-2.274	0.464	-1.584	0.746
		Trichet period				**Lagarde period**			
		FinBERT		**Lexicon**		**FinBERT**		**Lexicon**	
Variable	**ADF Type**	**ADF**	**p value**	**ADF**	**p value**	**ADF**	**p value**	**ADF**	**p value**
Subjectivity	Zero Mean	-2.477	0.379	-2.203	0.493	-0.815	0.952	-0.677	0.963
	Non-Zero Mean	-1.586	0.748	-2.423	0.401	-1.160	0.901	-0.310	0.985
	Trend	-1.603	0.741	-2.783	0.252	-0.976	0.928	-0.512	0.976
Polarity	Zero Mean	-2.423	0.401	-1.128	0.914	-1.716	0.684	-2.659	0.316
	Non-Zero Mean	-2.512	0.364	-1.381	0.834	-1.158	0.901	-2.786	0.266
	Trend	-2.390	0.415	-1.299	0.867	-1.435	0.794	-2.645	0.321
Positivity	Zero Mean	-3.263	0.081	-2.429	0.399	-0.022	0.990	-2.348	0.437
	Non-Zero Mean	-2.731	0.273	-3.293	0.076	0.010	0.990	-2.470	0.390
	Trend	-2.627	0.317	-2.563	0.343	0.054	0.990	-2.536	0.364
Negativity	Zero Mean	-1.859	0.635	-1.164	0.909	-1.785	0.657	-0.741	0.958
	Non-Zero Mean	-1.483	0.791	-1.215	0.901	-1.911	0.608	-0.526	0.976
	Trend	-1.968	0.590	-1.493	0.787	-1.759	0.668	-0.697	0.962

*Notes: *** p<0.01, ** p<0.05, * p<0.1.*

[Place Table 4 here]

Table 6 shows the results of testing the difference in mean (panel A) and variance (panel B) in four periods per pair. In panel A, the significant p-value indicates differences in the mean of the pair-period in almost all cases. Particularly, periods 1 and 2 have similar means in terms of the subjectivity of the finBERT model and the positivity of the LM model. The mean of the subjectivity of the LM model in periods 2 and 3 is similar, while periods 1 and 3 do not show a difference in the mean of subjectivity and negativity of the finBERT model. Periods 1 and 4 have the same mean of polarity in both models (finBERT and LM). Additionally, the pair of periods 3 and 4 have non-different means of polarity in the finBERT model.

In panel B, there is a significant difference in variance, except for periods 1 and 3. The other pairs of periods show some similarity in variance. Specifically, there is no significant difference in variance between period 2 and 1 for negativity in the finBERT model, as well as for polarity, positivity, and negativity in the LM model. Period 2 versus 3 also shows no significant difference in variance for polarity, subjectivity, and positivity in the finBERT model. On the other hand, there is no significant difference in the variance of subjectivity in the finBERT model between period 2 and 4. Additionally, there are various non-significant differences in variance between period 3 and 4 for subjectivity in the finBERT model, as well as for polarity.

Table 6: Pairwise t-tests

Panel A: T-test in mean						
Variable	**P2 vs P1**	**P2 vs P3**	**P2 vs P4**	**P1 vs P3**	**P1 vs P4**	**P3 vs P4**
Polarity Bert	0.0000***	0.0007***	0.0213**	0.0411**	0.1846	0.9371
Subjectivity Bert	0.3609	0.0000***	0.0000***	0.2420	0.0000***	0.0000*
Positivity Bert	0.0000***	0.0609*	0.0000***	0.0109*	0.0000***	0.0000***
Negativity Bert	0.0003***	0.0000***	0.0000***	0.9036	0.0051**	0.0005***
Polarity Lex	0.0000***	0.0000***	0.0000***	0.0003***	0.0973	0.0271*
Subjectivity Lex	0.0000***	0.4286	0.0000***	0.0000***	0.0000***	0.0000***
Positivity Lex	0.1468	0.0000***	0.0000***	0.0000***	0.0000***	0.0000***
Negativity Lex	0.0000***	0.0000***	0.0000***	0.0005***	0.0000***	0.0000***
Panel B: T-test in variance						
Variable	**P2 vs P1**	**P2 vs P3**	**P2 vs P4**	**P1 vs P3**	**P1 vs P4**	**P3 vs P4**
polarity_bert_fil	0.000***	0.111	0.006**	0.000***	0.120	0.000***
subjectivity_bert_fil	0.000***	0.357	0.774	0.000***	0.000***	0.676
positivity_bert_fil	0.000***	0.279	0.000***	0.000***	0.631	0.000***
negativity_bert_fil	0.095	0.002**	0.000***	0.000***	0.045*	0.000***
polarity_lex_fil	0.692	0.000***	0.000***	0.000***	0.000***	0.445
subjectivity_lex_fil	0.000***	0.001**	0.000***	0.000***	0.000***	0.249
positivity_lex_fil	0.258	0.000***	0.000***	0.000***	0.000***	0.370
negativity_lex_fil	0.187	0.000***	0.000***	0.000***	0.000***	0.008**

*P1 represents the period of ECB President Duisenberg, P2 for President Trichet, P3 for President Draghi, and P4 for President Lagarde. *** p<0.01, ** p<0.05, * p<0.1*

4.4 Structural breaks in tone

Whole press conferences In this section, we investigate the structural break and the period of the presidents to answer the question does tone change in communications of presidents.

To determine the structural breaks, we apply the Bai & Perron structural break model (Bai & Perron, 1998, 2003). The Bai & Perron structural break model is a modeling approach used to identify break dates.

The results related to the structural breaks, including the RSS (Residual Sum of Squares) and BIC (Bayesian Information Criterion) values, as well as the break dates of the entire ECB's press conferences, are in Table 7 and 8. The Bai & Perron results of the introductory and the Q&A part can be found in the appendix.

The results confirm the presence of significant structural breaks in all indicators of the press conferences. Based on the BIC value, all indicators are consistent with models that have five breaks, except for the finBERT subjectivity indicator, which has the smallest BIC and is consistent with a model that has four breaks.

Table 7: Structural break of full text (Part 1)

Subjectivity finBERT sup.F = 117.46, p-value < 0.000			
Breaks	**RSS**	**BIC**	**Break dates**
0	0.063	-1513.00	
1	0.029	-1710.00	"1999-01-07"
2	0.025	-1738.00	"1999-01-07", "2020-04-30"
3	0.023	-1739.00	"1999-01-07", "2014-02-06", "2020-06-04"
4	0.022	-1739.00	"1999-01-07", "2004-12-02", "2014-01-09", "2020-06-04"
5	0.021	-1737.00	"1999-01-07", "2001-07-05", "2003-05-08", "2014-02-06", "2020-06-04"
Subjectivity lexicon sup.F = 117.46, p-value < 0.000			
Breaks	**RSS**	**BIC**	**Break dates**
0	0.040	-1650.00	
1	0.024	-1783.00	"2017-12-14"
2	0.018	-1853.00	"2005-01-13", "2019-10-24"
3	0.014	-1903.00	"1998-12-22", "2005-01-13", "2019-10-24"
4	0.013	-1911.00	"1998-12-22", "2005-01-13", "2017-09-07", "2019-12-12"
5	0.012	-1927.00	"1998-12-22", "2005-01-13", "2008-02-07", "2017-01-19", "2019-12-12"
Polarity finBERT sup.F = 20.114, p-value = 0.000			
Breaks	**RSS**	**BIC**	**Break dates**
0	29.17	173.87	
1	24.22	133.57	"1999-07-15"
2	21.12	106.83	"1999-07-15", "2022-09-08"
3	19.63	97.86	"1999-07-15", "2016-04-21", "2022-09-08"
4	17.43	76.19	"1999-07-15", "2008-06-05", "2016-03-10", "2022-09-08"
5	15.09	47.53	"1999-06-02", "2003-09-04", "2007-07-05", "2016-04-21", "2022-09-08"
Polarity lexicon sup.F = 101.2, p-value < 0.000			
Breaks	**RSS**	**BIC**	**Break dates**
0	3.5075	-412.9049	
1	2.564	-488.4516	"2008-09-04"
2	1.2464	-677.0152	"2003-09-04", "2008-02-07"
3	1.0671	-708.7911	"1998-12-01", "2003-09-04", "2008-02-07"
4	0.9162	-739.7661	"1998-12-01", "2003-09-04", "2007-07-05", "2011-10-06"
5	0.8464	-750.4613	"1998-12-01", "2003-09-04", "2007-07-05", "2011-11-03", "2015-01-22"

Table 8: Structural break of full text (Part 2)

Positivity finBERT sup.F = 57.978, p-value = 0.000			
Breaks	**RSS**	**BIC**	**Break dates**
0	0.4659	-972.0799	
1	0.3506	-1039.5917	"1999-07-15"
2	0.2914	-1079.5529	"1999-06-02", "2016-03-10"
3	0.2307	-1132.9787	"1999-06-02", "2019-12-12", "2022-09-08"
4	0.2001	-1161.1974	"1999-07-15", "2007-08-02", "2016-03-10", "2022-12-15"
5	0.1685	-1197.5102	"1999-07-15", "2007-08-02", "2015-01-22", "2020-04-30", "2022-07-21"
Positivity lexicon sup.F = 204.8, p-value < 0.000			
Breaks	**RSS**	**BIC**	**Break dates**
0	0.0142	-1939.0000	
1	0.0081	-2082.0000	"2010-10-07"
2	0.0063	-2140.0000	"2001-07-05", "2007-02-08"
3	0.0046	-2220.0000	"2001-07-05", "2007-01-11", "2018-07-26"
4	0.0039	-2254.0000	"1998-12-01", "2001-07-05", "2007-01-11", "2018-07-26"
5	0.0033	-2290.0000	"1998-12-01", "2001-07-05", "2006-08-31", "2011-05-05", "2018-09-13"
Negative finBERT sup.F = 32.023, p-value = 0.0000			
Breaks	**RSS**	**BIC**	**Break dates**
0	0.5483	-926.9828	
1	0.3917	-1008.8754	"2022-10-27"
2	0.3641	-1017.8437	"2008-04-10", "2022-10-27"
3	0.2933	-1066.5142	"2003-09-04", "2003-03-06", "2022-10-27"
4	0.2648	-1083.5765	"2003-09-04", "2008-04-10", "2009-10-08", "2022-10-27"
5	0.2439	-1095.0637	"2000-10-05", "2003-09-04", "2008-04-10", "2009-10-08", "2022-10-27"
Negative lexicon sup.F = 87.081, p-value < 0.000			
Breaks	**RSS**	**BIC**	**Break dates**
0	0.0227	-1809.0000	
1	0.0172	-1874.0000	"2003-09-04"
2	0.0141	-1918.0000	"2003-10-02", "2008-02-07"
3	0.0095	-2016.0000	"2003-10-02", "2008-03-06", "2019-12-12"
4	0.0087	-2029.0000	"2003-10-02", "2008-03-06", "2016-07-21", "2020-04-30"
5	0.0081	-2040.0000	"1998-12-22", "2003-10-02", "2008-03-06", "2016-07-21", "2020-04-30"

Note

Figure 2 displays the structural breaks (green dotted lines) and the president breaks (red dashed lines).

The results also show an obvious change in the tone of communications between two presidents Wim Duisenberg and Jean-Claude Trichet (in October 2003). Specifically, four indicators were analyzed: the negativity of both the finBERT and LM models, as well as the polarity of both the finBERT and LM models. There is a downtrend in negativity after the transition from Wim Duisenberg to Jean-Claude Trichet. Moreover, the polarity of Jean-Claude Trichet exhibits wider fluctuations compared to the previous period.

For the second president break (from Jean-Claude Trichet to Mario Draghi on 31st October 2011), the second structural break of the polarity lexicon coincides with this change. Table 2 shows that this structural break indicates a dramatic change in the fluctuation of the polarity value. We argue that this structural break clearly signifies a different communication strategy between Jean-Claude Trichet and Mario Draghi. Specifically, Mario Draghi tends to avoid spreading a particular tone in his communications.

[Place Figure 2 here]

For the third transition in the presidency period (from Mario Draghi on October 31, 2019 to Christine Lagarde), only the subjectivity lexicon accurately captures the same transition as the presidential transition. Specifically, the subjectivity of the lexicon model decreases during Christine Lagarde's tenure, while the subjectivity of the finBERT model increases during the same period. The two methods (the lexicon approach and the large language approach) have different approaches, with the lexicon approach relying on a predetermined dictionary to determine sentiment. Therefore, we argue that when speakers use neutral, advanced, or less common vocabulary, the lexicon-based method may struggle to capture sentiment from the context compared to the large language model. In fact, Leippold (2023) indicates that keyword-based approaches are susceptible to sentiment adversarial manipulation, which is a technique commonly used in human communication to change the sentiment of sentences without significantly altering their meaning. Moreover, they recommend employing context-aware approaches to enhance the robustness of financial sentiment analysis results.

In general, based on the indicators from the full press conferences, the polarity of the lexicon model most accurately reflects the changes in presidents. Specifically, the polarity based on the LM model aligns with two different presidential transitions, resulting in two distinct breaks.

Introductory statements Table 3 displays the filter values of four indicators with the Bai-Perron structural breaks and president breaks of introductory parts.

The introductory parts do not connect fully with the president period as the full press conferences. Except for the negativity finBERT, which has one break that matches the change from president Jean-Claude Trichet to Mario Draghi. However, the introductory parts show a close connection with macro shocks and crises. Specifically, the negativity finBERT model plot displays four structural breaks during the period from 2008 to 2012. The negativity increases and reaches its peak in 2008, as a consequence of the global crisis. Then it declines rapidly in 2010, only to hit a new peak in 2011 and increase slightly again. Additionally, the negativity finBERT shows an uptrend during the COVID-19 period (2019) and again reaches its peak during the Ukraine war (2022). The structural breaks in the tone of communications from 2008 to 2011 are consistent across many indicators, including negativity lexicon, polarity lexicon, positivity lexicon, and subjectivity finBERT. Furthermore, tone structural breaks appear during the Covid period in indicators such as subjectivity lexicon and finBERT, positivity finBERT, and polarity finBERT. Moreover, an uptrend of the negativity indicator was found in both approaches during the Ukraine war (2022).

[Place Figure 3 here]

Q&A parts Contrary to the tones of introductory statements, which are related to macro events, the tones of Q&A sections show strong influences from the presidential periods. Specifically, the structural breaks of the LM model polarity changes coincide with the replacements of Wim Duisenberg by Jean-Claude Trichet, and later from Mario Draghi to Christine Lagarde.

The breaks in the structure of the positivity lexicon are also aligned with the changes in presidency, from Wim Duisenberg to Jean-Claude Trichet, and then from Jean-Claude Trichet to Mario Draghi.

However, the tone of the Q&A sections is affected by the structural breaks caused by the global crisis in 2008. More specifically, there are noticeable structural breaks in all indicators between 2008 and 2009.

When examining the overall dynamic trend of the Q&A sections, it becomes evident that the fluctuations decrease. High fluctuations are observed during Wim Duisenberg's initial period, while the range narrows down under the subsequent presidents. This phenomenon indicates an adjustment in the communication strategy of the presidents. They appear to be more professional and cautious in their responses to questions.

[Place Figure 4 here]

5 Conclusion

The European Central Bank is the heart of the European economic system, and its announcements and decisions can have a wide-ranging influence not only in Europe but also worldwide. The ECB's communications also received attention both in practice and in research. This study analyzes the sentiment of the ECB's press conferences using state-of-the-art models in finance, namely finBERT and the LM model.

The sentiment in the introductory parts of the press conferences is closely linked to economic shocks and crises. There is a noticeable increase in negativity during events such as the Ukraine war and the global crisis. This connection can be attributed to the nature of the introductory parts, which serve two main purposes: (1) announcing monetary policies and economic analysis, and (2) providing forecasts. Thus, the introductory parts cannot ignore the macro shocks since the ECB need to declare their viewpoints and their reactions to crises in order to maintain market stability and avoid market crashes.

On the other hand, the tone in the Q&A parts of the press conferences reflects the personal style of the ECB's president, as these responses are given in reaction to questions. Over time, there is a decrease in the range of tones used, indicating an evolution in the communication strategy of the presidents. The ECB's presidents become more cautious in their responses and tend to use less emotional vocabulary. Another interesting finding is the presence of structural breaks in all indicators around the global crisis in 2008.

This study makes a significant contribution to financial communication and sentiment analysis by utilizing advanced models like finBERT and the LM model to analyze the ECB's press conferences. It demonstrates the correlation between sentiment in ECB communications and economic shocks. Additionally, the research reveals an evolution in the ECB's communication strategy, showing a shift towards more cautious and less emotional responses over time in the Q&A sections. This provides valuable insights into how central banks adapt their messaging to shape market perceptions and maintain stability. The study enriches both academic literature and practical understanding for policymakers and financial analysts.

References

Anastasiou, D., & Katsafados, A. (2023). Bank deposits and textual sentiment: When an european central bank president's speech is not just a speech. *The Manchester School*, *91*(1), 55–87.

Apergis, N., & Pragidis, I. (2019). Stock price reactions to wire news from the european central bank: Evidence from changes in the sentiment tone and international market indexes. *International Advances in Economic Research*, *25*, 91–112.

Araci, D. (2019). Finbert: Financial sentiment analysis with pre-trained language models. *arXiv preprint arXiv:1908.10063*.

Bai, J., & Perron, P. (1998). Estimating and testing linear models with multiple structural changes. *Econometrica*, 47–78.

Bai, J., & Perron, P. (2003). Critical values for multiple structural change tests. *The Econometrics Journal*, *6*(1), 72–78.

Beirne, J. (2012). The eonia spread before and during the crisis of 2007–2009: The role of liquidity and credit risk. *Journal of International Money and Finance*, *31*(3), 534–551.

Bennani, H. (2020). Central bank communication in the media and investor sentiment. *Journal of Economic Behavior & Organization*, *176*, 431–444.

Bennani, H., & Neuenkirch, M. (2017). The (home) bias of european central bankers: new evidence based on speeches. *Applied Economics*, *49*(11), 1114–1131.

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.

Bochkay, K., Brown, S. V., Leone, A. J., & Tucker, J. W. (2023). Textual analysis in accounting: What's next? *Contemporary accounting research*, *40*(2), 765–805.

Borovkova, S., Garmaev, E., Lammers, P., & Rustige, J. (2017). Sensr: A sentiment-based systemic risk indicator.

Brand, C., Buncic, D., & Turunen, J. (2010). The impact of ecb monetary policy decisions and communication on the yield curve. *Journal of the European Economic Association*, *8*(6), 1266–1298.

Correa, R., Garud, K., Londono, J. M., & Mislang, N. (2021). Sentiment in central banks' financial stability reports. *Review of Finance*, *25*(1), 85–120.

Dong, M. M., Stratopoulos, T. C., & Wang, V. X. (2023). A scoping review of chatgpt research in accounting and finance. *Theophanis C. and Wang, Victor Xiaoqi, A Scoping Review of ChatGPT Research in Accounting and Finance (December 30, 2023)*.

Dybowski, T. P., & Kempa, B. (2020). The european central bank's monetary pillar after the financial crisis. *Journal of Banking & Finance*, *121*, 105965.

Ehrmann, M., & Fratzscher, M. (2009). Explaining monetary policy in press conferences. *International Journal of Central Banking*, *5 (2)*, 42–84.

Ehrmann, M., & Talmi, J. (2020). Starting from a blank page? semantic similarity in central bank communication and market volatility. *Journal of Monetary Economics*, *111*, 48–62.

Ehrmann, M., & Wabitsch, A. (2022). Central bank communication with non-experts–a road to nowhere? *Journal of Monetary Economics*, *127*, 69–85.

Ferrara, F. M., & Angino, S. (2022). Does clarity make central banks more engaging? lessons from ecb communications. *European Journal of Political Economy*, *74*, 102146.

Gao, G., Nikolsko-Rzhevskyy, A., & Talavera, O. (2023). Can central banks be heard over the sound of gunfire? *Journal of Financial Research*, *46*, S183–S203.

Gorodnichenko, Y., Pham, T., & Talavera, O. (2023). The voice of monetary policy. *American Economic Review*, *113*(2), 548–584.

Hayo, B., & Neuenkirch, M. (2013). Do federal reserve presidents communicate with a regional bias? *Journal of Macroeconomics*, *35*, 62–72.

Hayo, B., & Zahner, J. (2023). What is that noise? analysing sentiment-based variation in central bank communication. *Economics Letters*, *222*, 110962.

Hu, N., Liang, P., & Yang, X. (2023). Whetting all your appetites for financial tasks with one meal from gpt? a comparison of gpt, finbert, and dictionaries in evaluating sentiment analysis. In *Whetting all your appetites for financial tasks with one meal from gpt? a comparison of gpt, finbert, and dictionaries in evaluating sentiment analysis: Hu, nan/ uliang, peng/ uyang, xu.* [Sl]: SSRN.

Huang, A. H., Wang, H., & Yang, Y. (2023). Finbert: A large language model for extracting information from financial text. *Contemporary Accounting Research*, *40*(2), 806–841.

Hubert, P., & Fabien, L. (2017). Central bank sentiment and policy expectations.

Hussain, S. M. (2011). Simultaneous monetary policy announcements and international stock markets response: An intraday analysis. *Journal of Banking & Finance*, *35*(3), 752–764.

Jansen, D.-J. (2011). Mumbling with great incoherence: Was it really so difficult to understand alan greenspan? *Economics Letters*, *113*(1), 70–72.

Jansen, D.-J., & De Haan, J. (2006). Look who's talking: Ecb communication during the first years of emu. *International Journal of Finance & Economics*, *11*(3), 219–228.

Kalman, R. E., et al. (1960). A new approach to linear filtering and prediction problems [j]. *Journal of basic Engineering*, *82*(1), 35–45.

Kaminskas, R., & Jurkšas, L. (2023). *Ecb communication sentiments: how do they relate to the economic environment and financial markets?* Lietuvos Bankas.

Kanelis, D., & Siklos, P. L. (2024). The ecb press conference statement: deriving a new sentiment indicator for the euro area. *International Journal of Finance & Economics*.

Klejdysz, J., & Lumsdaine, R. L. (2023). Shifts in ecb communication: A textual analysis of the press conference. *International Journal of Central Banking*, *19*(2), 473–542.

Leippold, M. (2023). Sentiment spin: Attacking financial sentiment with gpt-3. *Finance Research Letters*, 103957.

Leombroni, M., Vedolin, A., Venter, G., & Whelan, P. (2021). Central bank communication and the yield curve. *Journal of Financial Economics*, *141*(3), 860–880.

Linzert, T., & Schmidt, S. (2011). What explains the spread between the euro overnight rate and the ecb's policy rate? *International Journal of Finance & Economics*, *16*(3), 275–289.

Loughran, T., & McDonald, B. (2011). When is a liability not a liability? textual analysis, dictionaries, and 10-ks. *The Journal of finance*, *66*(1), 35–65.

Moniz, A., & de Jong, F. (2014). Predicting the impact of central bank communications on financial market investors' interest rate expectations. In *The semantic web: Eswc 2014 satellite events: Eswc 2014 satellite events, anissaras, crete, greece, may 25-29, 2014, revised selected papers 11* (pp. 144–155).

Mullings, R. (2023). Do central bank sentiment shocks affect liquidity within the european monetary union? a computational linguistics approach. *The European Journal of Finance*, *29*(12), 1355–1381.

Niszczota, P., & Abbas, S. (2023). Gpt has become financially literate: Insights from financial literacy tests of gpt and a preliminary test of how people use it as a source of advice. *Finance Research Letters*, *58*, 104333.

Niţoi, M., Pochea, M.-M., & Radu, Ş.-C. (2023). Unveiling the sentiment behind central bank narratives: A novel deep learning index. *Journal of Behavioral and Experimental Finance*, *38*, 100809.

Picault, M., Pinter, J., & Renault, T. (2022). Media sentiment on monetary policy: Determinants and relevance for inflation expectations. *Journal of International Money and Finance*, *124*, 102626.

Picault, M., & Renault, T. (2017). Words are not all created equal: A new measure of ecb communication. *Journal of International Money and Finance*, *79*, 136–156.

Rosa, C., & Verga, G. (2007). On the consistency and effectiveness of central bank communication: Evidence from the ecb. *European Journal of Political Economy*, *23*(1), 146–175.

Schmeling, M., & Wagner, C. (2016). Does central bank tone move asset prices? *Journal of Financial and Quantitative Analysis*, 1–48.

Shapiro, A. H., Sudhof, M., & Wilson, D. J. (2022). Measuring news sentiment. *Journal of econometrics*, *228*(2), 221–243.

Stekler, H., & Symington, H. (2016). Evaluating qualitative forecasts: The fomc minutes, 2006–2010. *International Journal of Forecasting*, *32*(2), 559–570.

Welch, G., Bishop, G., et al. (1995). An introduction to the kalman filter.

Woodford, M. (2005). *Central bank communication and policy effectiveness.* National Bureau of Economic Research Cambridge, Mass., USA.

Zhang, C. L. (2023). Feel the market: an attempt to identify additional factor in the capital asset pricing model (capm) using generative pre-trained transformer (gpt) and bidirectional encoder representations from transformers (bert). *Available at SSRN 4521946*.

A Appendix 1: structural breaks

Table 9: Structural breaks of Q&A (Part 1)

Subjectivity finBERT sup.F = 54.926, p-value = 7.107e-12			
Lag	**RSS**	**BIC**	**Break dates**
0	0.309	-1037.00	
1	0.256	-1076.00	"2003-12-04"
2	0.142	-1223.00	"2004-05-06", "2012-06-06"
3	0.120	-1256.00	"2004-06-03", "2008-09-04" , "2012-11-08"
4	0.100	-1293.00	"1999-01-07", "2004-06-03", "2008-09-04", "2012-11-08"
5	0.086	-1323.00	"1999-01-07", "2004-06-03", "2008-09-04 ", "2012-11-08", "2019-03-07"
Subjectivity lexicon sup.F = 54.926, p-value = 7.107e-12			
Lag	**RSS**	**BIC**	**Break dates**
0	0.047	-1537.00	
1	0.026	-1683.00	"2020-01-23"
2	0.017	-1796.00	"2004-06-03", "2020-04-30"
3	0.013	-1841.00	"2004-12-02", "2009-10-08", "2020-01-23"
4	0.012	-1855.00	"1999-10-07", "2004-12-02", "2009-10-08", "2020-01-23"
5	0.010	-1904.00	"1999-11-04", "2001-05-10", "2004-06-03", "2009-10-08","2020-01-23"
Polarity finBERT sup.F = 16.845, p-value = 0.001032			
Lag	**RSS**	**BIC**	**Break dates**
0	9.439	-123.538	"2000-07-06"
1	8.400	-143.510	"2000-07-06"
2	6.488	-201.262	"2000-10-05", "2005-11-03"
3	5.645	-227.251	"2000-10-05", "2005-11-03", "2008-07-03"
4	5.120	-293.641	"2000-10-05", "2005-11-03", "2020-10-29", "2023-06-15"
5	4.049	-242.182	"2000-10-05", "2005-11-03", "2008-07-03", "2020-06-04", "2023-06-15"
Polarity lexicon sup.F = 31.535, p-value = 7.922e-07			
Lag	**RSS**	**BIC**	**Break dates**
0	2.413	-487.7113	
1	2.1564	-506.5575	"2010-09-02"
2	1.1368	-666.3173	"2003-11-06", "2007-07-05"
3	0.9742	-696.3621	"1999-03-04", "2008-03-06", "2020-04-30"
4	0.7868	-742.2209	"2003-11-06", "2007-07-05", "2010-11-04", "2019-12-12"
5	0.7273	-752.038	"1999-03-04", "2003-11-06", "2007-07-05", "2010-11-04", "2019-12-12"

Note

Table 10: Structural breaks of Q&A (Part 2)

Positivity finBERT sup.F = 119.96, p-value < 2.2e-16			
Lag	**RSS**	**BIC**	**Break dates**
0	0.057	-1489.00	
1	0.039	-1578.00	"2013-07-04"
2	0.026	-1679.00	"2001-03-01", "2013-07-04"
3	0.023	-1697.00	"2001-03-01", "2013-07-04", "2023-05-04"
4	0.021	-1704.00	"2001-03-01", "2004-06-03", "2009-06-04", "2013-07-04"
5	0.019	-1727.00	"2001-03-01", "2004-06-03", "2009-06-04", "2013-07-04", "2023-05-04"
Positivity lexicon sup.F = 122.21, p-value < 2.2e-16			
Lag	**RSS**	**BIC**	**Break dates**
0	0.008	-2022.00	
1	0.005	-2112.00	"2011-10-06"
2	0.004	-2153.00	"2003-10-02", "2007-01-11"
3	0.004	-2185.00	"2003-11-06", "2006-08-31", "2011-10-06"
4	0.003	-2221.00	"2000-02-03", "2001-07-05", "2007-02-08", "2019-04-10"
5	0.003	-2249.00	"2000-02-03", "2001-07-05", "2003-11-06", "2006-08-31", "2011-10-06"
Negative finBERT sup.F = 44.542, p-value = 1.258e-09			
Lag	**RSS**	**BIC**	**Break dates**
0	0.251	-1091.552	
1	0.215	-1121.860	"2003-12-04"
2	0.168	-1177.321	"2004-05-06", "2008-07-03"
3	0.135	-1224.447	"2000-10-19", "2004-01-08", "2008-08-07"
4	0.118	-1249.582	"2000-10-19", "2002-01-03", "2004-05-06", "2008-07-03"
5	0.103	-1273.0276	"2000-10-19", "2002-01-03", "2004-05-06", "2008-07-03", "2023-06-15"
Negative lexicon sup.F = 124.83, p-value < 2.2e-16			
Lag	**RSS**	**BIC**	**Break dates**
0	0.038	-1599.00	
1	0.025	-1698.00	"2020-04-30"
2	0.017	-1787.00	"2003-11-06", "2020-04-30"
3	0.010	-1922.00	"2004-01-08", "2009-12-03", "2019-12-12"
4	0.009	-1945.00	"2004-02-05", "2008-02-07", "2010-06-10", "2019-12-12"
5	0.008	-1964.00	"2004-02-05", "2008-02-07", "2010-06-10", "2014-08-07", "2020-04-30"

Note

Table 11: Structural breaks of introductory statement (Part 1)

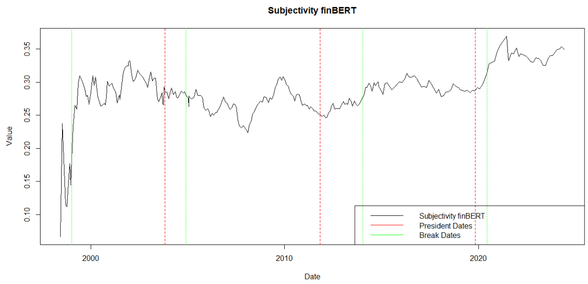
Subjectivity finBERT sup.F = 499, p-value < 2.2e-16			
Breaks	**RSS**	**BIC**	**Break dates**
0	5.806	-273.2916	
1	2.063	-548.683	"2005-09-01"
2	1.048	-725.0001	"1999-03-04", "2005-12-01"
3	0.892	-758.4146	"1999-03-04", "2006-01-12", "2023-05-04"
4	0.809	-774.2563	"1999-03-04", "2005-12-01", "2016-04-21", "2021-06-10"
5	0.688	-807.7364	"1999-03-04", "2006-01-12", "2007-07-05", "2008-05-08", "2021-06-10"
Subjectivity lexicon sup.F = 499, p-value < 2.2e-16			
Breaks	**RSS**	**BIC**	**Break dates**
0	0.067	-1510.00	
1	0.035	-1674.00	"2010-10-07"
2	0.025	-1758.00	"2001-02-01", "2010-09-02"
3	0.022	-1787.00	"2001-03-01", "2005-06-02", "2010-09-02"
4	0.018	-1827.00	"2001-02-01", "2010-09-02", "2017-01-19", "2021-06-10"
5	0.015	-1872.00	"2001-03-01", "2005-06-02", "2010-09-02", "2017-01-19", "2021-06-10"
Polarity finBERT sup.F = 8.2853, p-value = 0.05841			
Breaks	**RSS**	**BIC**	**Break dates**
0	44.640	291.73	
1	33.910	226.78	"1999-06-02"
2	27.950	184.51	"1999-06-02", "2022-09-08"
3	24.810	162.8	"1999-06-02", "2015-03-05", "2022-03-10"
4	20.780	124.9	"1999-06-02", "2008-01-10", "2015-03-05", "2022-03-10"
5	19.360	116.54	"1999-06-02", "2008-01-10", "2015-03-05", "2020-01-23", "2022-09-08"
Polarity lexicon sup.F = 149.3, p-value < 2.2e-16			
Breaks	**RSS**	**BIC**	**Break dates**
0	11.008	-96.088	
1	7.135	-204.966	"2018-12-13"
2	6.035	-240.091	"2015-03-05", "2018-12-13"
3	5.224	-268.828	"2008-01-10", "2015-03-05", "2018-12-13"
4	4.027	-329.655	"1999-03-04", "2008-01-10", "2015-03-05", "2018-12-13"
5	3.553	-353.1	"1999-03-04", "2008-03-06", "2009-08-06", "2015-03-05", "2018-12-13"

Note

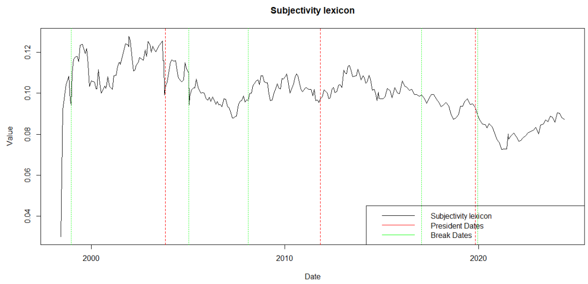
Table 12: Structural breaks of introductory statement (Part 2)

Positivity finBERT sup.F = 98.228, p-value < 2.2e-16			
Breaks	**RSS**	**BIC**	**Break dates**
0	7.279	-210.66	
1	5.345	-284.946	"1999-07-15"
2	4.361	-330.07	"2005-04-07", "2022-09-08"
3	2.944	-427.628	"1999-07-15", "2015-09-03", "2021-06-10"
4	2.414	-471.373	"1999-06-02", "2005-04-07", "2016-03-10", "2021-06-10"
5	1.884	-528.857	"1999-06-02", "2005-11-03", "2007-07-05", "2015-09-03", "2021-06-10"
Positivity lexicon sup.F = 202.82, p-value < 2.2e-16			
Breaks	**RSS**	**BIC**	**Break dates**
0	0.034	-1698.00	
1	0.020	-1839.00	"2018-04-26"
2	0.014	-1927.00	"1999-07-15", "2008-03-06"
3	0.007	-2094.00	"1999-07-15", "2008-01-10", "2018-07-26"
4	0.007	-2107.00	"1999-07-15", "2001-12-13", "2008-01-10", "2018-07-26"
5	0.006	-2122.00	"1999-07-15", "2008-01-10", "2011-07-07", "2012-10-04", "2018-06-14"
Negative finBERT sup.F = 60.648, p-value = 4.049e-13			
Breaks	**RSS**	**BIC**	**Break dates**
0	3.565	-408.39	
1	2.921	-452.346	"2008-03-06"
2	2.338	-502.74	"2008-03-06", "2014-05-08"
3	1.61	-594.785	"2008-03-06", "2014-05-08", "2022-03-10"
4	1.466	-609.478	"2008-03-06", "2011-12-08", "2014-05-08", "2022-03-10"
5	1.121	-672.685	"2008-04-10", "2010-05-06", "2011-11-03", "2014-05-08", "2022-03-10"
Negative lexicon sup.F = 29.251, p-value = 2.452e-06			
Breaks	**RSS**	**BIC**	**Break dates**
0	0.034	-1693.00	
1	0.030	-1718.00	"2022-03-10"
2	0.024	-1773.00	"2015-03-05", "2020-03-12"
3	0.020	-1817.00	"2008-04-10", "2010-08-05", "2022-02-03"
4	0.017	-1840.00	"2008-05-08", "2010-07-08", "2015-03-05", "2020-03-12"
5	0.015	-1865.00	"2001-02-01", "2005-05-04", "2008-03-06", "2010-08-05", "2022-02-03"

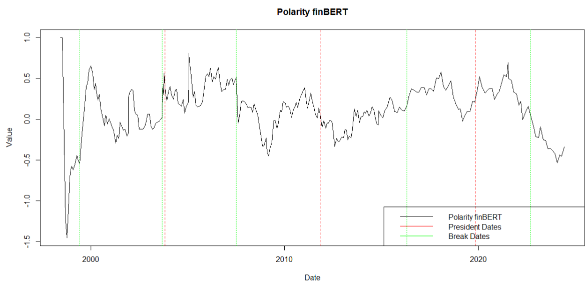
Note



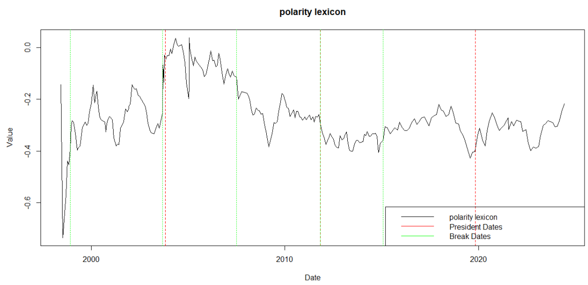
(a) Subjectivity finBERT



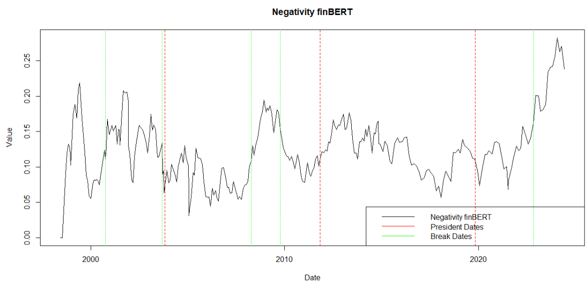
(b) Subjectivity lexicon



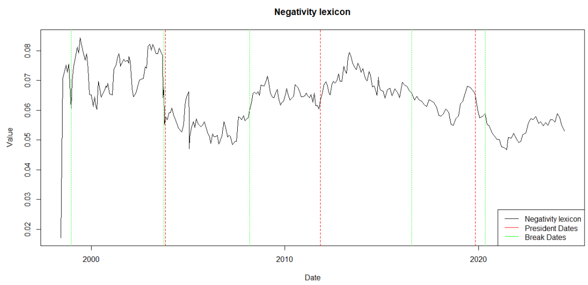
(c) Polarity finBERT



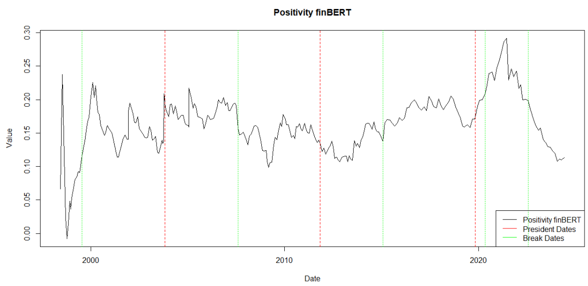
(d) Polarity lexicon



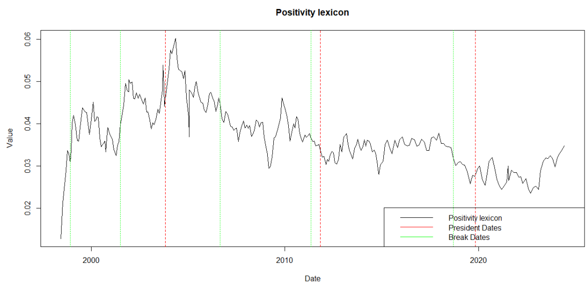
(e) Negativity finBERT



(f) Negativity lexicon

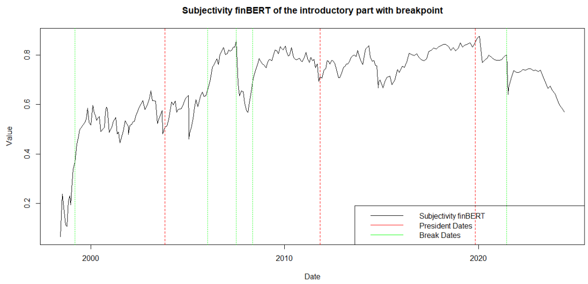


(g) Positivity finBERT

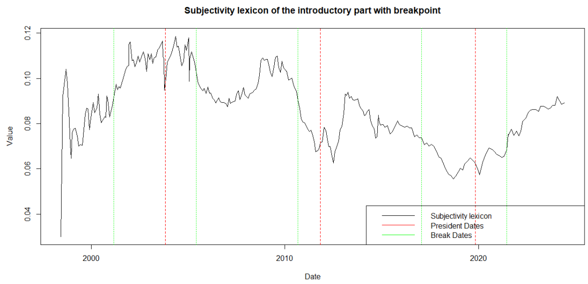


(h) Positivity lexicon

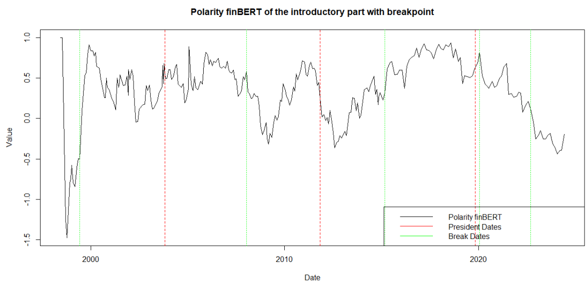
Figure 2: A comprehensive view of full press conferences with break structure



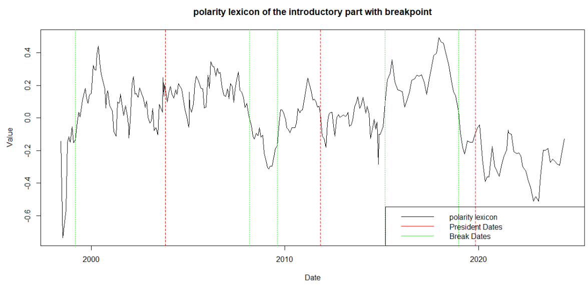
(a) Subjectivity finBERT



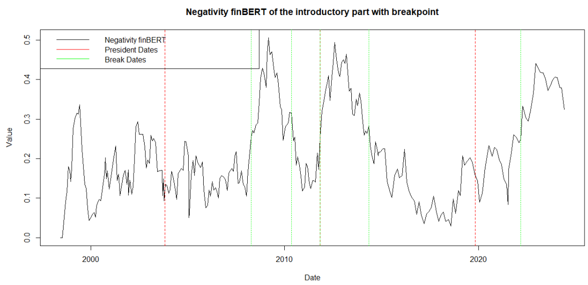
(b) Subjectivity lexicon



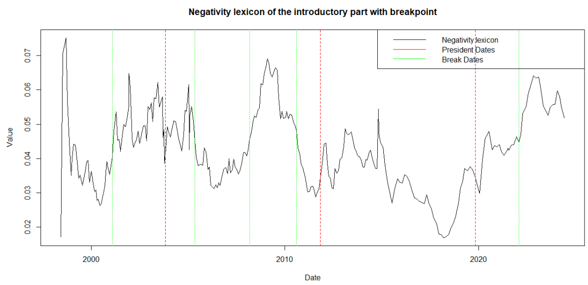
(c) Polarity finBERT



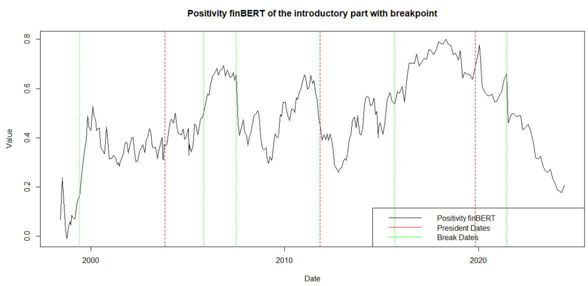
(d) Polarity lexicon



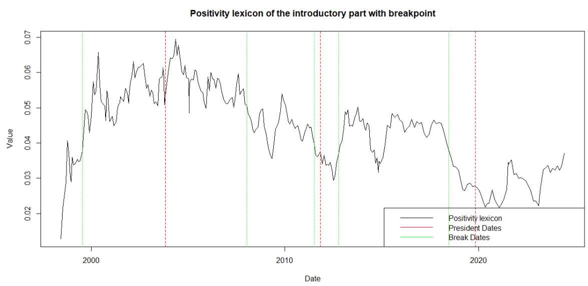
(e) Negativity finBERT



(f) Negativity lexicon

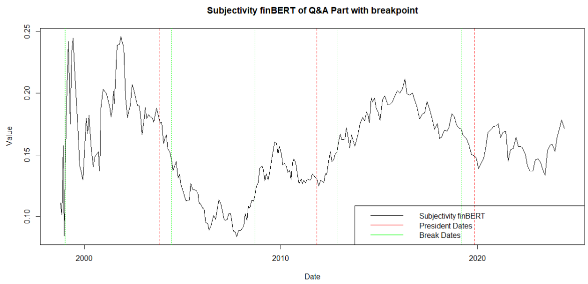


(g) Positivity finBERT

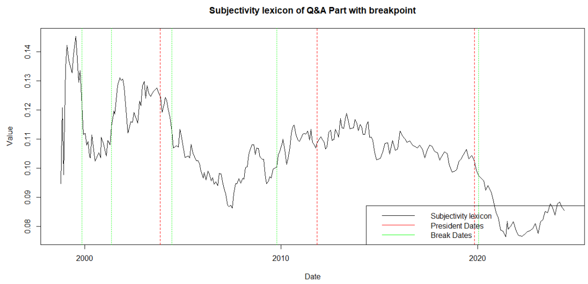


(h) Positivity lexicon

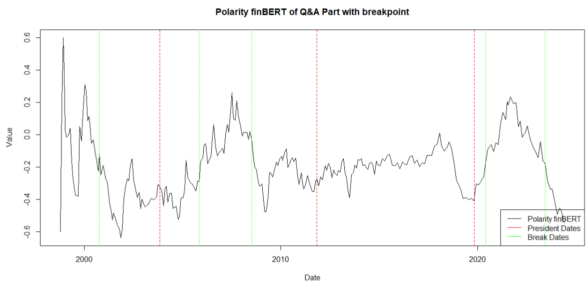
Figure 3: A comprehensive view of the introductory part with break structure



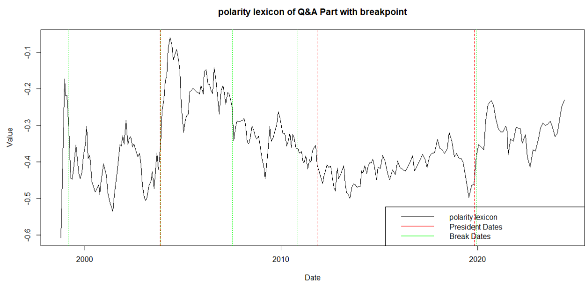
(a) Subjectivity finBERT



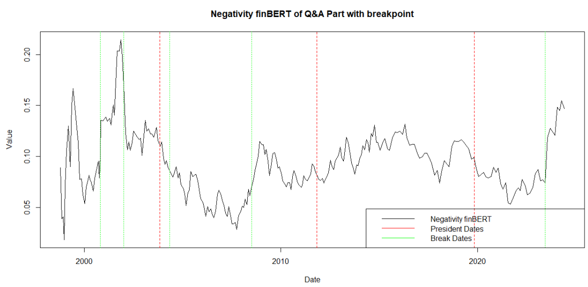
(b) Subjectivity lexicon



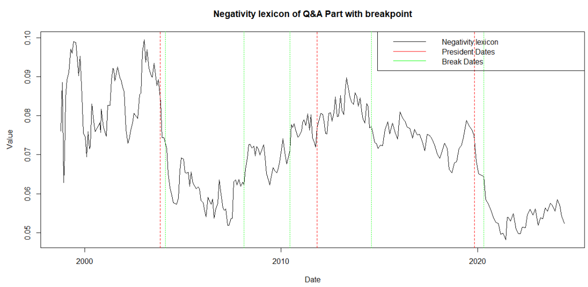
(c) Polarity finBERT



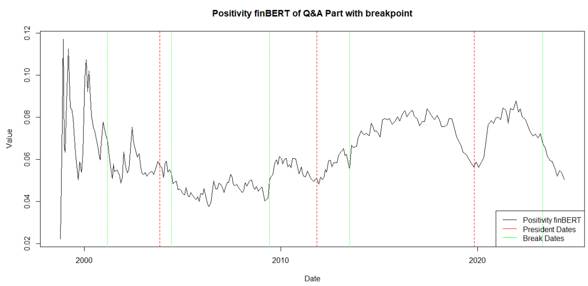
(d) Polarity lexicon



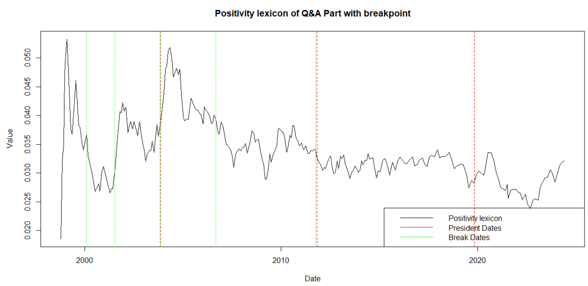
(e) Negativity finBERT



(f) Negativity lexicon



(g) Positivity finBERT



(h) Positivity lexicon

Figure 4: A comprehensive view of the Q&A part with break structure