MASTERARBEIT/ MASTER’S THESIS

Titel der Masterarbeit / Title of the Master’s Thesis

„Central bank announcements and foreign exchange rate movements“

verfasst von / submitted by

Marco Almbauer, BSc

angestrebter akademischer Grad / in partial fulfilment of the requirements for the degree of

Master of Science (MSc)

Wien, 2023 / Vienna 2023

Studienkennzahl lt. Studienblatt / degree programme code as it appears on the student record sheet:	UA 066 953
Studienrichtung lt. Studienblatt / degree programme as it appears on the student record sheet:	Masterstudium Research in Economics and Finance
Betreut von / Supervisor:	Assoz. Prof. Mag. Dr. Paul Pichler, Privatdoz.

Acknowledgements

The author wishes to express his gratitude to Paul Pichler for his excellent mentoring, patience, and support and for granting the author the freedom to pursue this work in any direction.

I want to thank my whole family for their love and support; every family member helped me develop into who I am today. Without my parents, Heidrun and Gerhard, I could have never fulfilled my studies in this depth and quality. I want to thank them for their support. My siblings, Lara and Philipp, have been amazing guinea pigs to test my ability to explain complicated topics to non-economists. I want to thank my grandmothers, Christine and Erika, and my aunt Sigrid, for always having being there for me and for educating me.

I would also like to thank Michael Greinecker for his unwavering support and insightful discussions that have played a significant role throughout my academic career. His comments have always been exceptional. I want to thank Ilya Archakov for discussing and developing ideas. Furthermore, I want to thank Tatiana Evdokimova, Luca Gonzato, Oliver Jakisch, and Maximilian Reiner for many valuable discussions.

Abstract

This thesis aims to quantify the impact of European Central Bank (ECB) communication during the press conference days on the foreign exchange (FX) market. First, textual data released by the ECB about its monetary policy decision and press conference is web-scraped from its internet page and analyzed. Second, the subtitles of the published press conference videos on YouTube are gathered. Then the sentiment of the texts is retrieved with a lexicon approach. This study demonstrates that since 2016, the monetary policy decision has increased in length, changed its structure, and began to vary in conveyed sentiment. To show the effects of these changes, intraday trading behavior on FX markets is analyzed and dissected into pre- and post-2016 periods. This indicates that the importance of the press conference has decreased. Furthermore, this work finds a positive relationship between trading volume on the FX markets and the polarization of the announcements, which is tested using various linear regression specifications. Lastly, a minor result suggests that YouTube subtitles may be used when working with press conference texts.

Kurzfassung

Ziel dieser Arbeit ist es, die Auswirkungen der Kommunikation der Europäischen Zentralbank (EZB) während der Pressekonferenztage auf den Devisenmarkt zu quantifizieren. Zunächst werden die von der EZB veröffentlichten Textdaten über ihre geldpolitischen Entscheidungen und Pressekonferenzen von ihrer Internetseite abgerufen und analysiert. Des Weiteren werden die Untertitel der auf YouTube veröffentlichten Pressekonferenzvideos gesammelt. Anschließend wird das Sentiment der Texte mit einem Lexikonansatz ermittelt. Zudem wird in dieser Arbeit gezeigt, dass die geldpolitischen Entscheidungen seit 2016 länger geworden sind, sich in ihrer Struktur verändert haben und das vermittelte Sentiment zu variieren begann. Um die Auswirkungen dieser Veränderungen aufzuzeigen, wird das Innertageshandelsverhalten an den Devisenmärkten analysiert und in Zeiträume vor und nach 2016 unterteilt. Dabei zeigt sich, dass die Bedeutung der Pressekonferenz abnahm. Darüber hinaus wird in dieser Arbeit ein positiver Zusammenhang zwischen dem Handelsvolumen an den Devisenmärkten und der Polarisierung der Ankündigungen festgestellt. Dies wird mit verschiedenen linearen Regressionsspezifikationen getestet. Zum Abschluss wird ein sekundäres Ergebnis präsentiert, welches besagt, dass YouTube Untertitel zur Bearbeitung von Pressekonferenztexten genutzt werden können.

Contents

1. Introduction

Central banks communicate both quantitative and qualitative information to markets. Since the mid-1990s, central banks have gradually increased transparency by publishing statements, speeches, and reports. The dominant strategy has moved from secrecy to clarity in monetary and exchange rate policy. Several authors have recognized that communication is a powerful and effective instrument to influence markets and help central banks to reach their goals. As Blinder et al. (2008) point out, financial markets react rapidly to central bank communication. Therefore, this thesis aims to measure the effects of central bank communication using high-frequency data. Oral strategies serve as a means for policymakers to communicate their beliefs rather than take direct action. This holds especially at times of high uncertainty (Eusepi & Preston, 2010), when operating at the effective lower bound on nominal interest rates (Coenen et al., 2017), and when utilizing unconventional monetary policy (De Santis, 2020).

With the advent of detailed market data and the increasing availability of information, economic research started to explore possible relationships between communication and market movements. An interdisciplinary approach was developed from three different scientific fields. These include, in addition to economics, linguistics, to gain insight into language characteristics, and machine learning, to enable the ability to create computational models. Given the necessity of interpreting large volumes of text, it was paramount that such information be quantified for proper research utilization. A multitude of potential metrics exist for the quantification of qualitative information. In one such method, the *sentiment* or *tone* is extracted from written text and subsequently included in econometric modeling. As coined by Algaba et al. (2020), this scientific approach is known as *sentometrics*. This methodology will be applied and thoroughly explained in the coming chapters of this thesis.

The European Central Bank's (ECB) communication is particularly interesting in this context. Every sixth week, the rate-setting Government Council meets. The decisions of the meeting are announced in two separate steps on the same day. The monetary policy decision is a brief press release at 14:15 Central European Time (CET)[1] that reveals the policy decision. At 14:45 CET,[2] a press conference takes place in which the President and the Vice President of the ECB explain the monetary policy decision. The conference consists of a planned introductory statement and an ad hoc question and answer (Q&A) session with the press. The ECB's press conferences are live-streamed,

[1]prior to 21 July 2022, it was 13:45.
[2]prior to 21 July 2022, it was 14:30.

and market participants pay close attention to every utterance central bank officials make. While a vast body of literature empirically shows that the tone of these press conferences impacts financial markets, press releases have been widely omitted in the current research agenda. This is due to their brevity and purely documentary nature, but this has changed since 2016. This change is outlined in this thesis.

I aim to contribute to the literature in the following ways: (i) Press releases have increased in length and content, and using linguistic analytical tools for these releases reveals shifts in the ECB's communication strategy, a development that has gone largely unnoticed in previous literature. This evolution has induced notable reactions within foreign exchange (FX) markets. (ii) the relevance of press conferences has decreased when considering price changes in FX markets. (iii) polarized language leads to more trading volume within announcement windows. (iv) when interested in high-frequency reactions, researchers may use the open-source YouTube subtitles of press conferences.

The thesis is organized as follows. After the introduction, chapter Two summarizes concepts needed to comprehend the analysis. Although essential, readers who are well-acquainted with these theories may choose to skip this segment. The third chapter discusses relevant literature, while the fourth chapter details the data used. Chapter Five presents evidence that press releases have increased in length. Chapter Six will elucidate the statistical evidence for varying hypotheses, followed by chapter Seven, which concludes the study and gives a brief outlook for future research.

2. Methodology

This section discusses the methods employed. To provide a concrete example, I illustrate the methods through their application to press conferences. The literature quantifies central bank communication with different methods. This thesis reviews exclusively the methods of sentiment analysis. Various concepts essential for comprehending this approach are delineated in the following chapter. This chapter is mainly based on Silge & Robinson (2017). The statistical software R is used throughout the text.

2.1. Natural language processing

Information is an economic good and nowadays can be transmitted instantaneously via the internet or other media. Frequently, information is transmitted in large quantities in written form, presenting comprehension challenges for a single individual. Computers enable us to extract information for content analysis and visualize large amounts of text. This field of study is called Natural Language Processing (NLP). This section briefly introduces NLP concepts used in the following analyses. Only the relevant topics are presented. For a more complete discussion of NLP in a central bank context, readers should consult Bholat et al. (2015).

A corpus consists of a number of unique documents, each a series of words. All distinct words in a corpus make up its vocabulary. Words must then be converted into numerical values to make texts computationally interpretable. Different software systems use different methods to achieve this. A commonly used method is the *bag of words* representation in which words are represented in a vector space. Each word i in the vocabulary is given an individual numerical value based on how often it appears in the document. In this method, the structure of the text is lost, meaning that the order of the text does not matter anymore. The numerical vector of the frequency of the words is now a mathematical object which can be taken as input for different machine-learning algorithms. Another pre-processing technique is to break down texts into individual *tokens*. A token is a meaningful text unit, such as a single word. Tokens may also be a sentence, paragraph, or even a whole document. In this context, a single word is called a *unigram*. Two consecutive words are a bigram. This may be arbitrarily extended, as we can deal with a sequence of consecutive words of length n, resulting in an *n-gram*, where n represents any feasible natural number. Each token is stored in the row in which it appears. Therefore, the method is called the *one-token-per-row*.

Consider, for instance, the following sentence *The euro area economy may contract in the current quarter.* When omitting punctuation and setting all words to lowercase, applying the bag of words approach to the sentence results in a vector of length 9:

$$[2, 1, 1, 1, 1, 1, 1, 1, 1]$$

The 2 indicates that the word *the* appears twice in the sentence. Table 2.2 exhibits the example sentence. Tokenizing this sentence into single words means splitting it into: "The" "euro" ... "quarter" (Table 2.3). The second column inhabits the bigram representation.

Table 2.1.: Comparison

Table 2.2.: Bag-of-words

word	frequency
the	2
euro	1
area	1
economy	1
may	1
contract	1
in	1
current	1
quarter	1

Table 2.3.: One-token-per-row

row	unigram	bigram
1	the	the euro
2	euro	euro area
3	area	area economy
4	economy	economy may
5	may	may contract
6	contract	contract in
7	in	in the
8	the	the current
9	current	current quarter
10	quarter	

To make the evaluation less intricate, many other steps are taken. These steps aim to reduce the dimensionality of the text rather than the interpretability of the text. Following Bholat et al. (2015) or Denny & Spirling (2018) some important steps and concepts are:

- Strip out punctuation or special characters such as periods, extra white spaces, or numbers. In the example above, this has already been done.
- Making all letters in every word lowercase, as in our example, helps identify similar words. Typically, this does not change meaning, but a particular problem arises with names. Lowercasing the name "Elon Musk" to "elon musk," for example, may lead to a misinterpretation, as musk might be understood as a type of scent. This, however, is not a major concern for the ECB text corpus.
- Omit stop words. Stop words are commonly used English words, such as "the", yet do not contribute any significant analytical value and may hinder analysis.
- Stemming involves removing affixes from words and retaining only their most basic form (root or stem). A popular method is the Porter Algorithm (Porter, 1980). For

instance, the stem of the word "economics" is "econom". The same stem is shared by other words such as "economical", "economically", or "economic".

- Omit infrequently used terms as they only convey a small amount of information. A heuristic approach excludes terms that appear in less than 1% of all documents.
- Other methods, including lemmatization or case folding, were not used in this thesis.

According to Denny & Spirling (2018), altering the stop word list impacts the outcomes generated by completely automated analytical tools. However, the significance of data cleansing becomes diminished when dealing with semi-automated methods such as sentiment analysis. The task for researchers is now to extract information from corpora that is interpretable and usable in their respective fields of study. Many methods exist for this task. I will illustrate one example, the tf-idf method, and then continue with sentiment extraction in the next section.

A common way to measure the relevance of a word to a document is to consider the term frequency-inverse document frequency (tf-idf) heuristic. The tf-idf is a statistical value that reveals how important an n-gram is to a document within the corpus. To illustrate this, I took all introductory statements of the ECB press conferences from 2006 to 2022 and removed stop words, years, and common introductory sentences such as ”Ladies and gentlemen, the Vice-President and I are” and the names of the presidents and vice presidents. Figure 2.1 shows this specification’s ten bigrams with the highest tf-idf value. During Lagard’s mandate, cost-push shocks and the Generation EU project dominated. Draghi was focused on adjustments and loan sales, while Trichet’s results seem ordinary.

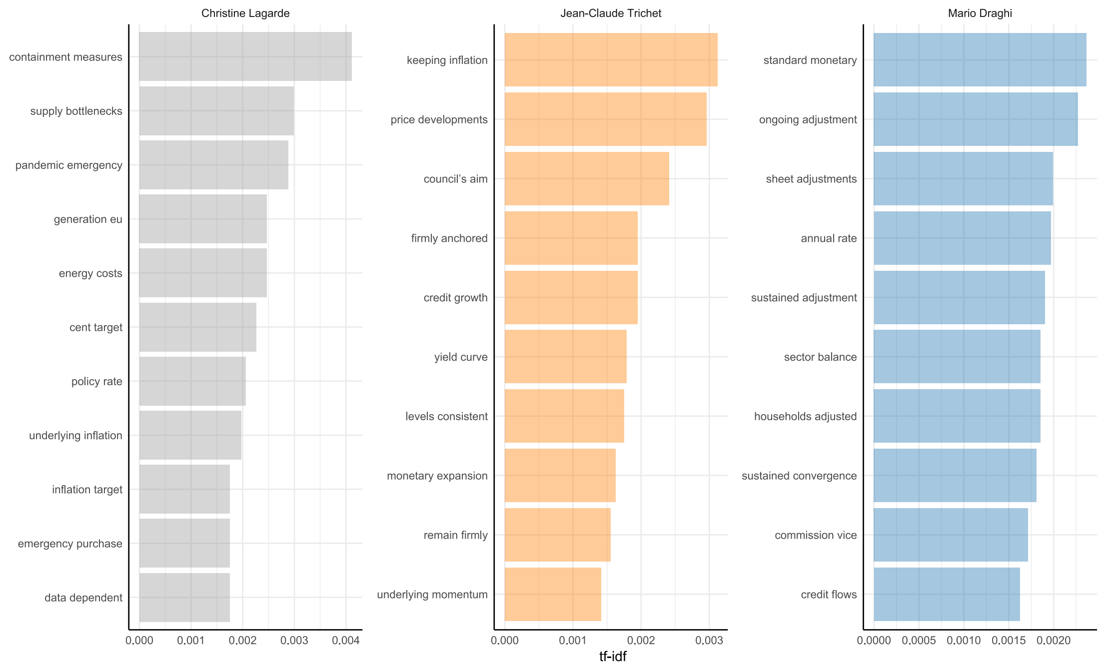

Figure 2.1.: The most important bigrams according to the tf-idf measure.

2.2. Sentiments

Algaba et al., 2020, (p.514) define sentiment as

> *"[...] the disposition of an entity toward an entity, expressed via a certain medium."*

This working definition is only one among many but is convenient for economists. It comprises three important features:

1. The expression happens through verbal and nonverbal communication through a medium. The ECB releases private information with subjective expressions through qualitative sentiment data such as textual, audio, and visual data in their press conference. Examples of sentiment in audio data include the tone of the speaker's voice, whereas sentiment in visual data might include facial expressions and bodily gestures. Most researchers focus on the sentiment transmitted through texts as crucial information. This analysis restricts itself to a dictionary-based sentiment analysis of written English text, dismissing any other form of sentiment.

2. The disposition is quantifiable on a discrete or a continuous scale. This measures the polarity of expressions conceivable at distinct levels of detail, such as single sentences or documents. The scale may be one-dimensional and only capture positive and negative sentiments, or it may be multi-dimensional. The one-dimensional case often has application-specific terminologies. In a central bank context, we may classify documents as dovish or hawkish. ECB press conferences may convey a pessimistic outlook by using negatively polarized words. The quantified sentiment is dovish, meaning the ECB will keep interest rates low to stimulate the economy. The hawkish sentiment is a more optimistic economic outlook and forecast in which the central bank is expected to increase interest rates because of greater inflationary pressure. Beyond the scope of economics multidimensional scales for text are employed. Mohammad & Turney (2013) created a dictionary to quantify text into different emotions, such as anger, trust, or joy.

3. The sentiment is expressed toward another entity. The ECB expresses its sentiment in its press conference about its current economic outlook to the public.

Much of the literature considers textual sentiment as critical text information. Three main techniques for identifying sentiment are manual, lexicon-based, and machine-learning approaches. This thesis solely concentrates on lexicon-based methods.

The lexicon-based approach is a straightforward, semi-automated method. The sentiment is attained through the appearance of important keywords in the text; therefore, a list of predetermined keywords with their corresponding positive or negative numeric sentiment scores is necessary. This list is called a *dictionary* or *lexicon*. First, the text and keywords are matched. Then, the frequency of the keyword and its numeric value is calculated, resulting in a sentiment score. Dictionaries may be general or domain-specific. The Henry lexicon (HENRY) (Henry, 2008) and the Loughran & McDonald lexicon (LM)

(Loughran & McDonald, 2011) are standard dictionaries used in the finance domain. This metric is often called the tone of documents in the relevant economic literature. Schmeling & Wagner (2015) in their influential paper *Does Central Bank Tone Move Asset Prices?* view tonality and a sentiment score measure with the LM dictionary as a synonym. They calculate the proportion of unfavorable terms to all the terms. This method tends to be unambiguous and transparent, making it suitable for any text. Nonetheless, lexicons can often be unreliable when confronted with domain shifts.

This thesis refers to the discussed measured score as the sentiment score or the sentiment of a document. The package SENTOMETRICS (Ardia et al., 2021) is used to perform sentiment analysis. Whenever the LM dictionary is used, the current form found on the complementary webpage,[1] is employed. For other dictionaries, default dictionaries provided by the package are used.

Here is an example to elucidate the method of sentiment calculation using the LM dictionary. The dictionary assigns a sentiment score by binary classification of unigrams as being either *negative* (-1) or *positive* ($+1$). Table 2.4 presents the first and last three words in the dictionary:

Table 2.4.: Vocabulary of the LM dictionary (4,150 unigrams)

Word	**Classification**	**Sentiment**
abandon	negative	-1
abandoned	negative	-1
abandoning	negative	-1
	...	
wrongful	negative	-1
wrongfully	negative	-1
wrongly	negative	-1

Consider the following sentence:

Critically, it can be a little bit contradictory, like forward guidance without the forward guidance.

This sentence carries complex information, but its tonality is assessed by only two words: *contradictory* and *critically*. Both have a negative connotation. The words *forward guidance* do not appear in the latest version of the LM dictionary. The simplest way is to calculate the sentiment as a weighted sum of all detected word scores within a document. We find two negative words; hence the weighted sum is $\frac{-1-1}{2} = -1$ for this sentence. Clearly, the score depends on the dictionary used. To illustrate this, four different lexicons were used for analysis to compare the sentiment scores of ECB introductory statements.

[1] Updated version of January 2022: https://sraf.nd.edu/loughranmcdonald-master-dictionary/

The sentiment scores are plotted in Figure 2.2. NRC (Mohammad & Turney, 2013) and SOCAL (Taboada et al., 2011) are non-financial lexica. Although the NRC captures similar movements as the financial lexicons, the SOCAL provides different values.

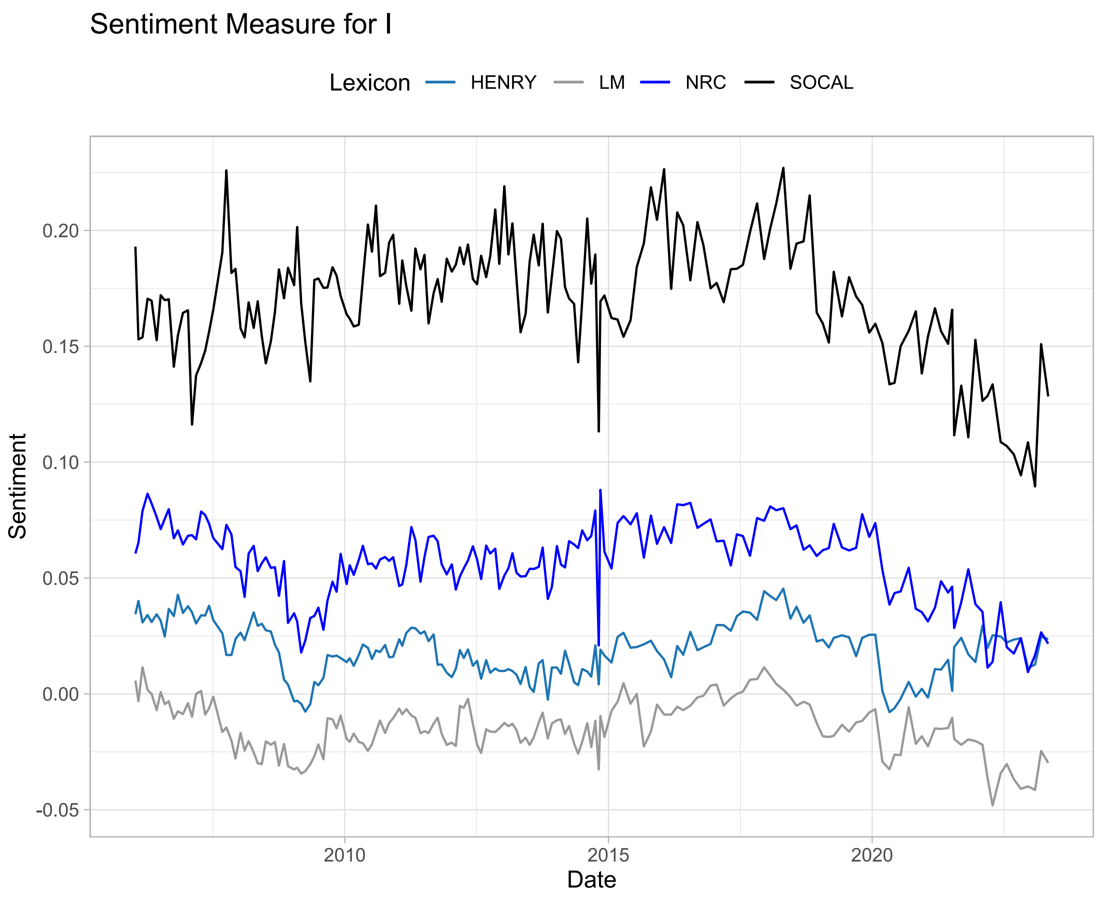

Figure 2.2.: Sentiment measure with different dictionaries of the introductory statement (I). The four dictionaries are: Henry (2008) (HENRY), Loughran & McDonald (2011) (LM) , Mohammad & Turney (2013) (NRC), and Taboada et al. (2011) (SOCAL).

An immediate drawback of this method appears with omitted negations or amplifications. To remedy this, one can include so-called valence-shifter words in the calculation. In this approach, bigrams are investigated. If a valence-shifting word appears before a keyword, the score of the keyword gets multiplied by the value of the valence shifter. In case we add a *not*, which has a value of -1, in front of the *contradictory* in our example sentence, the score changes to $-1 + (-1) \cdot (-1) = 0$. When adding the word *absolutely*, the score changes to $-1 + (-1) \cdot 1.8 = -2.8$. Table 2.5 displays the start and end of the list, where $\mathbf{y}$ shows the negation or amplification value.

Table 2.5.: Vocabulary of the valence shifters (131 unigrams)

Word	**y**	**t**
absolutely	1.8	2
acute	1.8	2
acutely	1.8	2
...		
wont	-1.0	1
wouldn't	-1.0	1
wouldnt	-1.0	1

Another layer of complexity can be implemented by clustering the words around a keyword. This examines at most four words before and two after the detected keyword. The score changes whenever the valence appears in this proximity to the keyword. Four types of valence shifters **t** are supported for the latter method, Table 2.5 shows the type of word i: negators (t = 1), amplifiers (t =2), and deamplifiers (t = 3). Adversative conjunctions (t = 4) are an additional type only picked up during a sentence-level calculation. Every dictionary used will incorporate this clustering on a sentence-level paradigm, ensuring sentiment measurement is carried out with the maximum granularity possible. To illustrate the case of the introductory statement, Figure 2.3 shows the three measures (simple, valence, cluster) for the LM sentiment score. Remarkably, these measures appear relatively indistinguishable. The ECB seemingly employs coherent language and seldom capitalizes on amplification or negation tactics.

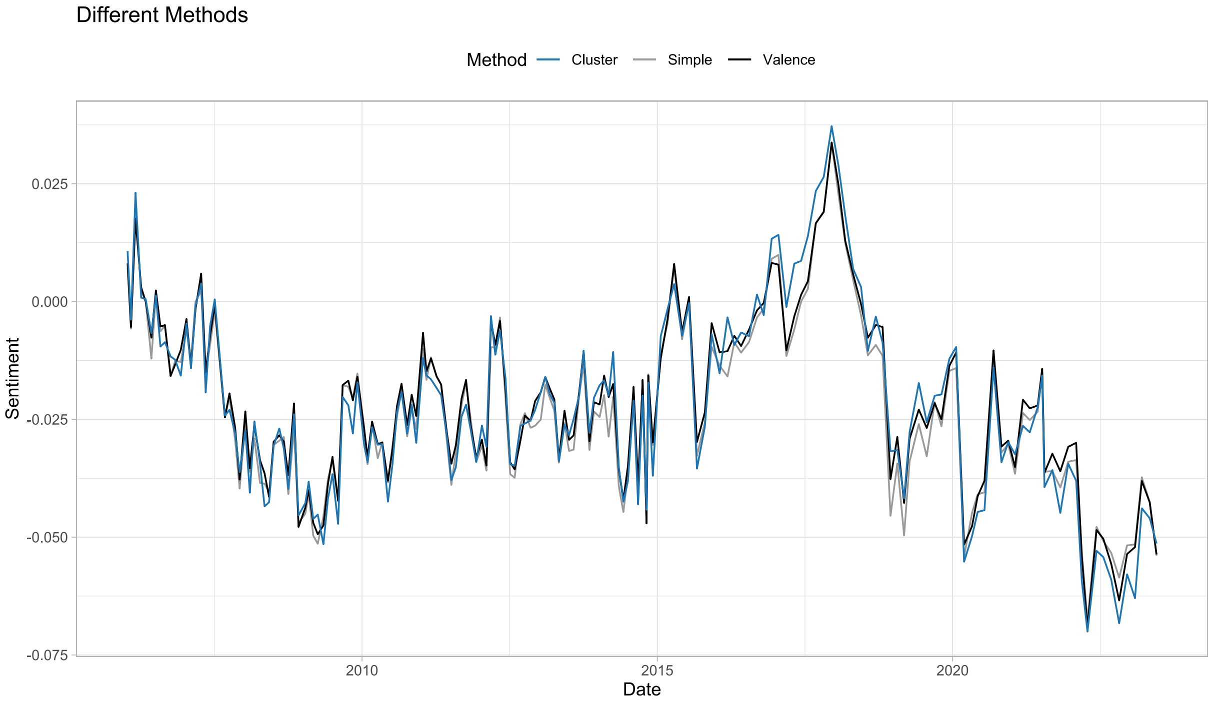

Figure 2.3.: The three different methods (simple, valence, cluster) applied to the introductory statements (I).

Nevertheless, some sentences are not captured due to insufficient signal words. Consider the following sentence:

The inflation outlook continues to be too high for too long.

Despite its apparent significance, this sentence does not receive a score in the LM lexicon, highlighting the limitations of this method. Picault & Renault (2017) clarify that a more extensive field-specific dictionary can enrich analyses. Their work analyzes the ECB introductory statements between January 2006 and December 2014. They transform the whole text into lowercase, use the Porter stemmer and remove stop words. They manually classify every sentence that appears, resulting in a total of 15,424. Then, for every n-gram that appears at least twice, they attribute one of two classes c. c consists of a monetary policy (MP) and economic outlook (EC) category and can be measured in three inclinations i: positive, neutral, and negative. Then, they count how often each n-gram appears in their sample and its associated inclination. They calculate a probability score that the n-gram is positive or negative in their sample.

$$P_n^{c,i} = \frac{numberofoccurrence_n^{c,i}}{totalnumberofoccurrence_n} \tag{2.1}$$

As an example, they show the bi-gram "consumption growth" which appears 22 times in their sample. It was always assigned to the category of economic outlook. Of these, it was used positively 20 (0.91) times and negatively 2 (0.09) times. This probability score was then taken to assess the tonality of the introductory statement s and is calculated in equation 2.2.

$$P_s^{c,i} = \frac{\sum_{n'=1}^{l} P_{n',s}^{c,i} * Occurrence_{n',s}}{\sum_{n'=1}^{l} P_{n',s}^{c} * Occurrence_{n',s}} \tag{2.2}$$

where $\mathrm{l} = 34,052$ is the number of n-grams in their field-specific lexicon. For each category $c = \{\mathrm{MP}, \mathrm{EC}\}$, the total tonality score adds up to 1; for example, $\sum_{c=1}^{3} P_s^{MP,c} = 1$ and $\sum_{c=1}^{3} P_s^{EC,c} = 1$. Only the most extensive n-gram was considered when multiple layered n-grams were encountered to improve the classification accuracy and avoid double-counting.

Finally, they aggregate the information content of the introductory statement into two indicators, defined by computing, for each statement s, the difference between the hawkish (positive) probability $P_s^{MP,hawk.}$ $\left(P_s^{EC,posi.}\right)$ and the dovish (negative) probability $P_s^{MP,dovi.}$ $\left(P_s^{EC,nega.}\right)$ as in equation 2.3. The two resulting variables are labeled I_s^{MP} and I_s^{EC}:

$$\left(I_s^{MP}, I_s^{EC}\right) = \begin{cases} I_s^{MP} = P_s^{MP,hawk.} - P_s^{MP,dovi.}, & I_s^{MP} \in [-1,1] \\ I_s^{EC} = P_s^{EC,posi.} - P_s^{EC,nega.}, & I_s^{EC} \in [-1,1] \end{cases} \tag{2.3}$$

The score of these indicators plays a central role in this thesis. The measurement in question is uniformly applied to press releases and Q&A sessions, underpinned by the unique lexicon tailored specifically to decipher ECB's communication. This implicitly presupposes a fair amount of constancy in terms of vocabulary and contexts across different documents. Whenever the tonality of a document is measured by the method of Picault & Renault (2017), the abbreviation PR is used. The study likewise adheres to their recommended acronyms for the introductory statement, with I^{MP} and I^{EC}. Similarly, for Q&A sessions, QA^{MP} and QA^{EC} are adopted; while monetary policy declaration press releases use MP^{MP} and MP^{EC}.

3. Related Literature

This review focuses on the literature examining ECB communication in general and the impact of such discourse on foreign exchange (FX) market dynamics. For reviews of other domains of central bank communication, interested readers should refer to Nassirtoussi et al. (2014), Gentzkow et al. (2019), or Algaba et al. (2020). Primary insights from this body of work suggest that trading activity and volatility increase up to approximately one hour after an ECB announcement, and unexpected announcements lead to higher volatility (Neely, 2011). Analysis without any textual interpretation might be regarded as an indirect approach, identifying financial market movements in a window around an event (Gürkaynak et al., 2004) or the ECB press conference itself (Brand et al., 2010; Altavilla et al., 2019; Jarociński & Karadi, 2020).

One strand of literature has attempted to identify textual information that moves markets, initially inspired by a *narrative approach* of monetary policy decisions (Romer & Romer, 1989, 2023). This was first achieved by manually encoding information (for the ECB, see Musard-Gies (2006); Rosa & Verga (2007); Heinemann & Ullrich (2007); Gerlach (2007); for the Federal Reserve (FED), see, e.g., Rosa (2013)).

Assigning sentiments to documents can be done purely manually. Musard-Gies (2006) categorizes ECB press conferences based on their perceived neutrality, hawkishness, or dovishness. The categorization stems from her interpretation of the ECBs statements and is, therefore, by definition, subjective and challenging to replicate. Rosa & Verga (2007) manually classify the introductory statements on a scale from -2 (very dovish) to 2 (very hawkish). In doing so, both authors disagree on 14 (22.58%) statements. Gerlach (2007) used this approach to classify the monthly ECB Bulletin. Literature has shifted to a lexicon-based method to implement objectivity partially.

To increase replicability, researchers focused on keyword detection, measured with different methods (see Section 2.2 for the specific method of measurement using one-dimensional textual sentiment). An early illustration of this type of keyword detection was put forward by Jansen & De Haan (2007), who counted the appearance of the word *vigilance* in the ECB communications and its effect on price stability. Lucca & Trebbi (2009) score Financial Open Market Committee (FOMC) statements on a scale from *hawkish* (+1) to *dawish* (-1) based on the Google semantic orientation score. This approach considers the semantic structure of statements and is based on a large information set obtained from search engines. Picault & Renault (2017) develop their own field-specific ECB dictionaries composed of n-grams (up to $n = 10$) to predict future monetary policy decisions and explain variation in the volatility and return in European stock markets.

Baranowski et al. (2021) adopt this strategy to predict future ECB decisions, including orthogonal tone shocks. They find that tone shocks significantly correlate to future ECB monetary policy decisions. Additionally, Picault et al. (2022) investigate Twitter data and use the data collected by Altavilla et al. (2019). They find that the press conference and intermeeting communication affect media sentiment.
Schmeling & Wagner (2015) detect negative words with the LM dictionary. They calculate how many negative keywords appear compared to the overall number of words. They find that a positive change in the sentiment correlates with increasing bond yields, lower implied equity volatility, lower variance risk premia, and lower credit spreads. Other examples that use the LM dictionary are Hubert & Fabien (2017) and Tillmann & Walter (2018). This approach is also used for central banks on continents other than North America or Europe (González & Tadle, 2020). Venturini (2020) augments the LM dictionary with additional dictionaries. His findings suggest a herding behavior in the European financial market after positive and negative spikes of the sentiment the ECB conveys. All the above research aligns with the research presented in this thesis, which presents the first press release analysis using the method of sentiment analysis with a dictionary approach.

Parallel to this development, researchers focusing on FX markets have begun to focus on the effects of oral interventions on euro exchange rates. It is hypothesized that these interventions possess a signaling effect on the exchange rate, which is viewed as an asset price under rational expectations. The authors are concerned with price levels and volatility and typically employ Generalised Autoregressive Conditional Heteroskedasticity (GARCH) models and extensions. Jansen & De Haan (2005) manually collected 930 unscheduled statements of central bankers in Europe and ECB officials and found that positive declarations had limited immediate influence on the level of exchange rates, instead increasing volatility. Fratzscher (2008) assesses oral interventions' effects on FX markets from 1990 to 2003. Statements of policymakers are manually classified as either positing a strengthened domestic currency (+1), a weakened exchange rate (−1), or ambiguous (0). The results conclude that oral interventions had a short-term effect and were a largely autonomous policy tool. Conrad & Lamla (2010) investigate the intraday seasonality of Thursdays. They quantified each sentence and compared absolute returns. Switching back to an OLS methodology and quantile regression Gertler & Horvath (2018) attempt to explain the price change of assets surrounding an unplanned announcement. The price change is the value of the financial asset 15 minutes before the communication event and 60 minutes after the event. The evidence suggests that central bank communication significantly affects interest rates and stock markets, with a lesser effect on exchange rates. Tadle (2022) uses his own dictionary and finds that hawkish FOMC minutes appreciate the dollar. By documenting a shift in intraday seasonality of announcements on Thursdays, this thesis will contribute to this strand of literature.

Most of the literature and more recent studies have largely focused on examining the market reactions to press conferences, especially introductory statements (Cour-Thimann, 2020; Cour-Thimann & Jung, 2021; Hayo et al., 2022). Anderes et al. (2021) analyze all press releases and do not focus exclusively on the monetary policy decision. An exception to the above is Parle (2022), who does consider the information from the announcement of the monetary policy decision. This thesis contributes to the literature by highlighting that the information in monetary policy press releases has become increasingly important since 2016.

Another interesting question pertains to how the market reacts in real-time to press conferences. Ehrmann & Fratzscher (2007) regressed 3-month Euribor futures returns on Reuters News *snaps*, which are real-time compressed descriptions of decisive points made during a press conference. Their results indicated that the measured and revealed information had a one-minute lag on absolute prices. Moreover, press conferences impacted the market response. Statements on inflation and interest rates had the highest leverage. Ulrich et al. (2019) built upon this idea by using Bloomberg newswire subtitles to conduct clustering of words to differentiate between topics. They examined the real-time effects of these topics on the returns of the Euro Stoxx 50 index futures and documented that different topics led to different trading activities. When analyzing live reactions to press conferences, a subscription to Bloomberg was necessary. In the present study, however, I document an open-source solution, thereby democratizing access to this data category and expanding the scope of researchers.

Finally, this thesis adds to a body of literature more closely related to Behavioral Finance. Motivated by Varian (1989), a particular strand of Behavioral Finance investigates market situations characterized by high trading volume coupled with low price changes. This phenomenon may be attributed to varying interpretations of information among agents when dealing with new information. Kandel & Pearson (1995) evaluated this hypothesis in a model that considered the different interpretations of public announcements made by corporations. They determined that relaxing the presumption of uniform information interpretation alters predictions concerning trade volume. Hong & Stein (2007) list multiple explanation mechanisms and argue that traded volume can be connected to investors' sentiments. In a central bank context, Smales & Apergis (2017) suggest that longer and more complex statements by the FED may result in more divergence of trader opinion and, therefore, higher trading volume. This results from either information asymmetry driven by knowledge of customer order flow or from a superior understanding of how asset prices will react to the information. Hayo et al. (2022) conducted a similar analysis for the ECB press conference and found that trading activity shifts to the Q&A session when statements are complex, resulting in a lower trading volume in the European stock index futures. These findings motivate the regression estimates in chapter Six. This thesis expands this research by finding that polarized language increases trading activity and that sentiment by the ECB leads to more trading volume on the FX markets. However, trading shifts resulting from more prolonged or polarized press releases at later times are not found.

4. Data

This chapter presents the different data sources, including the information obtained from the European Central Bank's website, data from YouTube subtitles, and an overview of the foreign exchange data.

4.1. Textual data

The relevant documents published by the ECB are monetary policy decisions and press conferences. The ECB provides transcripts of those documents on its website under two different links. One webpage contains *Press releases on Monetary policy*[1] and lists all press releases and monetary policy decisions. The other page of interest is *Press conferences*[2] which lists all transcripts of the press conferences.[3] Employing the package RVEST (Wickham, 2022) I created a corpus of all documents. Table 4.1 shows which documents have been gathered.

Table 4.1.: ECB data summary

Type of the statement	Number of occurrence	Percentage of occurrence
Press release	330	0.341
Monetary policy decision	294	0.304
Press conference	264	0.273
Other publication	46	0.048
Combined monetary policy decisions and statement	11	0.011
Monetary developments in the euro area	10	0.010
Monetary policy account	6	0.006
Euro money market	5	0.005
Annex	1	0.001
Eurep faq	1	0.001

This thesis is limited to the textual data of the monetary policy decisions and press conferences from 2006 to 2022, excluding other communication channels. This period is chosen for three reasons. Firstly, the extraction algorithm was unreliable for the earliest web pages due to inconsistent ECB page formatting. Secondly, results from Picault & Renault (2017) will be replicated. Their research involves the analysis of introductory

[1] https://www.ecb.europa.eu/press/pr/activities/mopo/html/index.en.html
[2] https://www.ecb.europa.eu/press/pressconf/html/index.en.html
[3] For the 27 of October 2022: https://www.ecb.europa.eu/press/pressconf/2022/html/ecb.is221027~358a06a35f.en.html

statements that originate from the year 2006. Thirdly, the FX data set began in 2009, so textual data prior to this date will not be used when working with the data. The analysis period covers the presidencies of Trichet, Draghi, and Lagarde, which are summarized in Table 4.2. A summary table of the corpus is presented in Table 4.3. In total, 173 press conferences and 169 press releases have been web-scraped.

Table 4.2.: Presidents of the European Central Bank

President	Took Office	Left Office	Days in Office
Wim Duisenberg	1 June 1998	31 October 2003	1739
Jean-Claude Trichet	1 November 2003	31 October 2011	3284
Mario Draghi	1 November 2011	31 October 2019	2922
Christine Lagarde	1 November 2019	Incumbent	-

Table 4.3.: Corpus summary

Document	**Result**
Number of press conferences	173
Total tokens	6,875,581
Total words	1,163,275
Total tokens I	1,788,174
Total words I	276,005
Total tokens Q&A	5,087,407
Total words Q&A	1,163,275
Number of monetary policy decisions	169
Total tokens MP	194,841
Total words MP	30,726
First observation	2006-01-12
Last observation	2022-12-15

The ECB publishes videos of its press conferences on YouTube, an online video-sharing and social media platform. YouTube provides transcripts for these videos, typically in the form of subtitles which include time stamps that indicate what has been said at each point. Transcripts can be manually uploaded to the videos or automatically created, which is the case for press conference videos. YouTube is opaque about how they create subtitles, but reviewing these transcripts makes it possible to track what has been said within each minute of the press conference. When transcripts are enabled for videos, the transcript section appears on the right-hand side or below of the video itself, and the time stamps are highlighted in blue. Table A.1 and Table A.2, presented in the appendix, document which videos have been used for the analysis. The dataset covers press conferences from 2016 to 2022. The press conference from 26 July 2018 has been omitted, as there are no subtitles available, and for the conference on 8 September 2022, only the introductory statement is available. Therefore these two dates are omitted in the further analyses. Ehrmann & Fratzscher (2007) attempted a similar approach, watching the taped versions of the press conference on Bloomberg from July 2001 to April 2006. They found that, on average, the press conference lasted 44 minutes, with 12 minutes dedicated to the introductory statement and 32 minutes allocated to the QA session. A more recent figure by Altavilla et al. (2019) or Ulrich et al. (2019) points out that it takes the president around 15 minutes to read the introductory statement and roughly 45 minutes to answer the follow-up questions.

4.2. Market data

The FX data comes from FirstRate Data,[4] and comprises equally-spaced 1-minute, 5-minute, 30-minute, and 1-hour intraday data of the ten most active FX crosses. It is sourced from major exchanges and four dark pools and tested for completeness and consistency. The time series starts in 2009 and ends on December 18^{th}, 2022, and contains the open, high, low, and close prices and the traded volume for each tick. The data contains the tick data for the following euro exchange rates: US-Dollar (USD), Japanese Yen (JPY), British Pound (GBP), and Swiss Franc (CHF). Table 4.4 displays summary statistics of the euro yen time series.

Table 4.4.: EUR/JPY 1 min data summary

	Min	1st Q	Median	Mean	3rd Q	Max	NAs	N
Open	94.115	118.720	127.267	125.235	132.730	149.755	0	$5,331,246$
High	94.149	118.738	127.286	125.250	132.745	149.776	0	$5,331,246$
Low	94.103	118.703	127.250	125.218	132.713	149.708	0	$5,331,246$
Close	94.114	118.721	127.268	125.235	132.730	149.752	0	$5,331,246$
Volume	1	74	149	207.119	267	$15,987$	0	$5,331,246$

The other summary statistics are not provided since the data comes from a single provider. Define the returns, r_t, as the logarithm of the relative price changes:

$$r_t = \ln\left(\frac{C_t}{C_{t-1}}\right). \tag{4.1}$$

In this thesis, returns are understood as price changes, and these terms are used interchangeably. Figure 4.1 depicts the hourly closing price C_t for the EUR/YEN currency pair for the whole-time sample. The plot below presents the respective hourly returns and marks the ten most significant price changes with blue dots. The appendix, specifically Table A.3, displays the dates and corresponding values for these items.

[4]FirstRate Data LLC. webpage: https://firstratedata.com/

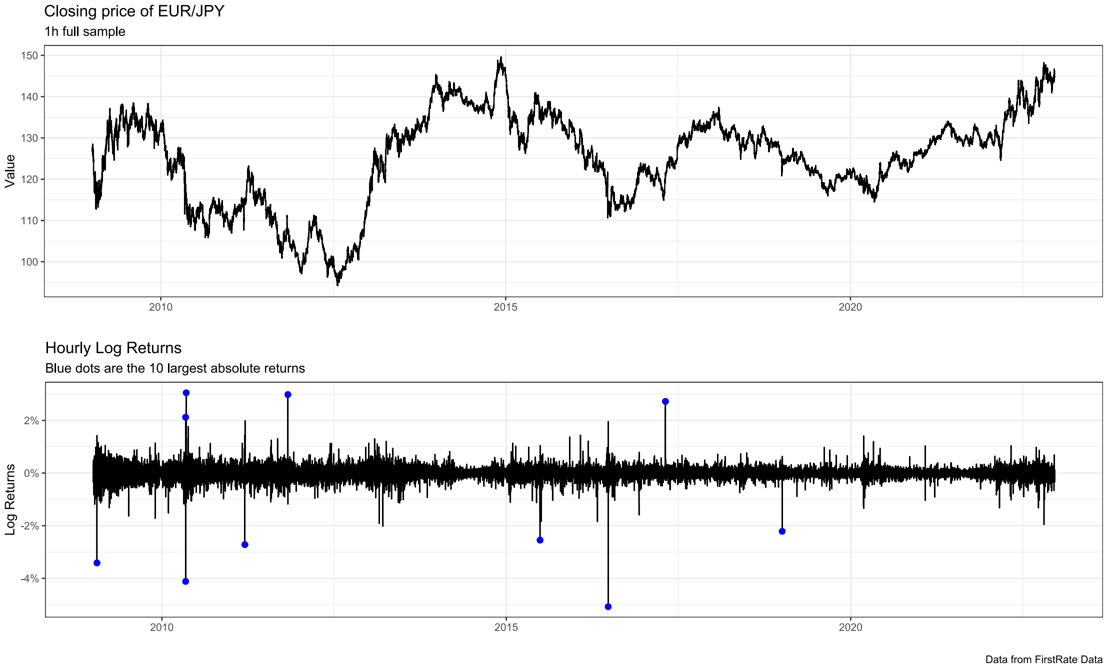

Figure 4.1.: 1h descriptives, list of the 10 largest absolute returns found in Table A.3 in the appendix, adopted from Perlin et al. (2020).

Additionally, the daily EURO STOXX 50 Volatility (VSTOXX) EUR Price Index and the change in the Main Refinancing Operations (MRO) of the ECB was obtained from the Refinitiv terminal of the University of Vienna, with the *high* prices being recorded. This time series is abbreviated as *VIX* and summarized in the appendix (see Table A.4).

Figure 4.2 plots the average traded volume of the EUR/JPY time series for each Thursday of the time series, as ECB decisions usually are released on Thursdays. Illustrated in black, the average traded volume per minute on days without a press conference is compared to the average traded volume per minute, denoted in blue, on press conference days. The grey dashed line depicts the point at which the monetary policy decision is released, and the red dashed line signifies the beginning of the press conference. As Conrad & Lamla (2010) point out, the trading day exhibits a noticeable intraday pattern, most significantly marked by an upsurge of traded volume at 14:00 CET when US markets open. Furthermore, the traded volume appears to diminish before ECB announcements and experiences a notable spike at their release. Traded volume is further accentuated at the onset of the press conference, sustains its elevated level afterward, and returns to its typical values post-decision in line with non-announcement days. Kerssenfischer (2019) makes a similar observation with the 2-year German Bonds and Euro STOXX 50 trading activity. This form of presenting trading volume is extensively employed in the following chapter.

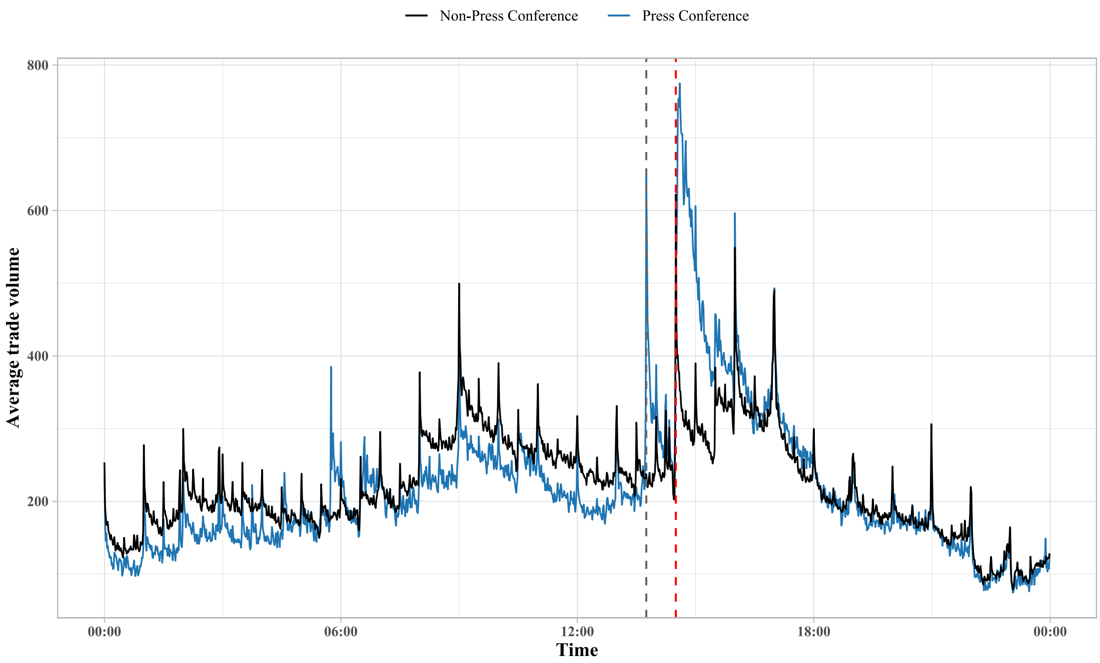

Figure 4.2.: The intraday seasonal pattern of the average traded volume of the EUR/YEN pair on Thursdays. In black are non-press conference Thursdays, and in blue are Thursdays when a press conference occurred.

5. Shift in the ECB Communication

Since 2016, the monetary policy decisions (MP) have increased in length, contain more content, and convey more significant meaning. Most of the literature refers to these decisions as press releases. Figure 5.1 shows the aggregated number of words for each year for the two types of documents. The supplied bar chart splits the press conference into the introductory statement (I) and the Q&A session. As of January 2015, the frequency of monetary policy meetings has moved to a six-week cycle, visible in the sharp drop during that year. Even though the number of decisions has decreased, the words transmitted through the press release have slowly risen.

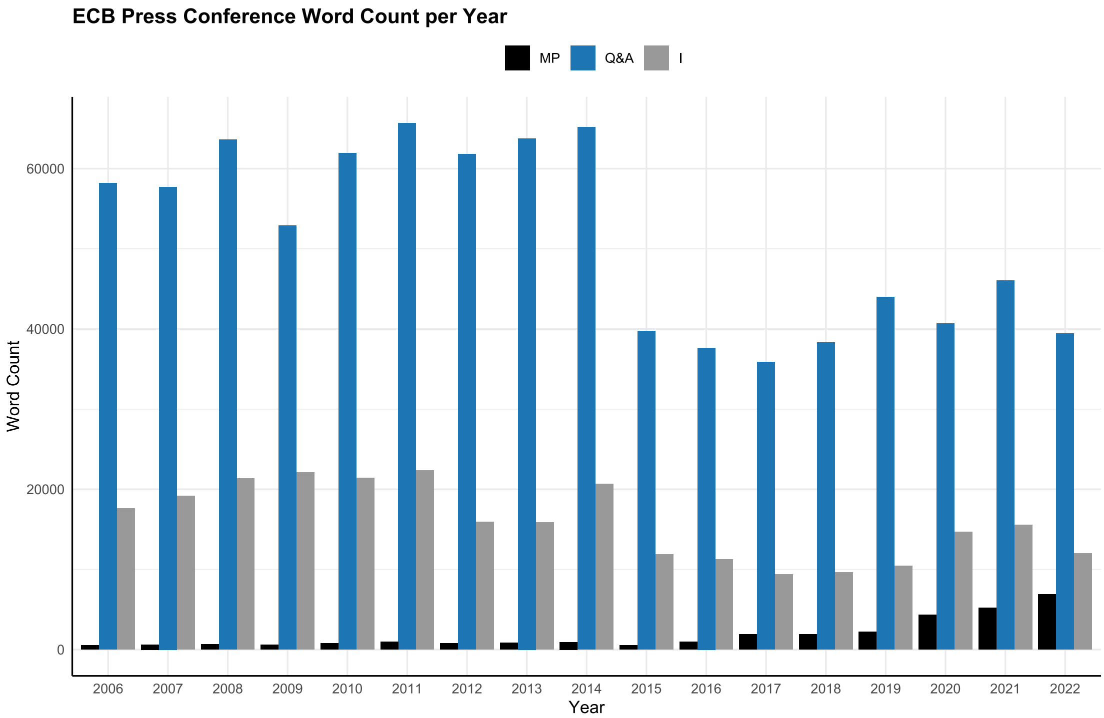

Figure 5.1.: Aggregated word count of the monetary policy release (MP) and the press conference (I and Q&A).

A comparison of the length of individual documents is shown in Figure 5.2. The MP documents experienced a marked increase in length around 2016, with the red dashed lines denoting changes in the presidency. The press conference is yet again broken down into introductory and Q&A parts. The Q&A session in light blue is the longest part of the conference, while the light grey introductory statements are shorter. There is a visible drop in the length of the introductory session for Draghi's Presidency. As research on releases is scarce, the length analysis of the documents provides a novel insight into the ECB's communication strategy.

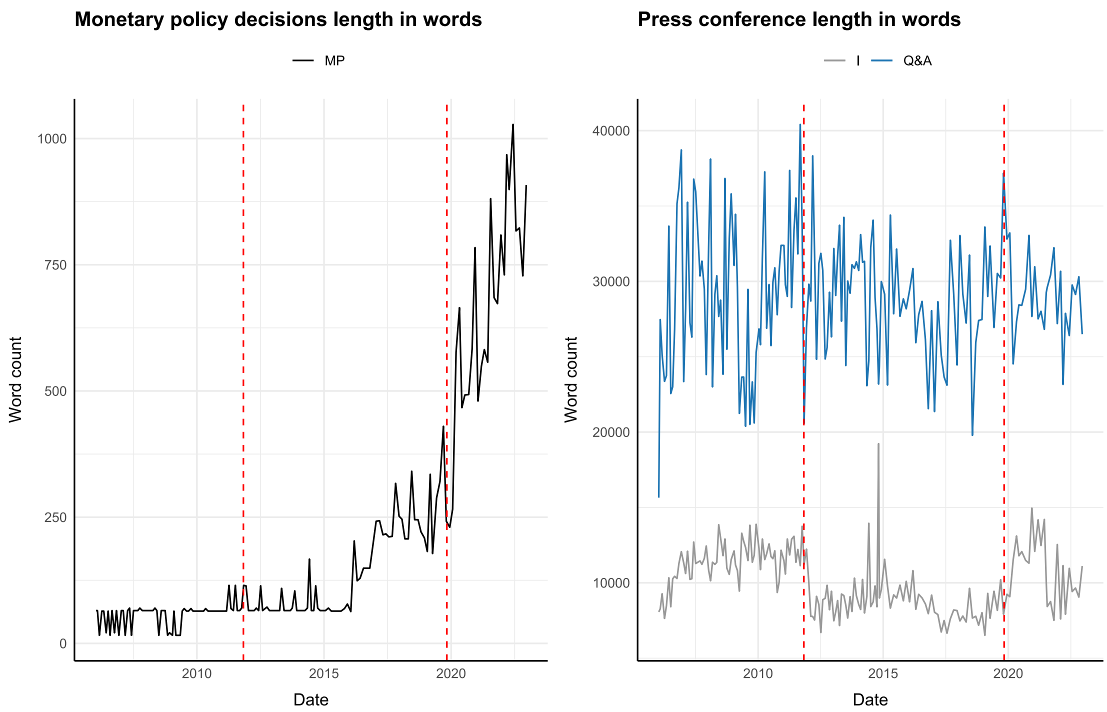

Figure 5.2.: Word count of the monetary policy decision (MP) in black and the press conference (I and Q&A). Red dashed lines indicating the point when the presidency changed.

Table 5.1 provides the results of linear regressions to obtain statistical evidence for this intuition. The number of words in each document is regressed against a year and presidency dummy variable. The data suggests that press conferences conducted by Draghi were shorter when controlling for years, primarily due to a shorter Q&A session. It further reveals an increase in the length of press releases starting in 2017 compared to 2006. However, this change does not necessarily imply that the press releases provide more relevant information. It could suggest that other measures were introduced as well. To investigate this issue further, consider the textual sentiment of the press releases measured by the LM dictionary and the Picault & Renault (2017) lexicon. Figure 5.3 visualizes the three different measures. The red dashed line marks the start of 2016.

With this visualization, it becomes apparent that sentiment variability has increased parallel to press release length. Yet, initial fluctuations failed to be detected by the LM dictionary, thus justifying the necessity to employ PR's measure as a more accurate tool.

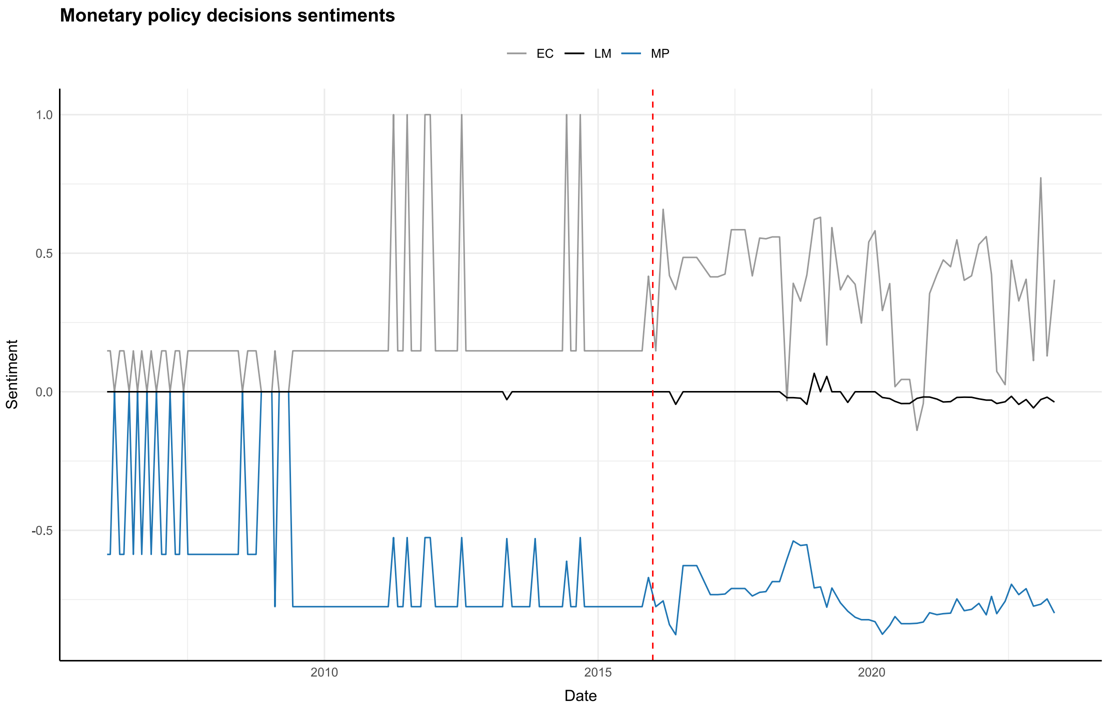

Figure 5.3.: Sentiment scores calculated with the LM (Loughran & McDonald, 2011) and PR lexicon (Picault & Renault, 2017). Red dashed line indicates the beginning of the year 2016.

Table 5.1.: Estimate of the length of the documents

	Dependent variable:			
	Total press conference	Introductory statement (I)	Q&A session	Press release (MP)
	(1)	(2)	(3)	(4)
base Trichet				
Lagarde	-1,119.871	64.200	-1,184.071	-13.386
	(958.130)	(319.123)	(945.177)	(80.677)
Draghi	-1,642.300**	-44.800	-1,597.500**	38.900
	(562.172)	(187.242)	(554.572)	(47.336)
base 2006				
2007	669.750*	273.447**	396.303	11.417
	(302.950)	(100.903)	(298.854)	(24.948)
2008	762.750*	312.000**	450.750	29.064
	(296.291)	(98.685)	(292.285)	(24.464)
2009	-66.667	375.250***	-441.917	3.667
	(296.291)	(98.685)	(292.285)	(24.948)
2010	627.667*	316.417**	311.250	19.250
	(296.291)	(98.685)	(292.285)	(24.948)
2011	1,290.550***	403.383***	887.167**	30.433
	(310.752)	(103.502)	(306.551)	(26.166)
2012	1,801.883**	-95.200	1,897.083**	-13.521
	(635.473)	(211.656)	(626.882)	(53.772)
2013	1,959.217**	-101.367	2,060.583**	-11.733
	(635.473)	(211.656)	(626.882)	(53.508)
2014	1,926.742**	165.960	1,760.782**	-5.650
	(632.810)	(210.769)	(624.256)	(53.508)
2015	1,777.300**	65.508	1,711.792**	-16.817
	(652.512)	(217.331)	(643.691)	(54.943)
2016	1,438.925*	-13.617	1,452.542*	53.933
	(652.512)	(217.331)	(643.691)	(55.547)
2017	981.800	-252.117	1,233.917	154.558**
	(652.512)	(217.331)	(643.691)	(54.943)
2018	1,314.925*	-216.742	1,531.667*	156.058**
	(652.512)	(217.331)	(643.691)	(54.943)
2019	2,061.621**	-129.117	2,190.738***	198.219***
	(659.681)	(219.719)	(650.762)	(55.547)
2020	1,721.246	304.508	1,416.738	509.219***
	(1,013.779)	(337.658)	(1,000.074)	(85.363)
2021	1,648.733	198.106	1,450.627	620.219***
	(1,010.165)	(336.454)	(996.508)	(85.363)
2022	1,226.996	-33.117	1,260.113	830.844***
	(1,013.779)	(337.658)	(1,000.074)	(85.363)
Constant	6,318.250***	1,466.917***	4,851.333***	45.167*
	(209.509)	(69.781)	(206.677)	(17.641)
Observations	173	173	173	171
R^2	0.294	0.497	0.203	0.935
Adjusted R^2	0.211	0.439	0.110	0.928

Note: $^{*}p<0.05$; $^{**}p<0.01$; $^{***}p<0.001$

Contemporary research examining the influence of textual sentiment on financial markets tends to focus solely on introductory statements. This could be problematic, as the press release provides comparable sentiment movements. Figure 5.3 illustrates the MP scores of the press release and introductory statement, rendered in blue and black, respectively. The red dashed line marks the start of 2016. Beginning this year, both sentiment scores move in parallel. Thus, any study attempting to ascertain the effect of the press conference without incorporating the prior press release may be prone to overestimate the significance of the introductory statement. Presented in Table 5.2 is a Pearson correlation between the sentiment score of the press release and that of the introductory statement from 2006 through 2022. The results show a significant correlation between the two measures for the entire period using the MP score but not the EC score. Between 2016 and 2022, the correlation between the EC scores also became significant. This suggests that the press release already states the information revealed during the press conference and has started to become more similar to the introductory statement.

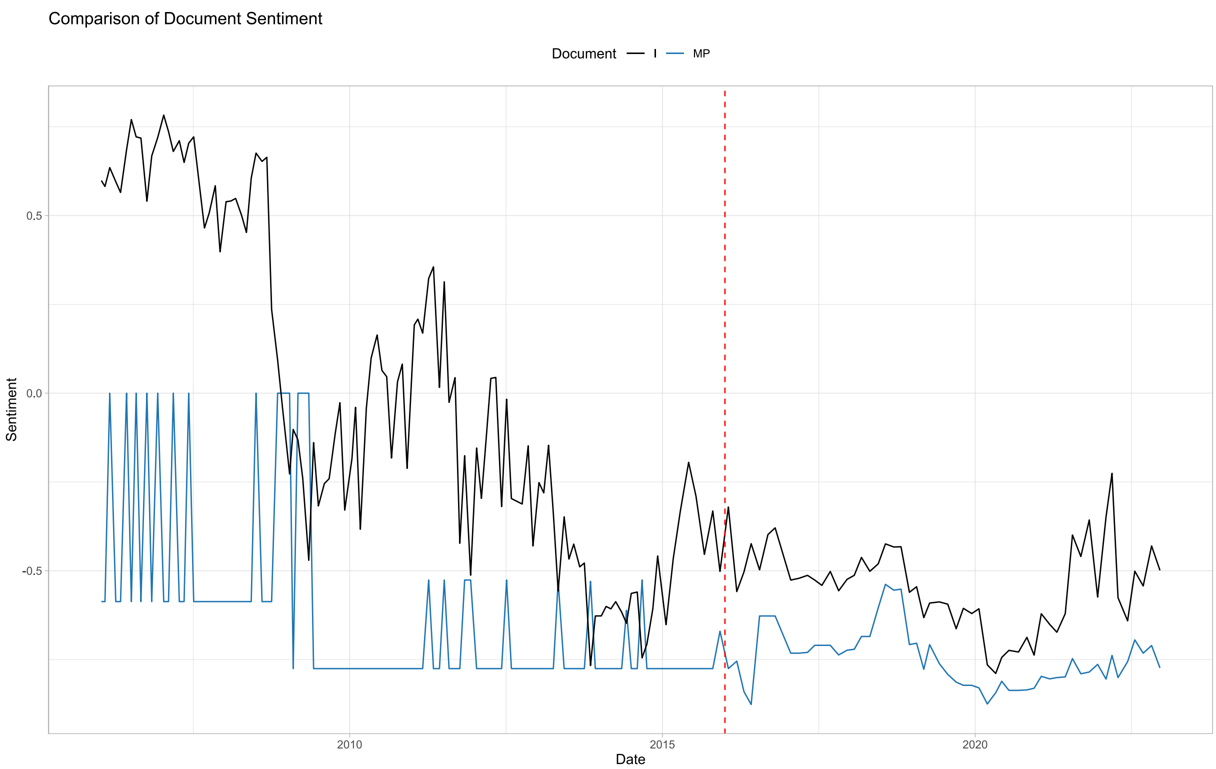

Figure 5.4.: Sentiment scores calculated with the PR lexicon (Picault & Renault, 2017). Red dashed line indicates the beginning of the year 2016. In black the sentiment score of the monetary policy decision (MP) and in blue the sentiment score of the introductory statement (I).

Table 5.2.: Pearson correlation

Pearson correlation table 2006-2022

	MP^{MP}	MP^{EC}	I^{MP}
MP^{MP}			
MP^{EC}	-0.18*		
I^{MP}	0.50***	-0.31***	
I^{EC}	0.11	0.08	0.32***

Pearson correlation table 2016-2022

	MP^{MP}	MP^{EC}	I^{MP}
MP^{MP}			
MP^{EC}	0.27*		
I^{MP}	0.53***	0.36**	
I^{EC}	0.62***	0.54***	0.49***

Topic models might be an attractive tool to identify changes in the discussed topics of the press releases. Topics are defined as a collection of similar words appearing in texts. Determining the topics provides information on otherwise hidden context and structure. Generally, the ECB press conference has been studied with various models. In these cases, either solely the introductory statement is examined (Klejdysz, 2018), the introductory and Q&A sections get dissected (Pavelkova, 2022), or the entire speech is studied (Cross & Greene, 2020). Similarly, a corpus of all documents published by the ECB has been used (Priola et al., 2021), though this does not focus exclusively on press release changes. Before 2016, press releases' lack of text, coherence, structure, and vocabulary rendered topic modeling infeasible. Topic modeling has the potential to offer unprecedented insights. Since this thesis does not introduce any topic modeling, it highlights some stylized facts about the press release. Before 2016, ECB press releases generally followed a simple structure: announcing a new policy decision in one to three sentences, followed by a sentence stating when the press conference would occur. The following press release from December 4^{th}, 2014, (ECB, 2014) says:

> At today's meeting the Governing Council of the ECB decided that the interest rate on the main refinancing operations and the interest rates on the marginal lending facility and the deposit facility will remain unchanged at 0.05%, 0.30% and -0.20% respectively.
> The President of the ECB will comment on the considerations underlying these decisions at a press conference starting at 2.30 p.m. CET today.

The press release on March 10^{th}, 2016 (ECB, 2016) modified this simple formula. It still contains the original sentences but adds forward guidance on the asset purchase and refinancing programs.

> At today's meeting the Governing Council of the ECB took the following monetary policy decisions:
>
> 1. The interest rate on the main refinancing operations of the Eurosystem will be decreased by 5 basis points to 0.00%, starting from the operation to be settled on 16 March 2016.
> 2. The interest rate on the marginal lending facility will be decreased by 5 basis points to 0.25%, with effect from 16 March 2016.
> 3. The interest rate on the deposit facility will be decreased by 10 basis points to -0.40%, with effect from 16 March 2016.
> 4. The monthly purchases under the asset purchase programme will be expanded to €80 billion starting in April.
> 5. Investment grade euro-denominated bonds issued by non-bank corporations established in the euro area will be included in the list of assets that are eligible for regular purchases.
> 6. A new series of four targeted longer-term refinancing operations (TLTRO II), each with a maturity of four years, will be launched, starting in June 2016. Borrowing conditions in these operations can be as low as the interest rate on the deposit facility.
>
> The President of the ECB will comment on the considerations underlying these decisions at a press conference starting at 14:30 CET today.

The releases afterward are organized in paragraphs and gradually given more content. Consider the following sentence from March 9^{th}, 2017 (ECB, 2017):

> If the outlook becomes less favourable, or if financial conditions become inconsistent with further progress towards a sustained adjustment in the path of inflation, the Governing Council stands ready to increase the programme in terms of size and/or duration.

More and more decisions concerning unconventional monetary policy begin to appear in the press releases. Studies that do not consider this break in communication might suffer a structural break.

Including the press release could have already been relevant to the literature. Table 5.3 presents a replication of the results found in Picault & Renault (2017) (table 7 specification 1 in their paper).[1] Since this thesis's text corpus includes other documents, it is possible to enrich their analysis. Specifications (1) to (5) take the time frame of the original publication. The findings confirm the literature. The introductory statement reveals the most information, and the Q&A session does not add more relevant information (Ehrmann & Fratzscher, 2007). A new result in this context is that the monetary press

[1] Unfortunately, the exact point estimates are not found. The surprise variable and market returns are taken from the package SENTOPICS. After public and private communication with the package's author, Olivier Delmarcelle, the results are still included since they convey the same interpretation.

release already had information pertinent to the markets as seen in the specification (4), even showing a greater R^2 than when only using the Q&A content (2). However, the introductory statement provided the most critical information to explain the returns. Specifications (6) and (7) restrict the data on Trichet's presidency. Within this period, the economic outlook of the introductory statement was the only relevant variable. Specifications (8) and (9) show the results for Draghi's term in office. The monetary policy indicator is more critical during this period. The presidency is relevant for the analyses done by Picault & Renault (2017). Additionally, the MP release also adds explanatory power to their analyses.

Table 5.3.: Replication of Picaul Renault

	Dependent variable:								
	R_t								
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
$Surprise_t$	0.043	0.046*	0.039	0.054*	0.045*	0.028	0.024	0.032	0.047
	(0.022)	(0.022)	(0.022)	(0.022)	(0.023)	(0.027)	(0.029)	(0.038)	(0.037)
R_{t-1}	0.215	0.247*	0.221	0.263*	0.237*	0.168	0.176	0.592*	0.383
	(0.114)	(0.116)	(0.114)	(0.114)	(0.114)	(0.137)	(0.138)	(0.267)	(0.253)
I^{MP}	-0.012**		-0.019**		-0.016*	-0.022	-0.022	-0.036*	-0.048**
	(0.005)		(0.007)		(0.007)	(0.011)	(0.011)	(0.016)	(0.016)
I^{EC}	0.012		0.012		0.013	0.027*	0.028*	-0.023	-0.031*
	(0.007)		(0.008)		(0.008)	(0.013)	(0.013)	(0.014)	(0.014)
QA^{MP}		-0.005	0.014		0.014	0.004	0.002	0.028	0.028*
		(0.006)	(0.011)		(0.011)	(0.015)	(0.016)	(0.015)	(0.014)
QA^{EC}		-0.009	-0.011		-0.009	-0.010	-0.005	-0.023	-0.010
		(0.013)	(0.013)		(0.013)	(0.020)	(0.021)	(0.016)	(0.016)
MP^{MP}				0.015*	0.011		0.009		-0.001
				(0.006)	(0.007)		(0.014)		(0.010)
MP^{EC}				0.0003	0.002		0.007		-0.049*
				(0.007)	(0.007)		(0.009)		(0.021)
Constant	-0.0001	-0.002	0.001	0.008	0.007	0.001	0.005	-0.010	-0.011
	(0.002)	(0.002)	(0.003)	(0.004)	(0.005)	(0.004)	(0.011)	(0.007)	(0.009)
Observations	107	107	107	107	107	69	69	38	38
R^2	0.139	0.093	0.157	0.124	0.175	0.149	0.168	0.370	0.512
Adjusted R^2	0.106	0.058	0.106	0.090	0.108	0.066	0.057	0.248	0.377

Note: $^{*}p<0.05$; $^{**}p<0.01$; $^{***}p<0.001$
Notice that I report sig. levels other than Picault and Renault.

Markets should react to this informational increase. The following figures provide visual evidence that, in recent years, the market response to press conference days has changed. Starting with intraday trading patterns of the EUR/JPY currency pair for 2009−2015 in Figure 5.5. The black line shows the average traded volume per minute on 288 non-press conference days. Meanwhile, the blue line denotes the average traded volume per minute on 75 press conference days. Seven press conferences occurred on another weekday than Thursday. The figure below shows the average absolute returns, calculated similarly. As anticipated in Chapter 4, a distinct trading pattern can be observed; however, only 11:00 to 17:00 CET is included here. Trading activity before the press release was lower than on normal Thursdays, which could be attributed to institutional regulations. However, upon the release of the announcement, there was a sharp surge in trading activity. Following this, trading activity rapidly returned to its pre-press release level. Simultaneously, trading activity during the press conference was higher than usual and remained at an elevated level for approximately an hour, consistent with the duration of the press conference.

However, absolute price changes indicated an inverse pattern, remaining similar to ordinary Thursdays until approximately 13:30 CET. After the press release, immediate price adjustments were made with low volume. The price adjustments were more prominent than during the press conference but with higher trading volume. The price adjustments were maintained over a longer period. This is in line with the efficient market hypothesis. The press release provides information for the fundamental value, most of which is priced in.

Figure 5.6 displays the trading activity of the EUR/JPY currency pair between 2016 and 2022, excluding Thursdays past July 2022 due to the altered release timing. This results in average data from 52 press conferences and 287 non-press conference Thursdays. Trading volume is higher on press conference days compared to 2009−2015, peaking at the time of the press release and remaining at a higher intensity over the following period. It is hypothesized that the more comprehensive nature of the press release requires a more extended period for market participants to comprehend the release. This generates extended trading activity and the consequent secondary price adjustment. Intriguingly, the data hints at opposite price adjustments for the press conference: trading activity remains higher than during the press release, although the price adjustment is lower. An identical behavior can be seen in other currency pairs. The EUR/USD dollar pair shows an even more significant change. In the appendix, the same plots are presented for the other currency pairs: EUR/USD in Figure A.1 and Figure A.2, EUR/GBP in Figure A.3 and Figure A.4, and EUR/CHF in Figure A.5 and Figure A.6. All point to the same conclusion.

In the following section, it will be tested whether these shifts depend upon the decision's length or the release's sentiment.

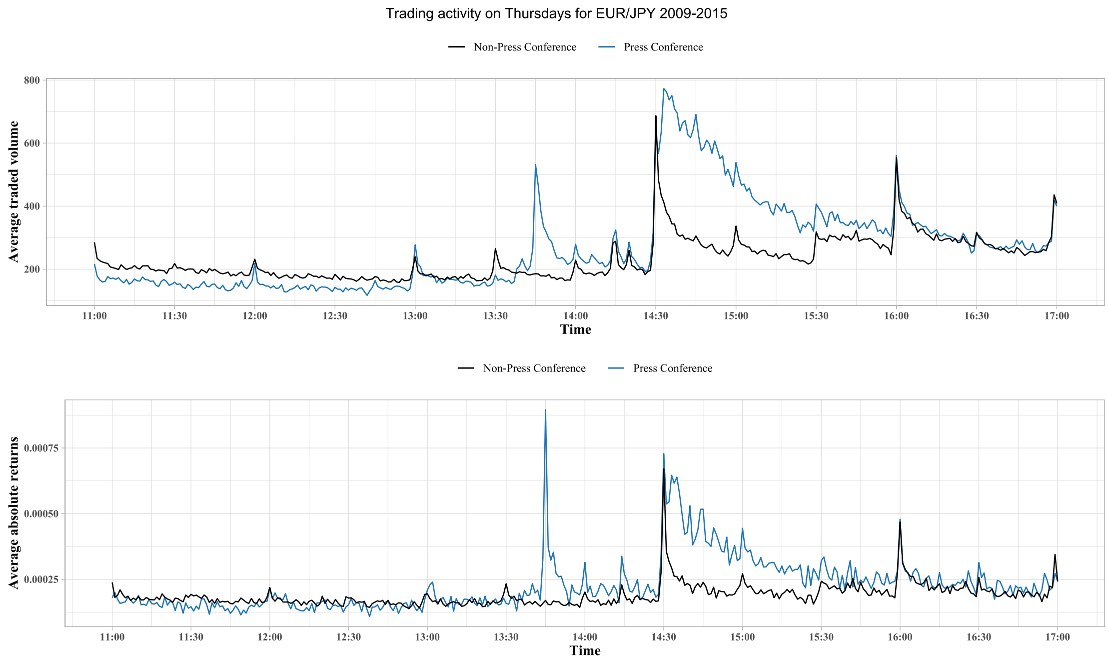

Figure 5.5.: Average traded volume and average absolute returns EUR/JPY 2009−2015.

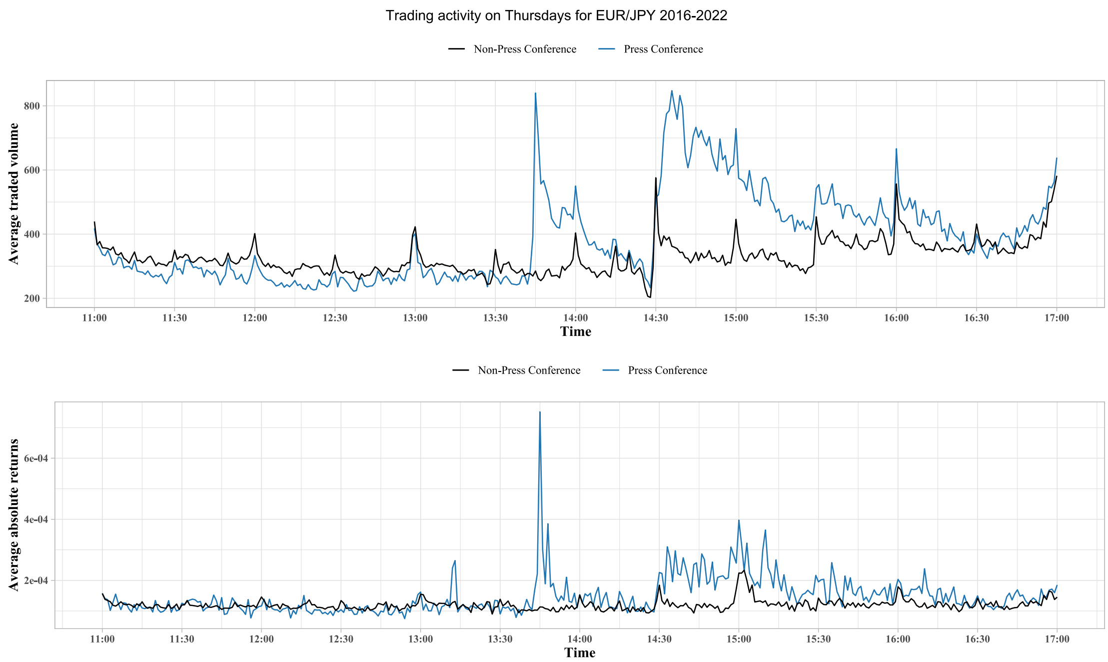

Figure 5.6.: Average traded volume and average absolute returns EUR/JPY 2016−2022.

6. The Impact on Exchange Rates: Hypotheses and Results

This section outlines the estimation strategy used to examine the assertions in the preceding chapters. Initially, a basic regression model incorporating a dummy variable for the period after 2015 is employed. Subsequently, sentiment scores are incorporated into the model. The results of the estimations will be provided for clarity and comprehensive understanding.

6.1. Event-based Regression Analysis

I analyze the impact of communication through event-based regressions that consider only press conference days. Although these regressions are not micro-founded and are ad-hoc, they provide empirical evidence regarding the subject matter. All time references will be denoted in Central European Time (CET). I adopt the event window framework proposed by Altavilla et al. (2019) to evaluate price changes and trading activity surrounding announcements. This framework divides the analysis into three windows: (i) The Policy Decision Window, constructed around 13:45 when the press release is published. (ii) the Press Conference Window, spanning from 14:30 to 15:30; and (iii) the Total Monetary Event Window (13:25−15:50). To assess the impact of the policy decision, I subtract the median closing prices during the pre-release window (13:25-13:35) and post-release window (14:00-14:10). The same is done with the press conference (pre-conference window 14:15−14:25 and post-conference window 15:40−15:50). Lastly, I measure the total monetary event effect by comparing median closing prices in the pre-release and post-conference windows. For a visual representation of the timeline and windows discussed above, refer to Figure 6.1, as taken from Altavilla et al. (2019). Let $R_{w,t}$ denote the return change for the currency under examination, then:

$$R_{w,t} = C_{w,t+1} - C_{w,t-1} \tag{6.1}$$

where t is the announcement date, and w is the respective window under consideration. I examine the average traded volume per minute within these windows to enhance this analysis further. In reference to the initial window, this establishes a timeframe from 13:25 to 14:10. This chosen timeframe is supported by empirical observations showing a substantial concentration of trading volume during this period (refer to Figure 4.2). The ending and starting points of the other windows are taken identically. Let $v_{w,t}$ denote the logarithm of the average traded volume within each window:

$$V_{w,t} = \log\left(\frac{\sum_{i=0}^{m} V_{w,t,m}}{m}\right) \tag{6.2}$$

where m denotes the length of window w in minutes.

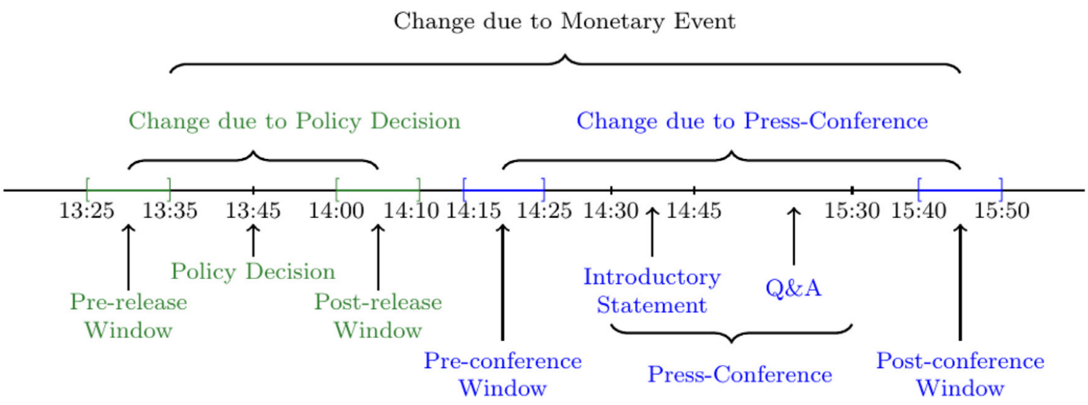

Figure 6.1.: ECB policy communication timeline, taken from Altavilla et al., 2019, (p.165).

The initial hypothesis associate lowers absolute price fluctuations in the press conference with more comprehensive press releases since 2016. Using regression analysis, I analyze the absolute price changes associated with press releases and press conferences in the pre-and post-2016 periods. The time frame for this study is limited to Thursdays from 2009 to July 2022 and yields 127 regular press conference days. To mitigate concerns about potential confounding factors arising from global market stress, the preceding day's volatility index (VIX) is considered. Furthermore, any absolute change in the ECB's main refinancing rate ($|\Delta MRO|$) is included to account for fluctuations in interest rates. Heteroscedastic Robust Standard Errors are employed as typical within the literature (Picault & Renault, 2017; Picault et al., 2022; Tadle, 2022). The SANDWICH package (Zeileis et al., 2019) is used, following the H0 specification of the standard errors. The following regression model is estimated:

$$|R_{i,w,t}| = \alpha + \beta_1 * D_{>2015} + \beta_2 * VIX_{t-1} + \beta_3 * |\Delta MRO|_t + \epsilon_t \tag{6.3}$$

where $|R_{i,w,t}|$ is the absolute price change for currency i during the event window w. $D_{>2015}$ is a dummy variable that takes a value of 1 if the date is after 2015. This equation is estimated separately for the press release and press conference windows. I present the estimation results of equation 6.3 in the forthcoming Table 6.1.

Table 6.1.: Estimation of change in returns for different currencies

	Dependent variable:							
	$\lvert R\rvert_{MP}$				$\lvert R\rvert_{P}$			
	$\lvert R\rvert_{YEN,MP,t}$	$\lvert R\rvert_{USD,MP,t}$	$\lvert R\rvert_{GBP,MP,t}$	$\lvert R\rvert_{CHF,MP,t}$	$\lvert R\rvert_{YEN,P,t}$	$\lvert R\rvert_{USD,P,t}$	$\lvert R\rvert_{GBP,P,t}$	$\lvert R\rvert_{CHF,P,t}$
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
$D_{>2015}$	−0.04	−0.001	−0.0000	−0.0001	−0.27***	−0.003***	−0.001**	−0.001
	(0.04)	(0.0004)	(0.0003)	(0.0001)	(0.08)	(0.001)	(0.0004)	(0.0003)
VIX_{t-1}	−0.002	−0.0000	0.0000	0.0000	−0.002	0.0000	0.0000	0.0000*
	(0.003)	(0.0000)	(0.0000)	(0.0000)	(0.004)	(0.0000)	(0.0000)	(0.0000)
$\lvert\Delta$ MRO$\rvert$	1.13**	0.01**	0.01*	0.003	−0.12	−0.004	0.0004	−0.002
	(0.40)	(0.004)	(0.002)	(0.001)	(0.49)	(0.004)	(0.002)	(0.002)
Constant	0.20*	0.002*	0.001	0.0004	0.60***	0.01***	0.002***	0.001*
	(0.09)	(0.001)	(0.0005)	(0.0002)	(0.11)	(0.001)	(0.001)	(0.0004)
Observations	124	124	124	124	124	124	124	124
R^2	0.17	0.18	0.12	0.16	0.08	0.11	0.09	0.06
Adjusted R^2	0.15	0.16	0.10	0.13	0.06	0.09	0.07	0.03

Note: *p<0.05; **p<0.01; ***p<0.001
Heteroscedasticity-Consistent standard errors.

The table displays the absolute price changes associated with the press release (estimations 1-4) and press conference (regressions 5-8). The results indicate that changes in the interest rate serve as the primary driver of price changes for the press release. This suggests that the press release holds significant value in confirming expected interest rates or revealing unexpected changes in interest rates. Note that the larger yen point estimates can be attributed to the differing yen levels compared to the other three currencies; these values have not been normalized. The absolute price changes of the press conference from 2016 onwards are significantly smaller than those from 2009 to 2015. This corroborates the findings presented in the preceding chapter and underscores the diminished importance of press releases in driving price changes over time.

The second hypothesis to be tested is whether trading volume has changed since 2016. Employing the same methodological approach, the equation is framed as follows:

$$v_{i,w,t} = \alpha + \beta_1 * D_{>2015} + \beta_2 * VIX_{t-1} + \beta_3 * |\Delta MRO|_t + \epsilon_t \tag{6.4}$$

$v_{i,w,t}$ is the logarithm of the average traded volume (definition 6.2). Table 6.2 presents the results for regression 6.4. Except for the USD, the logarithm of average traded volume was significantly higher post-2016 than pre-2016 for the press release window (specifications $1-4$). On the other hand, trading activity during press conferences exhibited a decline compared to the earlier period.

Table 6.2.: Estimation of average traded log volume for different currencies

	Dependent variable:							
	v_{MP}				v_P			
	$v_{YEN,MP,t}$	$v_{USD,MP,t}$	$v_{GBP,MP,t}$	$v_{CHF,MP,t}$	$v_{YEN,P,t}$	$v_{USD,P,t}$	$v_{GBP,P,t}$	$v_{CHF,P,t}$
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
$D_{>2015}$	0.44***	0.01	0.42***	0.71***	−0.20	−0.62***	−0.05	0.01
	(0.13)	(0.13)	(0.12)	(0.13)	(0.15)	(0.18)	(0.14)	(0.15)
VIX_{t-1}	−0.003	0.004	0.02*	0.02*	−0.01	−0.01	0.005	0.01
	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
$\lvert\Delta\ MRO_t\rvert$	0.89	1.13	0.90	1.25	−0.29	−0.38	−0.47	−0.46
	(0.77)	(0.84)	(0.69)	(0.79)	(0.59)	(0.67)	(0.53)	(0.70)
Constant	5.28***	5.15***	4.55***	3.99***	6.29***	6.20***	5.46***	4.97***
	(0.24)	(0.20)	(0.22)	(0.23)	(0.23)	(0.27)	(0.20)	(0.24)
Observations	124	124	124	124	124	124	124	124
R^2	0.11	0.03	0.14	0.22	0.03	0.11	0.01	0.02
Adjusted R^2	0.09	0.01	0.12	0.20	0.01	0.09	−0.02	−0.01

Note: *p<0.05; **p<0.01; ***p<0.001
Heteroscedasticity-Consistent standard errors.

6.2. The Effect of Sentiments on Volume

As previous literature has extensively discussed the relation between returns and sentiments, the third empirical task is to explore whether polarized language influences trading activity, potentially causing a shift in the trading volume to occur at later points in time. Smales & Apergis (2017) find that the complexity of the monetary policy statement language of the FED is an important determinant of trading volume. They claim that differences in trader opinion resulting from different interpretations of language results in more trading volume. Hayo et al. (2022) examine this for the ECB and find that more complex language in the introductory statement leads to a temporal shift of trading activity towards the subsequent Q&A session. In this context, I test the following questions: (i) is there a relationship between absolute sentiment and traded volume (ii) does the length of the press release alter the trading volume (iii) is trading activity shifted to later points of time when language is polarized, or statements get longer. I start with the base specification:

$$v_{i,w,t} = \alpha + \beta_1 * |S_{w,t,l}| + \beta_2 * VIX_{t-1} + \beta_3 * |\Delta MRO|_t + \epsilon_t \tag{6.5}$$

where $|S_{w,t,l}|$ is the absolute sentiment of the announcement measured with lexicon l. Given the various possible combinations of estimations, selecting one proves challenging. To address this issue, I conduct a regression analysis of the EUR/YEN currency pair against different sentiment scores obtained using the clustered approach (see Chapter 2). Specifically, I focus on the sentiment derived from the introductory statement, a widely employed strategy for analyzing monetary events.

The results of the simple regression analysis for the press conference window of various sentiment measures for the introductory statement (I) are presented in Table 6.3. Notably, the LM dictionary indicates significant negative estimates. Moreover, both LM and MP exhibit statistically significant effects when combined with the PR measure. The influence of sentiments in this context is somewhat ambiguous. Firstly, the non-financial dictionary lacks statistical significance, which can be interpreted as a positive outcome. On the other hand, employing the LM dictionary reveals significant negative effects. This observation might lead to a hypothesis that different market participants may interpret polarized words differently. Thus, I proceed to solely consider the PR methods in a manner consistent with prior research in this field.

Referring to a broader perspective, Table 6.4 focuses on the whole monetary event (ME). The results demonstrate a significant positive association between the absolute sentiment measure, MP, particularly during the Q&A session. This finding supports the existence of a fundamental relationship between trading volume and the sentiment expressed by the ECB, validating the use of these sentiment scores. As a result, this empirical evidence supports the claim that polarized language stimulates trading volume.

Table 6.3.: Estimation of the average traded volume of the EUR/YEN within the press conference window

	Dependent variable:							
	$v_{Y,P,t}$							
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
I^{LM}	−12.16*							−14.34**
	(5.05)							(4.96)
I^{HENRY}		−4.00						
		(4.80)						
I^{GI}			7.47					
			(5.90)					
$I^{SENTICENT}$				6.98				
				(7.46)				
I^{MP}					0.52		0.52	0.67*
					(0.30)		(0.29)	(0.31)
I^{EC}						−0.04	0.02	−0.16
						(0.46)	(0.45)	(0.44)
VIX_{t-1}	−0.004	−0.01	−0.01	−0.01	−0.01	−0.01	−0.01	0.003
	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
$\lvert\Delta\ MRO_t\rvert$	0.13	−0.19	−0.13	−0.14	−0.16	−0.09	−0.17	0.20
	(0.65)	(0.61)	(0.59)	(0.62)	(0.58)	(0.66)	(0.62)	(0.64)
Constant	6.28***	6.30***	5.89***	5.54***	5.82***	6.15***	5.82***	5.91***
	(0.24)	(0.23)	(0.30)	(0.63)	(0.28)	(0.21)	(0.26)	(0.27)
Observations	124	124	124	124	124	124	124	124
R^2	0.06	0.02	0.03	0.03	0.04	0.02	0.04	0.09
Adjusted R^2	0.04	0.0001	0.004	0.002	0.01	−0.01	0.01	0.06

Note: *p<0.05; **p<0.01; ***p<0.001
Heteroscedasticity-Consistent standard errors.

Table 6.4.: Estimation of the average traded volume within the whole monetary event (ME)

	Dependent variable:			
	$v_{Y,ME,t}$			
	(1)	(2)	(3)	(4)
$\lvert MP^{MP}\rvert$	0.56			
	(0.70)			
$\lvert I^{MP}\rvert$		0.77**		0.13
		(0.28)		(0.31)
$\lvert QA^{MP}\rvert$			1.26***	1.18**
			(0.30)	(0.37)
VIX_{t-1}	−0.01	−0.004	−0.003	−0.003
	(0.01)	(0.01)	(0.01)	(0.01)
$\lvert\Delta\ MRO_t\rvert$	0.81	−0.10	0.01	−0.01
	(1.17)	(0.57)	(0.64)	(0.64)
Constant	5.61***	5.55***	5.36***	5.32***
	(0.62)	(0.27)	(0.25)	(0.27)
Observations	124	124	124	124
R^2	0.02	0.07	0.14	0.14
Adjusted R^2	0.0003	0.05	0.12	0.11

Note: *p<0.05; **p<0.01; ***p<0.001
Heteroscedasticity-Consistent standard errors.

I now turn my attention to the press release and its impact on trading volume. As Parle (2022) does, I control for the surprise factors from Altavilla et al. (2019). This enables control over the surprise factor size rather than implementing a binary variable to indicate an announcement. I incorporate the surprise factors related to quantitative easing (QE) and policy target (T) variables using the replica package HFDSHOCKS (Baumgärtner, 2021). A robust relationship can be inferred if sentiments continue to influence trading volume. Additionally, I include the change in the length of each press release. The length of the press release is measured in terms of total characters, denoted with ΔL. The rationale behind this choice is that longer documents require more volume to incorporate new information, leaving more room for potential misinterpretations.

Table 6.5 shows the results when controlling for the surprise factors. I again employ the PR scores and ascertain that the absolute economic outlook (EC) score positively correlates with traded volume for the EUR/JPY pair, even after controlling for surprises. Specifications (3), (6), and (7) show that the change in the length of each press release is insignificant for the trading volume. In Table 6.6, I extend the regression to include other currencies. Like the JPY results, the GBP and CHF currency findings suggest that press releases may positively affect trading volume. However, in the case of the USD currency, the results are positive but not statistically significant. Overall, I find evidence that a target surprise (T) increases trading activity.

To examine hypothesis (iii), I investigate whether the trading activity is shifted to later points based on polarized language during the monetary event. To do so, I again control for the factors QE and T. Specifically, I focus on the time frame from 14:00 to 14:15, $M2$. Results are presented in Table 6.7. Unexpected quantitative easing announcements do increase the traded volume during the second half of the press release window. However, the sentiment scores and the change in the press conference length do not seem to impact the traded volume.

Additionally, I examine whether the sentiment influences the ratio between the average traded volume during the introductory statement and the Q&A session. I define this ratio as ϕ as well as the surprise factors of the press conference as QE^P and T^P. If sentiment shifts trading volume, we would expect to observe some correlation. Moreover, I investigate the average traded volume during the hour following the press conference, referring to this variable as $PostPress$. I find no evidence in Table 6.8 to suggest that either the control variables or the sentiments significantly affect the ratio of the trading volume. However, in specifications (5)-(8), I find that the sentiment of the Q&A session positively influences the trading volume after the press conference.

Table 6.5.: Estimation of the average traded log volume of the EUR/YEN within the press release window (M)

	Dependent variable:						
				$v_{Y,M,t}$			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
$\|MP^{MP}\|$	0.17			0.42			
	(0.61)			(0.63)			
$\|MP^{EC}\|$		0.84**			0.84**		
		(0.26)			(0.28)		
ΔL_t			0.0003			0.0002	
			(0.0002)			(0.0001)	
ΔL_t * $\|M^{EC}\|$							0.001
							(0.0003)
QE_t				1.28*	1.02	1.05	0.99
				(0.59)	(0.54)	(0.60)	(0.57)
T_t				0.05	0.06*	0.05	0.05
				(0.03)	(0.03)	(0.03)	(0.03)
VIX_{t-1}	−0.01	0.0002	−0.01	−0.01*	−0.01	−0.02*	−0.02*
	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
$\|\Delta\ MRO_t\|$	0.75	−0.28	0.66	1.80	0.47	1.13	0.96
	(1.14)	(0.81)	(0.74)	(1.32)	(0.87)	(0.90)	(0.90)
Constant	5.48***	5.18***	5.65***	5.43***	5.32***	5.76***	5.77***
	(0.56)	(0.28)	(0.20)	(0.52)	(0.23)	(0.19)	(0.19)
Observations	124	124	123	124	124	123	123
R^2	0.01	0.09	0.05	0.11	0.19	0.12	0.14
Adjusted R^2	−0.01	0.07	0.03	0.07	0.15	0.08	0.10

Note: *p<0.05; **p<0.01; ***p<0.001
Heteroscedasticity-Consistent standard errors.

Table 6.6.: Estimation of the average traded log volume for different currencies within the press release window

	Dependent variable:			
	$v_{YEN,M,t}$	$v_{USD,M,t}$	$v_{GBP,M,t}$	$v_{CHF,M,t}$
	(1)	(2)	(3)	(4)
$\|MP^{EC}\|$	0.84**	0.48	0.69*	1.25***
	(0.28)	(0.27)	(0.27)	(0.30)
VIX_{t-1}	−0.01	0.005	0.01	0.02*
	(0.01)	(0.01)	(0.01)	(0.01)
$\|\Delta\ MRO_t\|$	0.47	1.15	0.52	0.38
	(0.87)	(0.99)	(0.81)	(0.88)
QE_t	1.02	0.83	0.92	0.73
	(0.54)	(0.57)	(0.52)	(0.63)
T_t	0.06*	0.04	0.05*	0.08***
	(0.03)	(0.03)	(0.02)	(0.02)
Constant	5.32***	5.00***	4.63***	4.04***
	(0.23)	(0.24)	(0.22)	(0.24)
Observations	124	124	124	124
R^2	0.19	0.10	0.18	0.24
Adjusted R^2	0.15	0.06	0.15	0.20
Note:	*p<0.05; **p<0.01; ***p<0.001			
	Heteroscedasticity-Consistent standard errors.			

Table 6.7.: Estimation of the average traded log volume for different currencies within the press release window

	Dependent variable:							
	$v_{YEN,M2,t}$	$v_{USD,M2,t}$	$v_{GBP,M2,t}$	$v_{CHF,M2,t}$	$v_{YEN,M2,t}$	$v_{USD,M2,t}$	$v_{GBP,M2,t}$	$v_{CHF,M2,t}$
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
$\lvert MP^{EC}\rvert$	1.87	−7.25	3.12	5.51				
	(5.39)	(3.90)	(4.98)					
Δ L_t					0.0001	−0.0000	0.0001	0.0000
					(0.0002)	(0.0002)		(0.0001)
QE_t	1.49*	0.78	1.10*	1.02	1.35*	0.93	0.87	0.92
	(0.59)	(0.64)	(0.52)	(0.59)	(0.64)	(0.69)	(0.57)	(0.75)
T_t	0.04	0.01	0.03	0.07**	0.04	0.01	0.03	0.07
	(0.03)	(0.03)	(0.02)	(0.02)	(0.03)	(0.03)	(0.03)	(0.04)
VIX_{t-1}	−0.01	−0.0003	0.01	0.01	−0.01	0.001	0.01	0.01
	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
$\lvert\Delta$ $MRO_t\rvert$	0.91	1.41	0.73	1.35	0.85	1.80	0.63	1.17
	(0.91)	(0.89)	(0.70)	(0.75)	(1.00)	(1.04)	(0.90)	(0.91)
Constant	5.61***	5.24***	4.81***	4.47***	5.64***	5.15***	4.85***	4.52***
	(0.20)	(0.21)	(0.18)	(0.20)	(0.22)	(0.22)	(0.19)	(0.25)
Observations	124	124	124	124	123	123	123	123
R^2	0.10	0.08	0.11	0.11	0.10	0.07	0.11	0.10
Adjusted R^2	0.06	0.04	0.07	0.07	0.06	0.03	0.07	0.06

Note: $^{*}p<0.05$; $^{**}p<0.01$; $^{***}p<0.001$
Heteroscedasticity-Consistent standard errors.

Table 6.8.: Different estimations

	Dependent variable:							
	$\phi_{YEN,t}$	$\phi_{USD,t}$	$\phi_{GBP,t}$	$\phi_{CHF,t}$	$\mathrm{v}_{YEN,PostPress,t}$	$\mathrm{v}_{USD,PostPress,t}$	$\mathrm{v}_{GBP,PostPress,t}$	$\mathrm{v}_{CHF,PostPress,t}$
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
$\lvert \mathrm{I}^{EC} \rvert$	−0.43	−0.16	0.11	0.39	0.04	−0.44	−0.28	0.47
	(0.26)	(0.25)	(0.44)	(0.81)	(0.36)	(0.41)	(0.28)	(0.43)
$\lvert \mathrm{QA}^{EC} \rvert$	0.09	0.21	−0.45	−0.89	0.81*	0.48	1.06***	0.46
	(0.30)	(0.33)	(0.75)	(1.49)	(0.36)	(0.45)	(0.31)	(0.43)
QE^P_t	−0.02*	−0.02	−0.02	−0.04*	0.0004	0.01	0.01*	0.02**
	(0.01)	(0.01)	(0.01)	(0.02)	(0.01)	(0.01)	(0.01)	(0.01)
T^P_t	0.46	0.56	0.27	0.67	−0.49	−0.39	−0.94	−0.93
	(0.39)	(0.44)	(0.48)	(0.60)	(0.55)	(0.81)	(0.61)	(0.64)
VIX_{t-1}	0.04	0.04	−0.001	−0.004	0.005	−0.03	0.02	0.03
	(0.03)	(0.02)	(0.03)	(0.05)	(0.02)	(0.03)	(0.02)	(0.03)
$\lvert \Delta\ \mathrm{MRO}_t \rvert$	−0.02	−0.02	−0.03	−0.04	0.03	0.03	0.03*	0.01
	(0.01)	(0.02)	(0.02)	(0.03)	(0.02)	(0.02)	(0.01)	(0.02)
Constant	1.88***	1.74***	2.16***	2.99***	5.41***	5.35***	4.72***	4.19***
	(0.24)	(0.28)	(0.50)	(0.86)	(0.30)	(0.31)	(0.21)	(0.31)
Observations	124	124	124	124	124	124	124	124
R^2	0.05	0.04	0.02	0.04	0.09	0.05	0.15	0.09
Adjusted R^2	−0.004	−0.01	−0.03	−0.01	0.04	−0.004	0.11	0.05

Note: $^{*}p<0.05$; $^{**}p<0.01$; $^{***}p<0.001$
Heteroscedasticity-Consistent standard errors.

6.3. YouTube data

This section presents evidence that subtitles on YouTube can be employed to identify high-frequency sentiment movements during press conferences without the need to pay for sources. The objective here is not to test hypotheses but to establish some validity for using YouTube subtitle texts. Using a simple sentiment score calculation with the LM dictionary, the accuracy of subtitles is checked by comparing them against the official press conference text data from the website. It should be noted that the cluster method could not be applied due to inaccuracies in punctuation in the subtitles. The start of the press conference videos incorporates an introduction by the Director of General Communications that lasts only a few seconds. Nevertheless, the plot comparison in Figure 6.2 shows that the sentiment scores are nearly identical, thus confirming that YouTube subtitles are sufficient for textual analysis. The exact time stamps of the subtitles further enable researchers to identify both which words have been said and when they were said. To summarize, this could be a compelling resource available at no cost for utilization in future research.

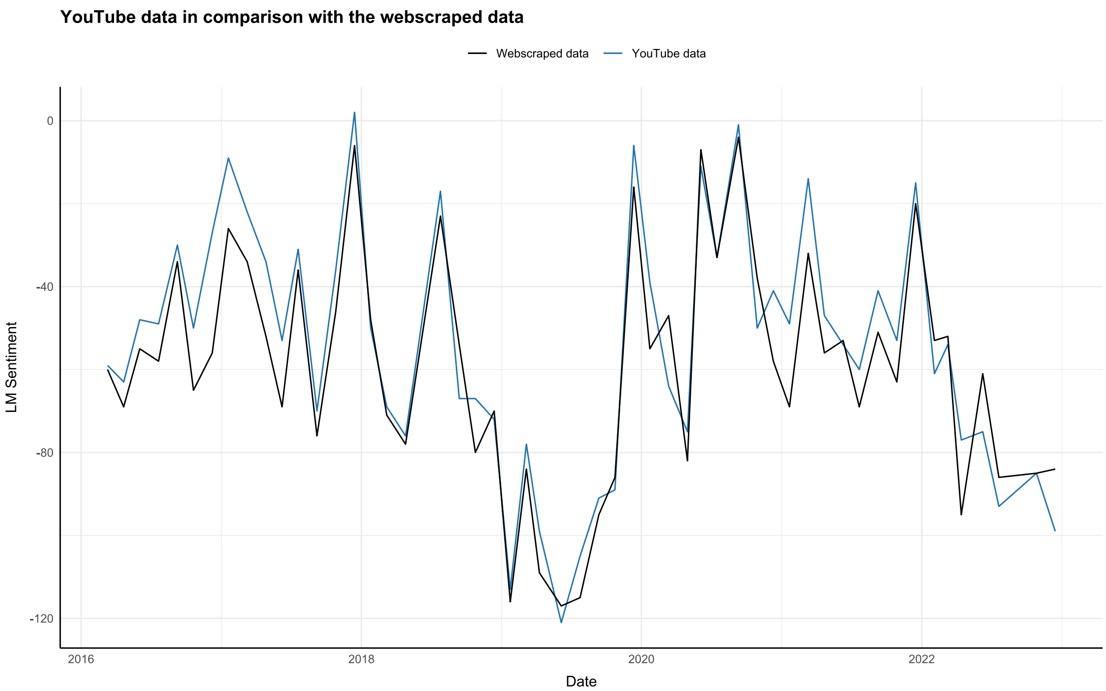

Figure 6.2.: Sentiment measure comparing the YouTube subtitles with the official web-scraped data of the press conference.

7. Conclusion

This thesis introduces the reader to basic Natural Language Processing (NLP) concepts while providing an in-depth dissection of sentiment analysis. The concepts are extrapolated and elucidated using a unique European Central Bank (ECB) document corpus. The primary finding is that press releases of the ECB changed in length and content in 2016. It replicates and extends the findings by Picault & Renault (2017), highlighting that press releases introduce interesting aspects to be analyzed. The study reveals that these press releases began to parallel the introductory statements of the press conference by 2016, evident from the graphical representations and a correlation table that was included.

In response to this shift in communication, the foreign exchange (FX) markets reacted twofold: (i) traded volume after the press release increased, and (ii) the pricing significance of press conferences decreased. This pattern is initially illustrated through graphical representations of the intraday trading behavior of absolute returns and trading volume for Thursdays. Event-based regressions were employed to conduct statistical tests. The results further support these observations. This framework allowed for testing various hypotheses contributing to the literature on central bank language and its impact on trading activity. The regressions primarily suggest a positive correlation between absolute sentiment scores and trading volume, measured using the method developed by Picault & Renault (2017). This especially holds true for the economic outlook measure. The analysis further shows that this particularly holds for unplanned question-and-answer sessions. However, no evidence is found supporting a shift in the timing of trading activity based on either polarized language or a change in the press release length. A minor additional result is that researchers can effectively analyze words per minute without solely relying on paid platforms such as Bloomberg. YouTube subtitles can also serve as a valuable resource for such analysis.

The empirical findings posited in this study may inspire theoretical models of central bank loss functions. The distinctive structure of the announcement day delineates an explicit trade-off that central banks face in managing price variances. Dewachter et al. (2014) distinguish between continuous volatility and discontinuous jumps within the EUR/USD exchange rate context. Based on these distinctions, it can be hypothesized that the ECB aims to strike a balance between continuous market volatility during its press conference and the occurrence of jumps at the moment of the press release. This trade-off suggests that the more information revealed in the press release, the lower the intensity of continuous volatility experienced. It is plausible to surmise that this approach also serves to minimize large jumps during the press conference while simultaneously maintaining accountability for communication made by the ECB president.

In conclusion, this study highlights the significance of understanding the role of lexicon variation in explaining trading behavior and interpreting textual data within financial markets. Expanding these analyses within a panel regression framework is crucial, accounting for fixed and time-fixed effects. This research holds promise for expanding our understanding of the intricacies involved in this process, ultimately fostering more informed announcements from central banks.

Bibliography

Algaba, A., Ardia, D., Bluteau, K., Borms, S., & Boudt, K. (2020). Econometrics meets sentiment: An overview of methodology and applications. *Journal of Economic Surveys*, 34(3), 512–547.

Altavilla, C., Brugnolini, L., Gürkaynak, R. S., Motto, R., & Ragusa, G. (2019). Measuring euro area monetary policy. *Journal of Monetary Economics*, 108, 162–179.

Anderes, M., Rathke, A., Streicher, S., & Sturm, J.-E. (2021). The role of ecb communication in guiding markets. *Public Choice*, 186(3-4), 351–383.

Ardia, D., Bluteau, K., Borms, S., & Boudt, K. (2021). The r package sentometrics to compute, aggregate and predict with textual sentiment. *arXiv preprint arXiv:2110.10817*.

Baranowski, P., Bennani, H., & Doryń, W. (2021). Do the ecb's introductory statements help predict monetary policy? evidence from a tone analysis. *European Journal of Political Economy*, 66, 101964.

Baumgärtner, M. (2021). *Package 'hfdshocks'*. https://github.com/martinbaumgaertner/hfdshocks.

Bholat, D., Hansen, S., Santos, P., & Schonhardt-Bailey, C. (2015). Text mining for central banks. *Available at SSRN 2624811*.

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.

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.

Coenen, G., Ehrmann, M., Gaballo, G., Hoffmann, P., Nakov, A., Nardelli, S., Persson, E., & Strasser, G. (2017). Communication of monetary policy in unconventional times.

Conrad, C. & Lamla, M. J. (2010). The high-frequency response of the eur-usd exchange rate to ecb communication. *Journal of Money, Credit and Banking*, 42(7), 1391–1417.

Cour-Thimann, P. (2020). Interest rate setting and communication at the ecb.

Cour-Thimann, P. & Jung, A. (2021). Interest-rate setting and communication at the ecb in its first twenty years. *European Journal of Political Economy*, 70, 102039.

Cross, J. P. & Greene, D. (2020). Talk is not cheap: Policy agendas, information processing, and the unusually proportional nature of european central bank communications policy responses. *Governance*, 33(2), 425–444.

De Santis, R. A. (2020). Impact of the asset purchase programme on euro area government bond yields using market news. *Economic Modelling*, 86, 192–209.

Denny, M. J. & Spirling, A. (2018). Text preprocessing for unsupervised learning: Why it matters, when it misleads, and what to do about it. *Political Analysis*, 26(2), 168–189.

Dewachter, H., Erdemlioglu, D., Gnabo, J.-Y., & Lecourt, C. (2014). The intra-day impact of communication on euro-dollar volatility and jumps. *Journal of International Money and Finance*, 43, 131–154.

ECB (2014). Monetary policy decisions. Original publication date: 4 December 2014, accessed: 4 August 2023, https://www.ecb.europa.eu/press/pr/date/2014/html/pr141204.en.html.

ECB (2016). Monetary policy decisions. Original publication date: 10 March 2016, accessed: 4 August 2023, https://www.ecb.europa.eu/press/pr/date/2016/html/pr160310.en.html.

ECB (2017). Monetary policy decisions. Original publication date: 9 March 2017, accessed: 4 August 2023, https://www.ecb.europa.eu/press/pr/date/2017/html/pr170309.en.html.

Ehrmann, M. & Fratzscher, M. (2007). Explaining monetary policy in press conferences.

Eusepi, S. & Preston, B. (2010). Central bank communication and expectations stabilization. *American Economic Journal: Macroeconomics*, 2(3), 235–271.

Fratzscher, M. (2008). Communication and exchange rate policy. *Journal of Macroeconomics*, 30(4), 1651–1672.

Gentzkow, M., Kelly, B., & Taddy, M. (2019). Text as data. *Journal of Economic Literature*, 57(3), 535–74.

Gerlach, S. (2007). Interest rate setting by the ecb, 1999-2006: Words and deeds. *Tenth issue (September 2007) of the International Journal of Central Banking*.

Gertler, P. & Horvath, R. (2018). Central bank communication and financial markets: New high-frequency evidence. *Journal of Financial Stability*, 36, 336–345.

González, M. & Tadle, R. C. (2020). Signaling and financial market impact of chile's central bank communication. *Economía*, 20(2), 127–178.

Gürkaynak, R. S., Sack, B. P., & Swanson, E. T. (2004). Do actions speak louder than words? the response of asset prices to monetary policy actions and statements. *The Response of Asset Prices to Monetary Policy Actions and Statements (November 2004)*.

Hayo, B., Henseler, K., Rapp, M. S., & Zahner, J. (2022). Complexity of ecb communication and financial market trading. *Journal of International Money and Finance*, 128, 102709.

Heinemann, F. & Ullrich, K. (2007). Does it pay to watch central bankersâ lips? the information content of ecb wording. *Swiss Journal of Economics and Statistics*, 143, 155–185.

Henry, E. (2008). Are investors influenced by how earnings press releases are written? *The Journal of Business Communication (1973)*, 45(4), 363–407.

Hlavac, M. (2022). stargazer: Well-formatted regression and summary statistics tables. r package version 5.2. 3. *Social Policy Institute: Bratislava, Slovakia*.

Hong, H. & Stein, J. C. (2007). Disagreement and the stock market. *Journal of Economic perspectives*, 21(2), 109–128.

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

Jansen, D.-J. & De Haan, J. (2005). Talking heads: the effects of ecb statements on the euro–dollar exchange rate. *Journal of International Money and Finance*, 24(2), 343–361.

Jansen, D.-J. & De Haan, J. (2007). The importance of being vigilant: Has ecb communication influenced euro area inflation expectations?

Jarociński, M. & Karadi, P. (2020). Deconstructing monetary policy surprises–the role of information shocks. *American Economic Journal: Macroeconomics*, 12(2), 1–43.

Kandel, E. & Pearson, N. D. (1995). Differential interpretation of public signals and trade in speculative markets. *Journal of Political Economy*, 103(4), 831–872.

Kerssenfischer, M. (2019). Information effects of euro area monetary policy: New evidence from high-frequency futures data.

Klejdysz, J. (2018). Shifts in ecb communication: a text mining approach. *MSc thesis*.

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.

Lucca, D. O. & Trebbi, F. (2009). *Measuring central bank communication: an automated approach with application to FOMC statements*. Technical report, National Bureau of Economic Research.

Mohammad, S. M. & Turney, P. D. (2013). Nrc emotion lexicon. *National Research Council, Canada*, 2, 234.

Musard-Gies, M. (2006). Do ecb's statements steer short-term and long-term interest rates in the euro zone?

Nassirtoussi, A. K., Aghabozorgi, S., Wah, T. Y., & Ngo, D. C. L. (2014). Text mining for market prediction: A systematic review. *Expert Systems with Applications*, 41(16), 7653–7670.

Neely, C. J. (2011). A survey of announcement effects on foreign exchange volatility and jumps. *FRB of St. Louis Review*.

Parle, C. (2022). The financial market impact of ecb monetary policy press conferencesâa text based approach. *European Journal of Political Economy*, 74, 102230.

Pavelkova, A. (2022). The ecb press conference: a textual analysis.

Perlin, M. S., Mastella, M., Vancin, D. F., & Ramos, H. P. (2020). A garch tutorial with r. *Revista de Administração Contemporânea*, 25.

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.

Porter, M. F. (1980). An algorithm for suffix stripping. *Program*, 14(3), 130–137.

Priola, M. P., Molino, A., Tizzanini, G., et al. (2021). The informative value of central banks talks: a topic model application to sentiment analysis.

Romer, C. D. & Romer, D. H. (1989). Does monetary policy matter? a new test in the spirit of friedman and schwartz. *NBER macroeconomics annual*, 4, 121–170.

Romer, C. D. & Romer, D. H. (2023). *Does Monetary Policy Matter? The Narrative Approach after 35 Years*. Technical report, National Bureau of Economic Research.

Rosa, C. (2013). The financial market effect of fomc minutes. *Economic Policy Review*, 19(2).

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. (2015). Does central bank tone move asset prices? *Available at SSRN 2629978*.

Silge, J. & Robinson, D. (2017). *Text mining with R: A tidy approach*. O'Reilly Media, Inc.

Smales, L. & Apergis, N. (2017). Understanding the impact of monetary policy announcements: The importance of language and surprises. *Journal of Banking & Finance*, 80, 33–50.

Taboada, M., Brooke, J., Tofiloski, M., Voll, K., & Stede, M. (2011). Lexicon-based methods for sentiment analysis. *Computational linguistics*, 37(2), 267–307.

Tadle, R. C. (2022). Fomc minutes sentiments and their impact on financial markets. *Journal of economics and business*, 118, 106021.

Tillmann, P. & Walter, A. (2018). *ECB vs Bundesbank: Diverging tones and policy effectiveness*. Technical report, MAGKS Joint Discussion Paper Series in Economics.

Ulrich, M., Jakobs, E., & Tran, R. (2019). The real-time impact of ecb press conferences on financial markets. *Available at SSRN 3436852*.

Varian, H. R. (1989). Differences of opinion in financial markets. In *Financial Risk: Theory, Evidence and Implications: Proceedings of the Eleventh Annual Economic Policy Conference of the Federal Reserve Bank of St. Louis* (pp. 3–37).: Springer.

Venturini, A. (2020). Analysing the impact of ecb communication on financial markets: A text mining approach.

Wickham, H. (2022). *rvest: Easily Harvest (Scrape) Web Pages*. https://rvest.tidyverse.org/, https://github.com/tidyverse/rvest.

Zeileis, A., Lumley, T., Berger, S., Graham, N., & Zeileis, M. A. (2019). Package 'sandwich'. *R package version*, (pp. 2–5).

A. Appendix

A.1. Tables Chapter 4

Table A.1.: ECB Governing Council Press Conference YouTube Videos 2020−2022

Date	Link	Youtube Channel
	2022	
2022-12-15	https://www.youtube.com/watch?v=z0HCd9g-VzA	t=2s
European Central Bank@ecbeuro		
2022-10-27	https://www.youtube.com/watch?v=5XYCEfjoFkU	t=2s
European Central Bank@ecbeuro		
2022-09-08	**only Introductory** https://www.youtube.com/watch?v=2XgKpi-J-P0	European Central Bank@ecbeuro
2022-07-21	https://www.youtube.com/watch?v=j1SVcroPuQM&t=1s	European Central Bank@ecbeuro
2022-06-09	https://www.youtube.com/watch?v=d$_{u}tpAxGMYo$	European Central Bank@ecbeuro
2022-04-14	https://www.youtube.com/watch?v=URG4k05tiAY	t=1s
European Central Bank@ecbeuro		
2022-03-10	https://www.youtube.com/watch?v=qdQYkgS-SoE	European Central Bank@ecbeuro
2022-02-03	https://www.youtube.com/watch?v=yXHehAdOYkI	European Central Bank@ecbeuro
	2021	
2021-12-16	https://www.youtube.com/watch?v=Ra6S1O6ZMb0	European Central Bank@ecbeuro
2021-10-28	https://www.youtube.com/watch?v=oggV578FTiA	European Central Bank@ecbeuro
2021-09-09	https://www.youtube.com/watch?v=V$_{G}nH_{g}GksQ$	European Central Bank@ecbeuro
2021-07-22	https://www.youtube.com/watch?v=hbRjSC32IGM	European Central Bank@ecbeuro
2021-06-10	https://www.youtube.com/watch?v=ChuPTLUT1YQ	European Central Bank@ecbeuro
2021-04-22	https://www.youtube.com/watch?v=BE9rF9yX4YY	European Central Bank@ecbeuro
2021-03-11	https://www.youtube.com/watch?v=D$_{7}eU4IXUjY$	European Central Bank@ecbeuro
2021-01-21	https://www.youtube.com/watch?v=7VMNOqI-cyo	European Central Bank@ecbeuro
	2020	
2020-12-10	https://www.youtube.com/watch?v=w1nMGcfJ8Eo	European Central Bank@ecbeuro
2020-10-29	https://www.youtube.com/watch?v=7wjfcS8Dals	European Central Bank@ecbeuro
2020-09-10	https://www.youtube.com/watch?v=AAAR37DFo3I	European Central Bank@ecbeuro
2020-07-16	https://www.youtube.com/watch?v=CNVjMYBeKqM	European Central Bank@ecbeuro
2020-06-04	https://www.youtube.com/watch?v=O1-nZ-31Vhw	European Central Bank@ecbeuro
2020-04-30	https://www.youtube.com/watch?v=5Qkxio9lNLE	European Central Bank@ecbeuro
2020-03-12	https://www.youtube.com/watch?v=ojOu7JP-1g0	European Central Bank@ecbeuro
2020-01-23	https://www.youtube.com/watch?v=pIslnWRL2NI	European Central Bank@ecbeuro

Table A.2.: ECB Governing Council Press Conference YouTube Videos 2016−2019

Date	Link	Youtube Channel
	2019	
2019-12-12	https://www.youtube.com/watch?v=ucuSp2QiK$_I$	European Central Bank@ecbeuro
2019-10-24	https://www.youtube.com/watch?v=VyKP1ykyewo	t=3s
European Central Bank@ecbeuro		
2019-09-12	https://www.youtube.com/watch?v=WrwQSFe$_D lo$	European Central Bank@ecbeuro
2019-07-25	https://www.youtube.com/watch?v=IBGTRRXF1vw	European Central Bank@ecbeuro
2019-06-06	https://www.youtube.com/watch?v=w14YTNsxPFE	t=16s
European Central Bank@ecbeuro		
2019-04-10	https://www.youtube.com/watch?v=ZcOVNY12ibQ	European Central Bank@ecbeuro
2019-03-07	https://www.youtube.com/watch?v=QTeYjQflIrI	European Central Bank@ecbeuro
2019-01-24	https://www.youtube.com/watch?v=-qq6GcDyhBI	European Central Bank@ecbeuro
	2018	
2018-12-13	https://www.youtube.com/watch?v=BmWkkrFS0cU	European Central Bank@ecbeuro
2018-10-25	https://www.youtube.com/watch?v=jAdM8RFrYL8	European Central Bank@ecbeuro
2018-09-13	https://www.youtube.com/watch?v=spp58hr-U2Y	European Central Bank@ecbeuro
2018-07-26	—	—
2018-04-26	https://www.youtube.com/watch?v=1phnPPwSLKc	European Central Bank@ecbeuro
2018-03-08	https://www.youtube.com/watch?v=Yl3cPy-6BVM	European Central Bank@ecbeuro
2018-01-25	https://www.youtube.com/watch?v=UTtnNapbjQo	European Central Bank@ecbeuro
	2017	
2017-12-14	https://www.youtube.com/watch?v=ZqkbVcu6jh8	European Central Bank@ecbeuro
2017-10-26	https://www.youtube.com/watch?v=9UqzddIO7i4	European Central Bank@ecbeuro
2017-09-07	https://www.youtube.com/watch?v=AjwGLTgeNSM	European Central Bank@ecbeuro
2017-07-20	https://www.youtube.com/watch?v=sTGKxfz2Mvw	European Central Bank@ecbeuro
2017-06-08	https://www.youtube.com/watch?v=VenJ$_C K2tPM$	European Central Bank@ecbeuro
2017-04-27	https://www.youtube.com/watch?v=OSy$_v W1RSSo$	European Central Bank@ecbeuro
2017-03-09	https://www.youtube.com/watch?v=T8fDJoc04AA	Statisma News@statismanews5090
2017-01-19	https://www.youtube.com/watch?v=6lN37ZrZP9E	Statisma News@statismanews5090
	2016	
2016-12-08	https://www.youtube.com/watch?v=9OhdSL7gmtc	European Central Bank@ecbeuro
2016-10-20	https://www.youtube.com/watch?v=NLuLw0qO834	
European Central Bank@ecbeuro		
2016-09-08	https://www.youtube.com/watch?v=C48YfIBBK7g	European Central Bank@ecbeuro
2016-07-21	https://www.youtube.com/watch?v=6FfQSvivnlA	Statisma News@statismanews5090
2016-06-02	https://www.youtube.com/watch?v=UMZEPmcq-1A	Statisma News@statismanews5090
2016-04-21	https://www.youtube.com/watch?v=Gc4urujBcOA	European Central Bank@ecbeuro
2016-03-10	https://www.youtube.com/watch?v=tMH8vmjnHjw	European Central Bank@ecbeuro

Table A.3.: 10 largest absolute returns in the 1h sample

Date	Time	Close	Returns	Volume	Volume Change
21.01.2009	15:00	112.82	-0.0341	17,542	0.681
06.05.2010	18:00	111.54	-0.0412	28,667	0.579
06.05.2010	19:00	113.93	0.0212	16,150	-0.574
09.05.2010	21:00	119.94	0.0305	3,329	6.03
16.03.2011	21:00	107.62	-0.0272	9,636	0.250
31.10.2011	01:00	110.39	0.0298	21,784	1.51
28.06.2015	21:00	134.82	-0.0255	11,045	0.660
24.06.2016	02:00	112.06	-0.0508	63,825	0.170
23.04.2017	13:00	120.10	0.0272	4,404	-0.829
02.01.2019	12:00	120.82	-0.0222	20,634	0.740

Table A.4.: VIX summary table

	Min.	1st Qu.	Median	Mean	3rd Qu.	Max.
VIX	10.97	17.82	22.25	23.95	27.89	95.02

A.2. Figures Chapter 5

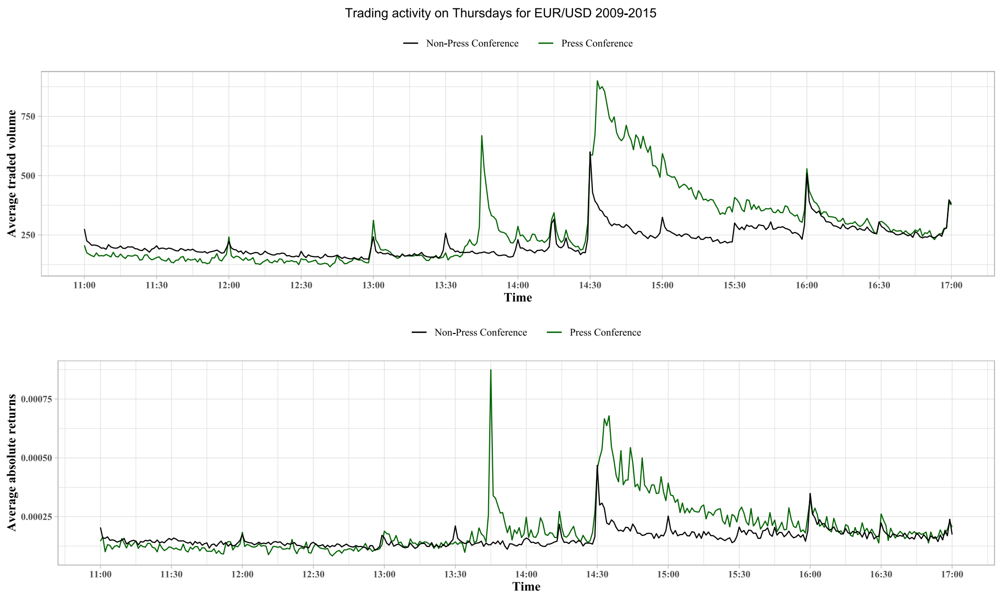

Figure A.1.: Average traded volume and average absolute returns EUR/USD 2009–2015.

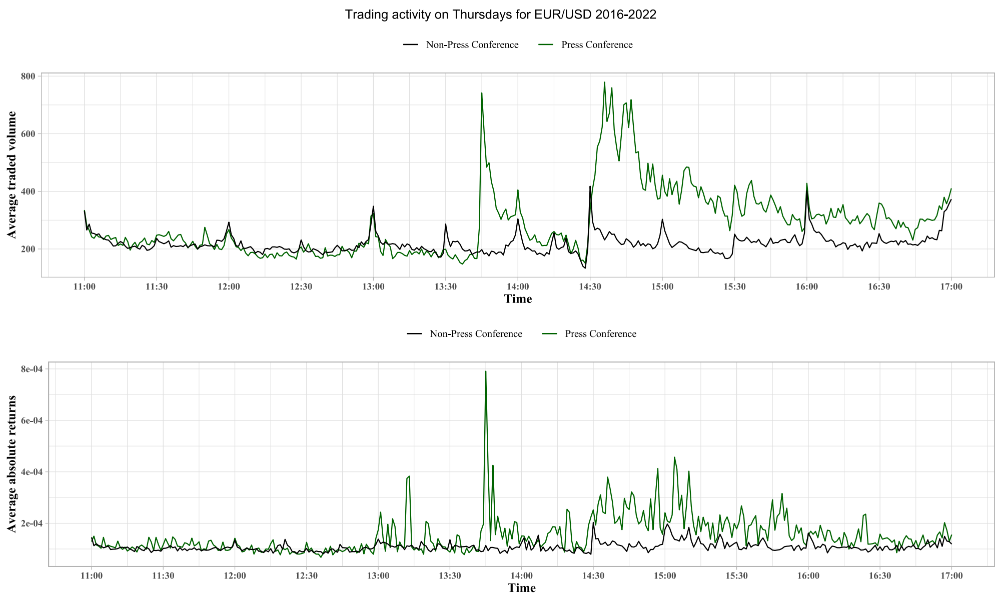

Figure A.2.: Average traded volume and average absolute returns EUR/USD 2016–2022.

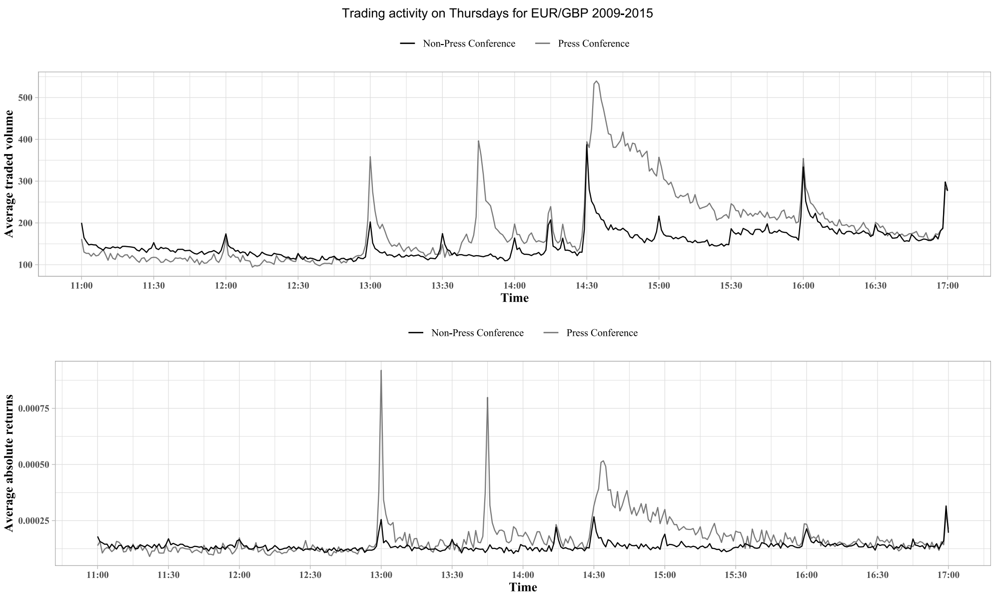

Figure A.3.: Average traded volume and average absolute returns EUR/GBP 2009–2015.

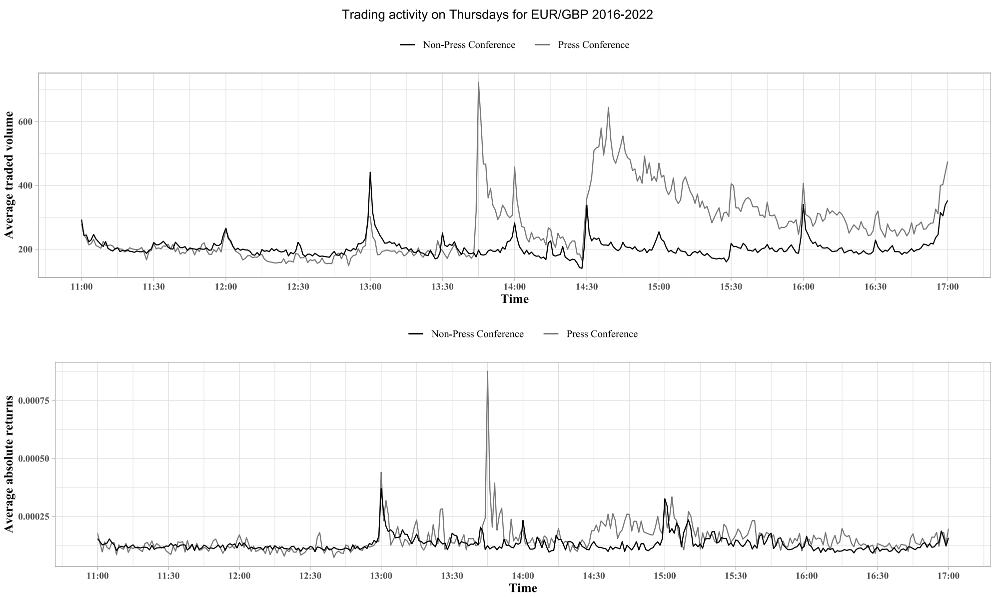

Figure A.4.: Average traded volume and average absolute returns EUR/GBP 2016–2022.

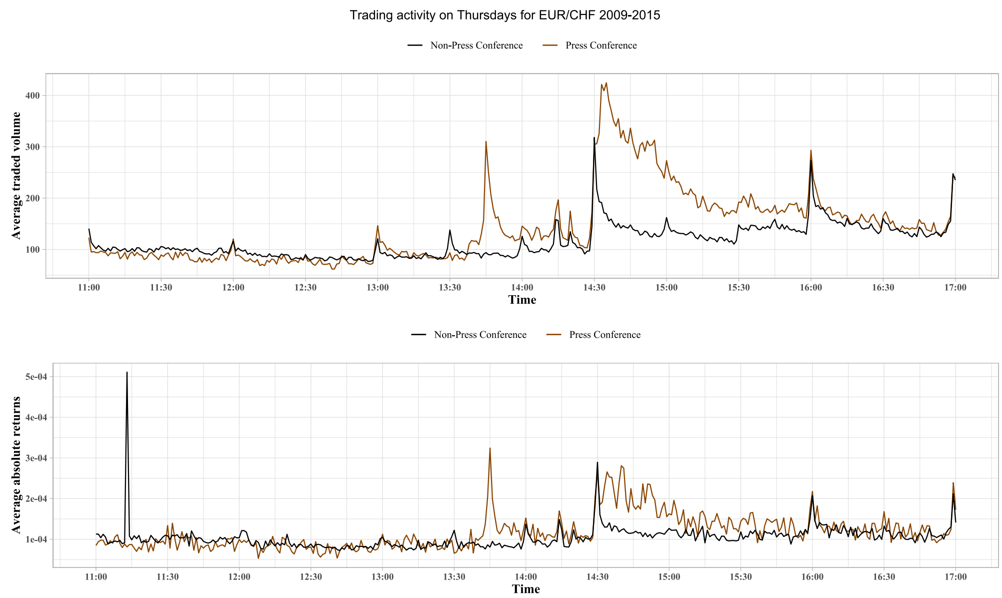

Figure A.5.: Average traded volume and average absolute returns EUR/CHF 2009–2015.

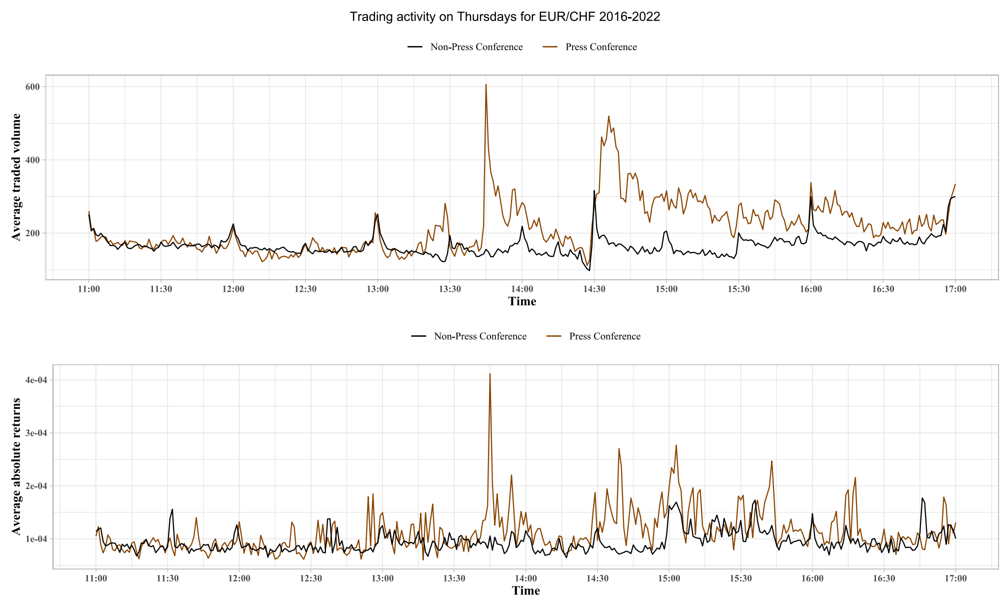

Figure A.6.: Average traded volume and average absolute returns EUR/CHF 2016–2022.