Australian National University

Crawford School of Public Policy

CAMA

Centre for Applied Macroeconomic Analysis

Emotion in Euro Area Monetary Policy Communication and Bond Yields: The Draghi Era

CAMA Working Paper 75a/2022
December 2022

Dimitrios Kanelis
Westfälische Wilhelms-Universität Münster

Pierre L. Siklos
Wilfrid Laurier University
CAMA, Crawford, ANU, Australia

Abstract

We combine modern methods from Speech Emotion Recognition and Natural Language Processing with high-frequency financial data to analyze how the vocal emotions and language of ECB President Mario Draghi affect the yield curve of major euro area economies. Vocal emotions significantly impact the yield curve. However, their impact varies in size and sign: positive signals raise German and French yields, while Italian yields react negatively, which is reflected in an increase in yield spreads. A by-product of our study is the construction and provision of a synchronized data set for voice and language.

Keywords

Communication, ECB, Neural Networks, High-Frequency Data, Speech Emotion Recognition, Asset Prices

JEL Classification

E50, E58, G12, G14

Address for correspondence:

(E) cama.admin@anu.edu.au

ISSN 2206-0332

The Centre for Applied Macroeconomic Analysis in the Crawford School of Public Policy has been established to build strong links between professional macroeconomists. It provides a forum for quality macroeconomic research and discussion of policy issues between academia, government and the private sector.

The Crawford School of Public Policy is the Australian National University's public policy school, serving and influencing Australia, Asia and the Pacific through advanced policy research, graduate and executive education, and policy impact.

Emotion in Euro Area Monetary Policy Communication and Bond Yields: The Draghi Era

Dimitrios Kanelis[1], Pierre L. Siklos[2]

December 2022

Abstract

We combine modern methods from Speech Emotion Recognition and Natural Language Processing with high-frequency financial data to analyze how the vocal emotions and language of ECB President Mario Draghi affect the yield curve of major euro area economies. Vocal emotions significantly impact the yield curve. However, their impact varies in size and sign: positive signals raise German and French yields, while Italian yields react negatively, which is reflected in an increase in yield spreads. A by-product of our study is the construction and provision of a synchronized data set for voice and language.

JEL Classification: E50, E58, G12, G14

Keywords: Communication, ECB, Neural Networks, High-Frequency Data, Speech Emotion Recognition, Asset Prices

[1]Corresponding Author: Department of Economics, Westfälische Wilhelms-Universität Münster, Am Stadtgraben 9, 48143 Münster, Germany, Phone: +49 251 83-22989, E-mail: Dimitrios.Kanelis@wiwi.uni-muenster.de, Declarations of Interest: None

[2]Department of Economics, Wilfrid Laurier University, Balsillie School of International Affairs, and the Centre for Applied Macroeconomic Analysis (CAMA), Declarations of Interest: None

"[The Press Conference] was originally an obligation, then it became a welcome obligation, and then even a pleasure. [...] Communication has become a tool of monetary policy, so your interaction has been essential in our monetary policy decisions all throughout these eight years."

Mario Draghi to Journalists, 24 October 2019

"[D]ue to this intervention of the activist at the beginning of the press conference I observed that you remained very calm. How do you manage this?"

Journalist to Mario Draghi, 15 April 2015

1. Introduction

When Mario Draghi was President of the European Central Bank (ECB) and spoke, financial markets listened. However, research on central bank communication has yet to quantify how financial markets reacted to the voice of the ECB President during press conferences. As the above quote from a journalist illustrates, it is not just written words that listeners take into account but the full range of communicative signals. Mehrabian (1971) describes, and others have since shown (see below), that this includes non-verbal forms of communication. For example, the New York Times (NYT) explicitly noticed and referred to the annoyance in Draghi's voice when he answered one of the questions a journalist asked.[1]

Since the introduction of forward guidance (FG), central bank communication has increased and reached a new level of importance for the conduct of monetary policy. Over time, communication became significantly more standardized as central bankers attempted to avoid surprising markets. Accordingly, it is to be expected that financial markets consider information that is difficult to systematize. One source of unscripted information is a person's voice since the voice does not necessarily deliver the same message as the words. Gorodnichenko et al. (2020) provide evidence of how vocal emotions of the Federal Reserve (FED)

[1]The NYT writes: *"'Go back and ask yourself, where were you two years ago?' Mr. Draghi said, with a hint of annoyance in his voice."* See: https://www.nytimes.com/2014/10/03/business/international/ecb-leaves-key-interest-rate-unchanged.html (Last Access: 1 September 2022).

chair have significant effects on different asset prices and show the importance of unscripted information in U.S. monetary policy communication.

In this study, we combine state-of-the-art methods from Speech Emotion Recognition (SER) and Natural Language Processing (NLP) with high-frequency financial data to estimate the effect on the yield curves of the four largest euro area economies of the interplay of vocal emotions and language during the Q&A sessions of the ECB press conferences under Mario Draghi's presidency. Focusing on yields has advantages since government bonds are highly liquid assets, and their price is widely considered a reference point for the overall financing conditions of an economy. Relying on French, Italian, and Spanish bonds, we focus on the spread that we define as the difference between the yield of the respective bonds and German bonds with the same duration to study how vocal emotions affect the risk behavior of investors. German yields have long been considered the benchmark against which financing conditions in the euro area are judged.

We conduct an event study and construct a novel data set consisting of timely synchronized audio and textual data for press conferences between May 2012 and October 2019. One challenge is that Draghi answers several questions in a row on totally different topics. To ensure that we accurately measure vocal emotions over a wide variety of topics, we exploit an interesting characteristic of the ECB press conference transcripts. The ECB staff identifies individual answers or focal points and structures these in writing. Following this structure, we are able to adjust the audio data for each answer and establish synchronicity between voice and words. We provide additional details when we describe the methodology.

To measure vocal emotions, we implement the Fully Convolutional Neural Network (FCN) of García-Ordás et al. (2021), which has the special property that it is able to process audio files with non-fixed length. This framework has the advantage that we do not need to standardize and process audio signals a priori and, therefore, we avoid a potential loss in audio information. To measure the framing of language, which is also primarily unscripted during the Q&A session, we implement Fin-BERT. This large-language neural network model can analyze economic and finance-related text and goes beyond the word-counting approach of dictionary methods.

The results provide evidence for the relevance of vocal emotions in the interplay with language with important policy implications that may be used to improve monetary policy communication. We estimate a significant effect of vocal emotions and language on the yields of all four major euro area economies. However, these effects vary considerably in size and sign. In the case of German bonds, which financial markets view as the benchmark, positive signals lead to increases in the yields though only for bonds with a two-year duration. Using data from France, Italy, and Spain and focusing on spreads, we can discriminate between positive and negative language and vocal emotions and their impact on yield spreads. For example, we find that negative communication increases the Italian spread.

Our study contributes in at least three different ways to the literature. First, to the best of our knowledge, we are the first to document the importance of vocal emotions of the President of the ECB. Second, our study applies the neural network literature to economic analysis by introducing and using the model of García-Ordás et al. (2021), which belongs to the class of FCN models. This methodological contribution is an important next step in using audio data for research in central bank communication and empirical finance due to its capability to process non-fixed audio files, which is necessary to properly evaluate answers with different duration, for example, during Q&A sessions or conference calls. The model allows a classification also in real time. Third, we construct and provide a novel data set of synchronized voice and language data and additional qualitative data for future research.

The remainder of the study is structured as follows. Section 2 provides a literature review on monetary policy event studies, its intersection with central bank communication, and the recently increasing literature on considering vocal emotions. Section 3 extensively describes the methodology used to construct the data set and implement the event study regression. Section 4 presents the results. The final section concludes.

2. Event Studies and Central Bank Communication

Event studies represents one approach to analyze the direct causal effect of monetary policy decisions. By using a narrow time window around an event and utilizing high-frequency data (daily or intra-daily), identification of the causal effects is obtained by disentangling and ordering the sequence of events (Ramey, 2016). Since the first seminal studies using this identification strategy (e.g., Kuttner (2001); Cochrane and Piazzesi (2002); Rigobon and Sack (2004)), the literature has grown considerably. Gürkaynak et al. (2005) use factor analysis with high-frequency federal funds futures (FFF) data and estimate two factors to explain the FFF interest rate movements in real-time. They label the first factor as "current federal funds rate target," because it measures the effect of changes in the current federal funds rate; the second factor is defined as the "future path of policy" since it captures the impact of future changes in the federal funds rate. The authors analyze how monetary policy announcements drive the second factor during press conferences. Following Gürkaynak et al. (2005), Rosa (2011) analyzes the effects of U.S. monetary policy on the U.S. Dollar exchange rate against different currencies. In addition to a monetary shock, Rosa (2011) quantifies unexpected statements of the Fed chair using a narrative approach. He reports that surprising statements explain a large part of exchange rate movements, not the monetary shock. Gertler and Karadi (2015) identify monetary shocks, which they derive using high-frequency financial data as external instruments in a SVAR analysis to estimate the dynamic effects of these shocks on the macroeconomy. Nakamura and Steinsson (2018) build on a high-frequency setting and document a "Fed information effect" which is able to explain the positive effect of tightening monetary policy on economic growth expectations, in contrast to standard economic theory. The authors explain this effect by arguing that the Fed reveals private information with their interest rate decision. Nonetheless, Hoesch et al. (2020) demonstrate that, in recent years, the "Fed information effect" has been dwindling due to forward-looking policies like FG. Cieslak and Schrimpf (2019) distinguish between monetary and non-monetary news that markets perceive during central bank press conferences and estimate separate effects on the comovement of stocks and interest rates. The authors find that the composition of news in central bank statements varies considerably, with non-monetary news driving communication between 2008 - 2013 and monetary news thereafter. Swanson (2021) extends the identification approach of Gürkaynak et al. (2005) to the years following the Great Financial Crisis (GFC) and finds that additional factors are necessary to explain real-time movements of asset prices. These are motivated by the unconventional monetary policies of the Fed.

Event studies focusing on the ECB generally exploit the unique structure of monetary policy decision announcements, that is, the time difference between the press release at 13:45 CET and the press conference at 14:30 CET (e.g., Brand et al. (2010)).[2] Altavilla et al. (2019) use factor analysis on real-time Overnight Index Swap (OIS) data and consider that three latent factors necessary to explain the variance of changes in OIS during the press conference since the GFC. Analogous to Swanson (2021), they attribute this to the increased importance of unconventional monetary policy instruments such as FG and quantitative easing (QE).

Another strand of the literature relevant for our study are studies of central bank communication. Blinder et al. (2008) provide an early survey that has recently been updated by Blinder et al. (2022). Typical strategies to quantify qualitative central bank communication rests on using dictionary methods (e.g., Loughran and McDonald (2011); Apel and

[2]Since July 21, 2022, the ECB has changed the times of its announcement.

Blix Grimaldi (2014)) or textual indicators[3] on textual data from various central bank documents like introductory statements (e.g., Picault and Renault (2017)), transcripts (e.g., Shapiro and Wilson (2021)), speeches (e.g., Bohl et al. (2023)), press releases (Ehrmann and Talmi, 2020), or contributions to social media (Ehrmann and Wabitsch, 2022). Hubert and Labondance (2021) demonstrate that the tone of the introductory statements explains monetary shocks and that it can predict future monetary policy decisions, while Schmelling and Wagner (2019) and Parle (2022) provide evidence that the framing of the introductory statements has real-time effects on stock prices. Lombardi et al. (2019) develop a novel dictionary and estimate international spillovers resulting from the language used by the Fed chair. Since the GFC, the effects of spillovers on the yield curve have become even more significant and persistent.

A third strand of the literature we build on and contribute to is in the area of machine learning methods (Mullainathan and Spiess, 2017). In the case of central bank communication, several studies use different variants of topic modeling on central bank documents to disentangle the different and connected topics and intentions (e.g., Hansen et al. (2018), Dybowski and Kempa (2020), Feldkircher et al. (2021), Ferrara et al. (2021), Parle (2022)). Our study instead builds on the literature on neural networks.

Since press releases following policy rate decisions often change little (Ehrmann and Talmi, 2020), markets may also want to take into account less scripted information that provides additional hints about the conduct of monetary policy or the sentiment of the policy-making committee. The emotional attitude of central bankers is one source of unscripted information that may help explain asset price movements. Indeed, as financial research demonstrates, investors observe vocal emotions. In an early approach to analyzing non-verbal communication, Mayew and Venkatachalam (2012) provide evidence that vocal emotions influence

[3] Among others, textual complexity or similarity.

investors and that positive (negative) vocal emotions of managers precede positive (negative) news about corporate performance. Hu and Ma (2020) show that vocal emotions are an important factor for start-up entrepreneurs when applying for venture capital funds.

In the case of central banking, Q&A sessions, following the introductory statements, offer a potential source of unscripted information that journalists actively draw on by asking questions on topics that central bankers may prefer to avoid discussing. For example, Gorodnichenko et al. (2020) measure the vocal sentiment of the Fed chair and estimate a significant effect on stock prices in the days following FOMC press conferences. The authors refer to asymmetric information as a potential explanation for the market's interest in non-verbal behavior. This explanation is also consistent with the argument of Mayew and Venkatachalam (2012) who acknowledge that negative private information contradicting one's beliefs results in an uncomfortable state of mind that manifests itself in corresponding non-verbal communication.[4] Complementary studies by Curti and Kazinnik (2021) and Alexopoulos et al. (2022) estimate real-time effects of the chair's facial emotions on stock prices.

In the euro area, one can also find examples of how observers detect non-verbal reactions of former ECB president Mario Draghi, like his calm behavior[5] or his annoyed reaction when a journalist refers to his German critics[6]. One should also note that the ECB does not reveal much information about the actual debates or the climate during discussions inside the governing council meetings.[7] One can only deduce from non-verbal behavior how satisfied

[4]Hobson et al. (2012) show that vocal dissonance of CEOs is positively associated with the likelihood of financial misreporting.

[5]See the article of Insider: https://www.businessinsider.com/who-was-the-protester-who-got-into-the-ecb-and-glitter-bombed-mario-draghi-2015-4 (Last access: August 4, 2022). Also, the article of Bloomberg: https://www.bloomberg.com/news/articles/2017-05-10/draghi-stays-calm-on-stimulus-as-dutch-warn-of-risks-with-tulip#xj4y7vzkg (Last access: August 4, 2022).

[6]See the article of Independent: https://www.independent.ie/business/world/angry-draghi-fights-back-at-german-ecb-critics-34648966.html (Last access: August 4, 2022)

[7]The monetary policy accounts, which the ECB has been publishing since February 2015, provide only

the president is with monetary policy decisions, given the economic outlook and the different views within the council about how monetary policy should be conducted (Brunnermeier et al., 2016).

3. Methodology

To study the effect of the emotion of the ECB president on government bond yields and spreads, we adopt an event study approach. During the entire presidency of Mario Draghi, the ECB released monetary policy decisions in a press release at 13:45 CET. At 14:30 CET, the press conference begins with the president reading a prepared introductory statement (IS) and then provides a Q&A session for journalists. By looking at yield changes a few minutes after the end of the press conference, we estimate an unbiased and direct causal effect (Altavilla et al., 2019). Furthermore, we ensure that our estimations are not affected by new monetary policy decisions since financial markets are already aware of them following the press release.

To estimate the effect of emotion and language on government bond yields during the press conference, we estimate the following regression model:

$$\begin{split} y_t = \beta_0 + \beta_1 * Voice_t \times PositivityAN_t + \beta_2 * Voice_t + \beta_3 * PositivityAN_t + \\ \beta_4 * TextComplexity_t + \beta_5 * \Delta PositivityIS_t + \sum_{i=6} \beta_i X_{ti} + \epsilon_t \end{split} \tag{1}$$

y_t is either the change in yield in the case of German bonds or the change in the spread in the case of French, Italian, and Spanish bonds with two, five, or ten-year duration.[8] $Voice_t$ is the

an overview. See: https://www.ecb.europa.eu/press/accounts/html/index.en.html (Last Access: August 1, 2022).

[8]We use data of the *press conference window* of the Euro Area - Monetary Policy Event-Study Database (EA-MPD) from Altavilla et al. (2019). For further information about how to derive the asset yields, we refer to the second section of their paper and the appendix of their study.

net vocal sentiment that we derive quantitatively from the vocal emotions of the president (see section 3.1). An increase in $Voice_t$ implies more positive vocal emotions. $PositivityAN_t$ is the net positivity of the individual answers that one can consider as the textual analog to vocal emotions, and $TextComplexity_t$ measures the average complexity of all answers during a press conference (for the measurement, see section 3.2). Furthermore, we include an interaction term consisting of the vocal and textual sentiment of the answers to account for potential non-linear effects arising from the interplay of voice and words. As a control variable, we include $\Delta PositivityIS_t$ to measure the change in net textual sentiment of the IS to the previous press conference. Furthermore, we include monetary shock variables, and control variables to include other forward-looking indicators such as forecasts of the ECB/Eurosystem staff ($\sum_{i=6} \beta_i X_{ti}$).

3.1. Measuring Vocal Sentiment

3.1.1. Design of the Speech Emotion Recognition Model

Measuring the vocal emotions of the ECB president is the most challenging task. To quantify the emotions and generate a numerical variable for our event regression estimation, we utilize methods developed in SER, a sub-area of machine learning (Pérez-Espinosa et al., 2022). The ultimate aim of SER is to accurately identify emotions based on the voice of the person of interest without directly considering the language used. Recently, economists have started to utilize SER to analyze the vocal sentiment of the Fed chair to estimate the effects on asset prices (Gorodnichenko et al. (2020); Alexopoulos et al. (2022)). The study closest to ours is Gorodnichenko et al. (2020), in which the authors implement a Convolutional Neural Network (CNN) to classify the vocal emotions of Ben Bernanke, Janet Yellen and Jerome Powell during press conferences of the Fed. In contrast to Gorodnichenko et al. (2020), we utilize a Fully Convolutional Neural Network (FCN) based on García-Ordás et al. (2021), which generates a higher Out-of-Sample accuracy[9] and has clear advantages when measuring emotions during Q&A sessions, which are characterized by answers with high variability in length.

García-Ordás et al. (2021) introduce a FCN model framework that can process non-fixed length audio files and classify the audio data underlying vocal emotions.[10] The model architecture of our FCN model consists of three convolutional layers: the first contains 64 filters and a kernel size of (11, 7); the second also contains 64 filters and a kernel size of (7, 11); the number of filters in the third layer equals the number of emotions in this classification task. The first two layers use ReLu activation functions, while we use a Dropout layer following the third layer to prevent overfitting. A central part of the model is the GlobalAveragePooling layer that reduces the amount of data by extracting a unique value from each filter that corresponds to the average value of filter weights and enables the model to process audio files without a priori fixing the length of the data (García-Ordás et al., 2021). The model generates a feature map for each corresponding category of the classification and then uses a Softmax activation function as the final step. We construct the model using Keras on Python.[11]

The strength of the FCN model introduced by García-Ordás et al. (2021) lies in its ability to process non-fixed audio data without a priori cutting or averaging the acoustic features of the data.[12] This represents a significant scientific and technical contribution, which we utilize to examine whether emotions plays a role in communicating European monetary

---

[9]The FCN model generated state-of-the-art Out-of-Sample accuracy at the time the study was published.

[10]In this section we provide a brief description of the FCN model and offer a more detailed description and explanations of the individual parts of the neural network in the Appendix. Nonetheless, we also refer the interested reader to the original article by García-Ordás et al. (2021).

[11]Appendix A provides a visualization of our FCN model

[12]Gorodnichenko et al. (2020), average the acoustic features due to the varying length of the answers. However, this procedure may lead to the loss of valuable information.

policy.

3.1.2. Model Training and Validation

Following the literature on SER (García-Ordás et al., 2021) and similar to Gorodnichenko et al. (2020), we train and validate our model framework using prepared and labeled emotions using the Ryerson Audio-Visual Database of Emotional Speech and Song (RAVDESS) and the Toronto Emotional Speech Set (TESS). RAVDESS offers 1440 vocal speech emotion audio files generated by 24 actresses and actors (12 female and 12 male) reading two statements with eight different emotions. These emotions are *Neutral, Calm, Happy, Sad, Angry, Fear, Disgust, and Suprised,* and they are available in two different intensities (*normal emotional intensity* and *strong emotional intensity*)[13] (Livingstone and Russo, 2018). TESS contains speech emotion data generated by a young and an old actress, who speak 200 different words with seven different emotions: *Neutral, Happy, Sad, Angry, Fear, Disgust, and (pleasantly) Surprised.* These emotions are considered the "basic emotions" in the neuroscientific literature (Bear et al., 2015, pp. 626 - 628) and are widely used in SER tasks (Pérez-Espinosa et al., 2022).

We follow Gorodnichenko et al. (2020) and remove the emotions *Fear* and *Disgust* due to the low probability that these emotions appear during a central bank press conference.[14] Otherwise, we use all emotions and combine both data sets to a combined emotion set for training and validation of our FCN, that is, we use six emotions for our classification task: *Neutral, Calm, Happy, Sad, Angry, and (pleasantly) Surprised.*[15]

[13]For *Neutral* a strong emotional intensity is not available for obvious reasons.

[14]At the beginning of the press conference on 15 April 2015, an activist jumped on the table and disturbed the press conference of the ECB. One could think that this intervention may lead the president to show some fear. Nonetheless, even then, Draghi remained calm such that a journalist even asked: *"[...] And maybe allow me a little add-on question, but due to this intervention of the activist at the beginning of the press conference I observed that you remained very calm. How do you manage this?"*

[15]We also remove, from RAVDESS, sad emotions with high intensity, that is, basically very sad or even crying. After a manual inspection of all Q&A sessions, we can state that Mario Draghi did not cry.

To engineer appropriate features from our emotion set for the FCN, we use the Python package *Librosa* and extract the first 100 Mel-Frequency Cepstrum Coefficients (MFCC) from each audio file.[16] The audio files are not processed or cut in any way. In general, the literature on audio analysis uses different acoustic features, though users decide based on classification accuracy.[17]

For the training and validation of the FCN, we split the emotion set into a training and validation set containing 80% and 20% of the emotions, respectively. To ensure that our model is generalized, that is, the classification is unaffected by the random distribution of emotions into the training and validation set, we use Monte-Carlo Cross Validation (MCCV). We generate seven different training and validation sets, train and validate seven independent FCN models, and combine the models with the highest Out-of-Sample accuracy into a model ensemble. To avoid overfitting, we use an "Early Stopping" - Callback.[18] Table (1) shows the Out-of-Sample accuracy of each FCN model and the average accuracy, that is, the accuracy of our model ensemble.[19]

- Table (1) around here -

3.1.3. Using ECB Press Conference Data

Now that we have a prepared model framework that can classify emotions based on voice, we can utilize our model to quantify the vocal emotions expressed by the ECB president. We download the audio data for all press conferences from the ECB website that also contain

[16] In Appendix B, we explain how we extract MFCCs.

[17] García-Ordás et al. (2021) use and compare Mel Spectrograms with MFCC and conclude that MFCC outperforms the former features regarding classification accuracy. Currently, the literature on speech processing considers that MFCC contains the best characteristics representing the human voice. Nonetheless, a disadvantage of MFCC is their sensitivity to noise (Pérez-Espinosa et al., 2022). Due to this weakness, we cautiously avoid noise and background voices when preparing the actual audio data of Draghi's voice.

[18] We use an Intel Core i7 CPU with 2.3 GHz. The whole training process of the ensemble needs about 16 days.

[19] Using a model ensemble ensures generalization and reduces the overall classification variance.

the Q&A session (ECB, 2022).[20] Next, we consider only the audio data[21] and remove the IS, the questions of the journalists, answers by the vice president, and interventions by the moderator such that only the answers of the ECB president during the Q&A session remain.

Next, we identify the different answers the ECB president gives to journalists. During the Q&A session, journalists can ask two or even three questions. An obvious way to proceed would be to define an answer as the time when Draghi starts to answer a question until he stops talking.[22] However, Draghi usually answers several questions in a row or uses the opportunity to summarize the debates during the ECB governing council meetings and talks about several different topics.[23] Based on such a long answer, we would only classify average emotions and thereby lose the information about emotions that briefly appear during answers to other issues the ECB president provides to the different questions from journalists. To solve this problem, we exploit an interesting characteristic in the written transcripts of the ECB press conferences.[24] The ECB staff already identifies the president's individual answers and identifies them in separate paragraphs of the transcripts. Therefore, we follow the structure of the ECB press conference transcripts and cut all audio files manually so that the voice in each audio file is identical to the respective paragraph in the transcripts for all press conferences.[25] The result is a novel data set consisting of synchronized voice and

[20]Following a formal request, the ECB provide us written confirmation to use their publicly available video and audio data for research.

[21]Using the video material on the Q&A sessions, it is possible to generate emotions based on facial expressions (Curti and Kazinnik, 2021; Alexopoulos et al., 2022). In this study, we focus on vocal emotions and keep the analysis of facial expressions open for future research.

[22]This is the way Gorodnichenko et al. (2020) define the answers of the Fed chair. Nonetheless, this may not be a good choice for the ECB president both because of the possibility of multiple answers and the presence of the ECB Vice-President.

[23]To clarify this, we provide examples in the Appendix.

[24]https://www.ecb.europa.eu/press/pressconf/html/index.en.html (last access: 1 August, 2022)

[25]In Appendix C, we provide an explanation and example of our approach.

language data for future research, which is another contribution of our study.[26] Our voice data consists of 2,336 answers in 71 press conferences between 1 May 2012 and 31 October 2019.[27]

Next, we label the emotions *Happy* and *(pleasantly) Surprised* as **Positive**, *Angry* and *Sad* as **Negative**, and *Neutral* and *Calm* as **Neutral**.[28] We use the voice data of the ECB president, so our model ensemble classifies all answers regarding the underlying vocal sentiment,[29] and we calculate the net vocal sentiment for the whole press conference as follows:

$$Voice_t = \frac{Positive_t - Negative_t}{Positive_t + Negative_t} \tag{2}$$

$Positive_t$ measures the number of positive answers and $Negative_t$ the number of negative answers during the Q&A session of the ECB press conference at t.[30]

- Table (2) and (3) around here -

Tables (2) and (3) show the development of the net vocal sentiment during all press con-

[26]We have permission from the ECB confirming that the audio data are not confidential, and the rights of use are publicly available.

[27]Nonetheless, due to bad audio quality we have to remove the following ten press conferences from the sample: 4 October 2012, 2 May 2013, 2 October 2013, 4 December 2014, 22 October 2015, 2 June 2016, 8 September 2016, 8 December 2016, 8 June 2017, and 6 June 2019.

[28]Aggregating the emotions in this way has the additional advantage that it improves the classification precision of our FCN model ensemble, due to the similarity of the acoustic features of the emotions aggregated (García-Ordás et al., 2021). Our approach is also consistent with Gorodnichenko et al. (2020) and explicitly considers *calm* emotions.

[29]We use the first six FCN models to classify an answer. If the models disagree, we use the seventh FCN model for the final decision. In only seven responses, the whole ensemble disagrees, so we do not consider these answers

[30]We use this quantitative measure of vocal emotions for the main analysis. In addition, we provide in Appendix E a robustness check using a qualitative definition of vocal emotions. Our conclusions remain unchanged.

ferences in our sample.[31] It is plausible that the emotion of the ECB president is strongly affected by monetary policy decision-making, the debates inside the governing council, and the press conference that follows.[32]

During the European Sovereign Debt Crisis (ESDC),[33] the voice sentiment is continuously negative, and one can interpret this to be the result of pressure and stress during a crisis management period.[34] Notice the less negative vocal sentiment for the press conference on 2 August 2012, which is the most positive moment during the crisis and is observable a few days following Draghi's famous "Whatever it Takes" speech which is considered a turning point during the ESDC.[35] Following the end of the ESDC, a temporary increase in vocal sentiment is observable before it becomes again more negative during most of 2014, a challenging year for the ECB Governing Council due to the environment of low inflation and economic growth, increasing financial fragility and risks of deanchoring inflation expectations (Hartmann and Smets (2018); Rostagno et al. (2021)). The introduction of the Asset Purchase Programme (APP) goes along with a more positive vocal sentiment, possibly due to Draghi's success in pushing through unconventional monetary policies despite the controversy surrounding of the policy inside the Governing Council (Brunnermeier et al., 2016). The decline in the average vocal sentiment is observed again during 2018, a time of increasing challenges

---

[31] We remove the press conferences on 22 January 2015 as we consider this observation an outlier due to the high number of neutral answers (see table (3)). It is also the first time the ECB publishes an account of governing council meetings.

[32] Mayew and Venkatachalam (2012) provide an interesting review of emotions in the psychology literature and emphasize the role of social and interpersonal communication and events that trigger emotions and influence a person's affective state.

[33] We utilize the crisis dates of Hartmann and Smets (2018).

[34] As Bernanke (2015) makes clear in his review of the 2008/2009 GFC, enormous psychological stress accompanies crisis management in a financial crisis.

[35] Despite his efforts during the press conference on 2 August 2012, Draghi was heavily criticized for not delivering as much as the markets were expecting, see The Guardian: https://www.theguardian.com/business/blog/2012/aug/02/eurozone-crisis-live-markets-await-ecb-decision (Last access: 4 August 2022). Following this criticism, our model measures a very negative vocal sentiment for the following press conference on 6 September 2012.

(Draghi, 2018) and reaches a new low when the ECB restarts its QE program only a few months after the Governing Council started the beginning of an attempted exit.

3.2. Measuring Textual Sentiment

Methods from the area of NLP are an established part of the methodological toolkit economists use to analyze central bank communication (Bholat et al., 2015; Benchimol et al., 2022). Furthermore, economists utilize these methods to detect changes in the language used by central bankers. Until now, the literature primarily focuses on the analysis of the IS due to the high relevance of this document accompanying the announcement of monetary policy decisions (Picault and Renault, 2017; Schmelling and Wagner, 2019; Baranowski et al., 2021). However, the IS is a carefully crafted text, and it would be naive to assume that it contains unscripted information.[36] In contrast, less research has focused on the role of unscripted information on asset prices that may appear during the Q&A session directly following the IS. While in the previous section, we describe the measurement of vocal sentiment, we also aim to capture communication using the language of the ECB president: the tone or framing and the complexity of the answers.

First, we quantify the net positivity or language sentiment of the answers of the ECB president, that is, we classify whether an answer is creating a more positive or negative framing. Following Curti and Kazinnik (2021), we use Fin-BERT,[37] a deep learning method from the area of NLP (Yang et al., 2022) that goes beyond simple word counting and can identify the sentiment of economic and financial language of even sophisticated sentences within a context.[38] The analysis happens at the individual sentence level, and we classify

[36]Focusing on the Bank of Canada, Ehrmann and Talmi (2020) provide evidence that even changes in the similarity of the press release has a significant effect on volatility. Nonetheless, these documents also cannot be considered unscripted information.

[37]For the implementation, we use the Transformers package on Python (Wolf et al., 2020).

[38]Manela and Moreira (2017) provide evidence that automated methods are increasingly superior to

each sentence within an answer as either *positive, negative,* or *neutral.* If we count more positive than negative sentences, we classify the answer as on balance positive and vice versa. If an answer only consists of neutral sentences, we classify the answer as neutral. We use the same formula as for the calculation of the vocal sentiment to derive the textual sentiment for each press conference:

$$PositivityAN_t = \frac{Positive_t - Negative_t}{Positive_t + Negative_t} \tag{3}$$

$Positive_t$ measures the number of positive answers and $Negative_t$ the number of negative answers during the Q&A session of the ECB press conference at t regarding the language.[39]

Next, we consider the complexity or clarity of the language used by the ECB president. The clarity of the language used significantly affects media attention and engagement (Ferrara and Angino, 2021) and the real-time trading activity level.[40] To capture this important aspect of verbal communication, we follow the literature and use text complexity indicators (Ferrara and Angino, 2021). For the main analysis, we use the Flesch-Kincaid (F-K) Grade Level (Kincaid et al., 1975) that estimates the difficulty of text using a score that approximates the number of years of education a person needs, on average, to understand the content:

$$F-K\ Grade\ Level = 0.39 \times \left(\frac{Total\ Words}{Total\ Sentences}\right) + 11.8 \times \left(\frac{Total\ Syllables}{Total\ Words}\right) - 15.59 \tag{4}$$

lexicographic methods. Yang et al. (2022) show that Fin-BERT has a higher classification precision than alternative machine learning methods, for example, support vector machines.

[39]Table (2) and (3) provide the indicator for each press conference.

[40]Hayo et al. (2022) use high-frequency data for European stock index futures trading between 2009 and 2017 and demonstrate that less clear language reduces trading activity and directs the focus of traders on the Q&A session.

We calculate the F-K Grade Level for each answer i individually and average over all N answers of the Q&A session in t:[41]

$$TextComplexity_t = \frac{1}{N} * \sum_{i=1}^{N} F - K\ Grade\ Level_i \tag{5}$$

For the robustness analysis, we calculate and use the Gunning FOG Index, which considers explicitly *complex words*, analogously.[42]

3.3. Control Variables

Exploiting and combining the timely structure of the ECB press conference with high-frequency data ensures that our regression estimations remain robust to endogeneity problems (Altavilla et al., 2019), a critical issue in identifying monetary policy effects. Nonetheless, we include additional control variables to control for several aspects of the press conference.

To control for the content of the IS, we use Fin-BERT on the sentence level and calculate the net sentiment of the IS using the same formula as in equation (3), just with the individual sentences instead of answers. Next, we calculate the difference between the net sentiment of the IS at the press conference in t and $t-1$:

$$\Delta PositivityIS_t = PositivityIS_t - PositivityIS_{t-1} \tag{6}$$

In doing so we explicitly focus on the textual changes of the IS, which the literature considers essential for understanding the reactions of markets due to the otherwise high similarity

---

[41] To ensure that short answers do not bias this indicator, we remove all answers that consist of only one sentence.

[42] We provide the estimation results when using the FOG indicator instead of the F-K Grade Level in Appendix E. Our conclusions remain unchanged.

of the content of this document over time (Hubert and Labondance, 2021). In practical terms, positive (negative) $\Delta PositivityIS_t$ means that the ECB staff positively (negatively) rephrased existing sentences, included new positive (negative) sentences, or removed previous negative (positive) ones.

To control for possible news that Draghi may reveal along with the quarterly ECB / Eurosystem staff projections, we include the difference between the newly published Next Calendar Year (NCY) forecast and the previous NCY forecast for inflation and real GDP growth. To take into account potential asymmetries between positive and negative forecast news, we include variables $\Delta Inflation_t^{NCY,\, Positive}$ and $\Delta RGDP_t^{NCY,\, Positive}$ which equals the difference between the inflation or real GDP forecasts in t and $t-1$ in case the difference is positive and is otherwise zero. Analogously, we include $\Delta Inflation_t^{NCY,\, Negative}$ and $\Delta RGDP_t^{NCY,\, Negative}$ which equals the difference between the inflation or real GDP growth forecast in t and $t-1$ in case the difference is negative and is otherwise zero.

We also control for monetary shocks and follow the identification strategy of Altavilla et al. (2019) and Swanson (2021) by using factor analysis to extract the monetary shocks. We use the interest rate data from the *press conference window* of the EA-MPD for the time until the end of Draghi's presidency and calculate, in line with Altavilla et al. (2019), three monetary shocks: *Timing, FG,* and *QE.*

4. Empirical Analysis

For the empirical analysis, we focus on the period from July 2013 until October 2019. Therefore, we consider all press conferences since the formal introduction of FG (Hartmann and Smets, 2018) until the end of Draghi's presidency. Parle (2022) argues that the formal introduction of FG led to a decline in the amount of surprising information during press conferences and reduced the informational advantages of the ECB.[43]

4.1. Effects on German Yields

First, we analyze whether the interplay of vocal emotions and language affects the yield of German government bonds, which investors consider as the benchmark asset in the euro area (Altavilla et al., 2019). Therefore, we estimate the following regression model, based on equation (1):

$$\begin{aligned} y_t^{DE} = \beta_0 + \beta_1 * Voice_t \times PositivityAN_t + \beta_2 * Voice_t + \beta_3 * PositivityAN_t + \\ \beta_4 * TextComplexity_t + \beta_5 * \Delta PositivityIS_t + \sum_{i=6} \beta_i X_{ti} + \epsilon_t \end{aligned} \tag{7}$$

where y_t^{DE} is the change in the German yield of government bonds with either two-, five-, or ten-year duration. We test the following first hypothesis:

I: *The interplay of vocal emotions and language during the Q&A session influences the yield of German government bonds. Positive emotions raise yields; negative emotions reduce yields.*

Table (4) shows the estimation results for the yields of German bonds with varying duration:

- Table (4) around here -

We estimate significant effects of the textual variables for the IS ($\Delta PositivityIS_t$) and the Q&A session ($PositivityAN_t$) on the yield of two-year bonds (see column (1) of table (4)). Indeed, we estimate a significant positive effect of an increase in the net positivity of the IS relative to the previous press conference. One explanation could be that financial markets perceive the intended positive rephrasing of the IS, relative to the previous press conference,

[43] Our available data for the period May 2012 until June 2013 covers the ESDC (Hartmann and Smets, 2018). However, as this consists of a crisis period we avoid additional complication that can arise due to the possibility of a structural break. To avoid these potential problems, we focus on the non-crisis period.

and interpret this as a reflection of greater satisfaction in the Governing Council with current and anticipated economic developments and a higher likelihood of monetary tightening.

Consistent with this interpretation, we further estimate a significant positive effect of unscripted communication on yields. Even without considering the role of vocal emotions ($Voice_t$), an increase in the net positivity of the language used by the president to answer the questions leads to an increase in yields. As before, one explanation is that financial markets perceive the more positive framing of the language and associate it with higher satisfaction about underlying economic conditions so that future tightening becomes more likely. Interestingly, vocal emotions amplify this relationship. Figure (1) provides margin plots for the interaction term such that one can measure the marginal effects of vocal emotions given a fixed level for language ($PositivityAN_t$):[44]

- Figure (1) around here -

More positive vocal emotions amplify the effect of the language and lead to even higher yields. The voice is an important complement to the words and the framing that the president uses. Nonetheless, we estimate insignificant effects of the voice and language variables at the longer end of the yield curve (i.e., terms of five and ten years).

Finally, we note that FG_t and QE_t raise German yields at all maturities. Since positive values of the monetary shock variables are scaled to be restrictive monetary policy, the results are consistent with expectations assuming that these instruments are intended to reduce all yields in the euro area. Inflation expectations have an asymmetric impact on German yields; they raise yields at the short-end of the yield curve (i.e., two years) and reduce yields of the ten-year term. Negative inflation news signals higher expected inflation, while positive news heralds the opposite. Both outcomes are as theory predicts how authorities should conduct

[44]For illustration, we choose levels of 0.25 and -0.25 for $PositivityAN$ in the marginal plots since this variable has a mean of -0.07 and a standard deviation of 0.27. Our results remain robust when using different values than 0.25 and -0.25.

monetary policy.

4.2. Effects on Spreads

Next, we analyze the influence of vocal emotions and language on bond yield spreads vis-ā-vis France, Italy, and Spain. The importance of monetary policy in ensuring low spreads has been discussed at least since Draghi's famous "Whatever it Takes" speech (Brunnermeier et al., 2016). This results in the question of whether the unscripted communication during the Q&A session affects bond spreads and to what extent vocal emotions contribute to altering yield spreads. We subtract the yields of French, Italian, and Spanish bonds from the yield of German bonds with identical duration to obtain yield spreads:

$$\overline{y}_t^C = y_t^C - y_t^{DE},\ with\ C \in \{FR, IT, ES\} \tag{8}$$

We use the spread $\overline{y}_t^C$ as the dependent variable and estimate the following regression model based on equation (1):

$$\begin{gathered}\overline{y}_t^C = \beta_0 + \beta_1 * Voice_t \times PositivityAN_t + \beta_2 * Voice_t + \beta_3 * PositivityAN_t + \\ \beta_4 * TextComplexity_t + \beta_5 * \Delta PositivityIS_t + \sum_{i=6} \beta_i X_{ti} + \epsilon_t,\ with\ C \in \{FR, IT, ES\}\end{gathered} \tag{9}$$

Where $\overline{y}_t^C$ is the change in the spread for French, Italian, or Spanish bonds with either two-, five-, or ten-year duration. We then test a second hypothesis:

II: *The interplay of vocal emotions and language during the Q&A session influences the spread of French, Italian, and Spanish government bonds vis-ā-vis German yields. Positive vocal emotions raise yield spreads; negative vocal emotions reduce them. Since the ECB sets monetary policy for the entire euro area, a positive emotion signals economic* *conditions one expected to improve in core countries.*

Table (5) provides the results. We report a significant effect of the interaction term ($voice_t \times PositivityAN_t$) on the spreads of government bonds for nearly all specifications.

- Table (5) around here -

Furthermore, in the case of French and Spanish bonds with five and ten-year duration, we estimate a significant positive effect of the change in the framing of the IS on the spreads, that is, a positive change in the statement leads to a larger spread while a negative rephrasing leads to a declining spread. To investigate the relevance of the complex interplay of vocal emotions and language, we use marginal plots to measure the effects of vocal emotions given a specific level of positivity in the language used during the Q&A session.[45] This permits us to identify more subtle influences on yield spreads than is provided by the coefficient estimates in table (5). Figure (2) shows the marginal effect of vocal emotions on the bonds spreads with two-year duration:

- Figure (2) around here -

The interplay of vocal emotions and language during the Q&A session has a significant and asymmetric influence on the spread of Italian bonds. The combination of negative vocal emotions and negative language leads to an increased spread. One explanation could be that investors consider negative unscripted communicative signals, particularly when searching for potential future risks. This result is consistent with Mayew and Venkatachalam (2012), who also demonstrate that investors value negative vocal emotions more than positive ones when pricing risks. Next, figure (3) visualizes the marginal effect of vocal emotions on spreads of government bonds with a five-year duration:

[45] For illustration, we choose levels of 0.25 and −0.25 for $PositivityAN$ in the marginal plots since this variable has a mean of −0.07 and a standard deviation of 0.27. Our results remain robust when using different values than 0.25 and −0.25.

- Figure (3) around here -

The interplay of vocal emotions and language has a more versatile effect on the spread of five-year government bonds. In the case of French bonds, contrasting vocal and verbal signals lead to a decline in the spread ranging from 0.5 to 1.0 basis point. In contrast, negative language and vocal emotions, and vice versa, do not affect the spread. We find a similar relationship for Spanish bonds. When the ECB president's language is negatively framed while his voice communicates positive signals, the spread declines. It would appear that Draghi's vocal emotions outweigh the rise in the spread generated by the language used alone. In the case of Italy, the results are similar as in Figure (2). Consistent negative communication, that is, negative voice and language, leads to an increase in the spread. Finally, Figure (4) visualizes the marginal effect of vocal emotions on the spreads of ten-year bonds:

- Figure (4) around here -

We find a statistically significant negative effect of the ECB president's communication on the French spread in case of a positive verbal and a negative vocal sentiment, consistent with the findings for bond spreads with lower duration. Furthermore, ten-year Italian bonds also react differently, depending on the degree to which voice and language signals conflict with each other. A positive-positive combination of voice and language and a negative-negative combination leads to an increase in the spread, while more conflicting signals reduce it. In contrast, we cannot reject the null hypothesis that the interplay of vocal emotions and language influences the spread concerning Spanish ten-year bonds.

Financial markets clearly take note of the ECB President's vocal emotions and draw conclusions for each country. While the bond spread declines for France in case of conflicting vocal and verbal signals and increases for Italy in case of negative vocal and verbal framing, Spanish bonds seem to be less influenced by communicative signals from the Q&A session and more by the prepared information from the IS. Nevertheless, we can detect an effect on the bond spreads for all three major economies.

Finally, we estimate negative effects of FG_t and QE_t on bond spreads. The impact of QE_t on bond spreads is insignificant except for five-year Spanish bonds, while FG_t exerts significantly negative effects on short- to medium-term spreads only. Interestingly, greater text complexity also reduces yield spreads for five- and ten-year bonds. Inflation expectations are found to have virtually no impact on spreads but, with exception (five-year spread vis-ā-vis Spain), positive growth expectations exert a strong positive impact on spreads with the largest impact on French bonds. This implies that markets interpret worsening growth expectations as unfavorable to yields in core euro area economies vis-ā-vis German yields.

5. Conclusions

We provide new evidence on the importance of non-verbal communication in the euro area on bond yields. Emotions displayed in press conferences are found to have relevance on yields and yield spreads. Furthermore, we find that there exists an interplay between the language used in the Q&A session by the ECB president and the emotions he displays. Using data from four major economies of the euro area, we estimate a statistically significant impact of non-verbal and verbal communication on yields, with varying effects across different maturities. Equally important, however, we find the impact of vocal emotions to be asymmetric.

In the case of Germany, non-scripted communication has a positive effect on yields. However, this effect is limited to the short end of the yield curve and is asymmetric regarding the kind of vocal emotion. More positive communication signals an increase in German bond yields. The results suggest that investors react to a positive framing of the language by expecting a future monetary policy tightening. Similarly, we report an effect of unscripted communication on the spread of French, Italian, and Spanish bonds. In the case of France and Spain, we identify a similar impact of the interplay of vocal emotions and language on the spread. Negative communication has only minor effects on yields while conflicting vocal and verbal signals generate on a net a declining spread. In the case of Italy, we observe a relationship between negative communication signals and increasing spreads. Negative vocal emotions and language during the Q&A session increase Italian yields and provide evidence for the hypothesis that investors process the information for each country individually when pricing bonds, while monetary policy continues to matter significantly for the spread of euro area member states.

Our results also provide some evidence about the role of inflation and growth expectations, FG, and QE in influencing yields and yield spreads when the analysis is conditional on vocal emotions of the ECB President. FG and QE have reduced yields while having a lesser effect on spreads.

These results add to the growing body of literature which finds that vocal cues influences financial markets. Hence, communication *goes beyond just words.* Future research should focus on how financial markets perceive and process vocal cues during crises like the COVID-19 pandemic or the rising inflation since 2021 and how central bankers' emotions can influence asset prices in times of rising economic and geopolitical uncertainty.

Acknowledgments

We are grateful to María Teresa García-Ordás for her valuable comments on replicating her neural network model. Furthermore, we are thankful for valuable comments from Martin T. Bohl, Nicole Branger, Bernd Kempa, Elissa A. M. Iorgulescu, Fiona Höllmann and other participants of the 2022 Mantua Workshop on "Commodity Prices and Central Bank Communication" of the University of Munster. We also thank Niklas Humann for his outstanding research assistance. Part of the research was conducted, while the first author was a visiting scholar at the Balsillie School of International Affairs. He is grateful for their hospitality.

References

Alexopoulos, M., Han, X., Kryvtsov, O., Zhang, X., 2022. *More than Words: Fed Chairs' Communication During Congressional Testimonies*. Working Paper .

Altavilla, C., Brugnolini, L., Gürkaynak, R.S., Motto, R., Ragusa, G., 2019. *Measuring Euro Area Monetary Policy*. Journal of Monetary Economics 108, 162–179.

Apel, M., Blix Grimaldi, M., 2014. *How Informative Are Central Bank Minutes?* Review of Economics 65, 53–76.

Baranowski, P., Bennani, H., Doryn, W., 2021. *Do the ECB's Introductiory Statements Help Predict Monetary Policy? Evidence from a Tone Analysis*. European Journal of Political Economy 66.

Bear, M.F., Connors, B., Paradiso, M., 2015. *Neuroscience: Exploring the Brain.* Fourth, international ed., Jones & Bartlett Learning.

Benchimol, J., Kazinnik, S., Saadon, Y., 2022. *Text Mining Methodologies with R: An Application to Central bank Texts*. Machine Learning with Applications 8.

Bernanke, B., 2015. *The Courage to Act: A Memoir of a Crisis and Its Aftermath.* 1 ed., WW Norton.

Bholat, D., Hansen, S., Santos, P., Schonhardt-Bailey, C., 2015. *Text Mining for Central Banks*. Centre for Central Banking Studies Handbook No. 33.

Blinder, A.S., Ehrmann, M., De Haan, J., Jansen, D.J., 2022. *Central Bank Communication with the General Public: Promise or False Hope?* CEPR Discussion Paper Series: DP17441 .

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, 910–45.

Bohl, M.T., Kanelis, D., Siklos, P.L., 2023. *Central Bank Mandates: How Differences can Influence the Content and Tone of Central Bank Communication.* Journal of International Money and Finance 130, 102752.

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.

Brunnermeier, M.K., James, H., Landau, J.P., 2016. *The Euro and the Battle of Ideas.* Princeton University Press.

Cieslak, A., Schrimpf, A., 2019. *Non-Monetary News in Central Bank Communication*. Journal of International Economics 118, 293–315.

Cochrane, J.H., Piazzesi, M., 2002. *The FED and Interest Rates - A High-Frequency Identification*. American Economic Review 92(2), 90–95.

Curti, F., Kazinnik, S., 2021. *Let's Face It: Quantifying the Impact of Nonverbal Communication in FOMC Press Conferences*. Working Paper Draft: March 18, 2022.

Draghi, M., 2018. *The Outlook for the Euro Area Economy*. Speech by Mario Draghi, President of the ECB, Frankfurt European Banking Congress, Frankfurt am Main, 16 November 2018.

Dybowski, P.T., Kempa, B., 2020. *The European Central Bank's Monetary Pillar after the Financial Crisis*. Journal of Banking & Finance 121, 105965.

ECB, 2022. *Webcasts of ECB Press Conferences*. Retrived from: https://www.ecb.europa.eu/press/tvservices/webcast/html/index.en.html. Last Access: 30 May 2022.

Ehrmann, M., Talmi, J., 2020. *Starting from a Blank Page? Semantic Similarity in Central Bank Communication and Market Volatility*. Journal of Monetary Economics 111, 48–62.

Ehrmann, M., Wabitsch, A., 2022. *Central Bank Communication with Non-Experts - A Road to Nowhere?* Journal of Monetary Economics 127, 69–85.

Feldkircher, M., Hofmarcher, P., Siklos, P., 2021. *What do central banks talk about? A European perspective on central bank communication*. Focus on European Economic Integration Q2, 61–81.

Ferrara, F.M., Angino, S., 2021. *Does Clarity Make Central Banks More Engaging? Lessons From ECB Communications*. European Journal of Political Economy. In Press.

Ferrara, F.M., Masciandao, D., Moschella, M., Romelli, D., 2021. *Political Voice on Monetary Policy: Evidence from the Parliamentary Hearings of the European Central Bank*. European Journal of Political Economy. In Press.

García-Ordás, M.T., Alaiz-Moretón, H., Benítez-Andrades, J.A., García-Rodríguez, I., García-Olalla, O., Benavides, C., 2021. *Sentiment Analysis in Non-Fixed Length Audios Using a Fully Convolutional Neural Network*. Biomedical Signal Processing and Control 69.

Gertler, M., Karadi, P., 2015. *Monetary Policy Surprises, Credit Costs, and Economic Activity*. American Economic Journal: Macroeconomics 7(1), 44–76.

Gorodnichenko, Y., Pham, T., Talavera, O., 2020. *The Voice of Monetary Policy*. Working Paper Draft: June 14, 2022.

Gürkaynak, R.S., Sack, B., Swanson, E.T., 2005. *Do Actions Speak Louder Than Words? The Response of Asset Prices to Monetary Policy Actions and Statements*. International Journal of Central Banking , 55–93.

Hansen, S., McMahon, M., Prat, A., 2018. *Transparency and Deliberation within the FOMC: A Computa-*

tional Linguistics Approach. Quarterly Journal of Economics 133, 801–870.

Hartmann, P., Smets, F., 2018. *The European Central Bank's Monetary Policy during Its First 20 Years*. Brookings Papers on Economic Activity, Fall, 1-146.

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.

Hobson, J.L., Mayew, W.L., Venkatachalam, M., 2012. *Analyzing Speech to Detect Financial Misreporting*. Journal of Accounting Research 50(2), 349–392.

Hoesch, L., Rossi, B., Sekhposyan, T., 2020. *Has the Information Channel of Monetary Policy Disappeared? Revisiting the Empirical Evidence*. Working Paper. Draft: May 2022.

Hu, A., Ma, S., 2020. *Persuading Investors: A Video-Based Study*. Working Paper Draft: 13 July 2021.

Hubert, P., Labondance, F., 2021. *The Signaling Effects of Central Bank Tone*. European Economic Review 133.

Kincaid, J.P., Fishburne Jr., R.P., Rogers, R.L., Chissom, B.S., 1975. *Derivation of New Readability Formulas (Automated Readability Index, Fog Count And Flesh Reading Ease Formula) For Navy Enlisted Personnel*. Research Branch Report 8-75 .

Kuttner, K.N., 2001. *Monetary Policy Surprises and Interest Rates: Evidence from the Fed Funds Future Market*. Journal of Monetary Economics 47, 523–544.

Livingstone, S., Russo, F., 2018. *The Ryerson Audio-Visual Database of Emotional Speech and Song (RAVDESS): A dynamic, multimodal set of facial and vocal expressions in North American English.* PLoS One 5.

Lombardi, D., Siklos, P.L., St. Amand, S., 2019. *Asset Price Spillovers from Unconventional Monetary Policy: A Global Empirical Perspective*. International Journal of Central Banking 15(2), 43–74.

Loughran, T., McDonald, B., 2011. *When Is a Liability Not a Liability? Textual Analysis, Dictionaries, and 10-Ks*. Journal of Finance 66(1), 35–65.

Manela, A., Moreira, A., 2017. *News Implied Volatility and Disaster Concerns*. Journal of Financial Economics 123(1), 137–162.

Mayew, W.J., Venkatachalam, M., 2012. *The Power of Voice: Managerial Affective States and Future Firm Performance*. Journal of Finance 67(1).

Mehrabian, A., 1971. *Silent Messages: Implicit Communication of Emotions and Attitudes.* Belmont, CA: Wadsworth.

Mullainathan, S., Spiess, J., 2017. *Machine Learning: An Applied Econometric Approach*. Journal of

Economic Perspectives 13(2), 87–106.

Nakamura, E., Steinsson, J., 2018. *High-Frquency Identification of Monetary Non-Neutrality: The Information Effect*. Quarterly Journal of Economics 133(3), 1283–1330.

Parle, C., 2022. *The Financial Market Impact of ECB Monetary Policy Press Conferences - A Text Based Approach*. European Journal of Political Economy. In Press.

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.

Pérez-Espinosa, H., Zatarain-Cabada, R., Barrón-Estrada, M.L., 2022. *Emotion Recognition: From Speech and Facial Expressions.* chapter 15. Published In: Torres-García, Alejandro A. and Reyes-García, Carlos A. and Villaseñor-Pineda, Luis and Mendoza-Montoya, Omar (Ebs.): Biosignal Processing and Classification Using Computational Learning and Intelligence. Elsevier.

Ramey, V.A., 2016. *Macroeconomic Shocks and Their Propagration.* chapter 2. Published In: Taylor, John B. and Uhlig, Harald (Ebs.): Handbook of Macroeconomics. Volume 2. Elsevier.

Rigobon, R., Sack, B., 2004. *The Impact of Monetary Policy on Asset Prices*. Journal of Monetary Economics 51, 1553–1575.

Rosa, C., 2011. *The High-Frequency Response of Exchange Rates to Monetary Policy Actions and Statements*. Journal of Banking & Finance 35, 478–489.

Rostagno, M., Altavilla, C., Carboni, G., Lemke, W., Motto, R., Guilhem, A.S., Yiangou, J., 2021. *Monetary Policy in Times of Crisis: A Tale of Two Decades of the European Central Bank.* 1 ed., Oxford University Press.

Schmelling, M., Wagner, C., 2019. *Does Central Bank Tone Move Asset Prices?* Working Paper .

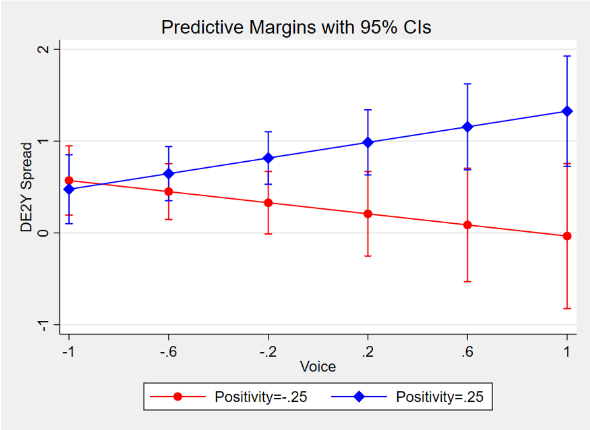
Shapiro, A.H., Wilson, D., 2021. *Taking the Fed at its Word: A New Approach to Estimating Central Bank Objectives using Text Analysis*. Review of Economic Studies .

Swanson, E.T., 2021. *Measuring the Effects of Federal Reserve Forward Guidance and Asset Purchases on Financial Markets*. Journal of Monetary Economics 118, 32–53.

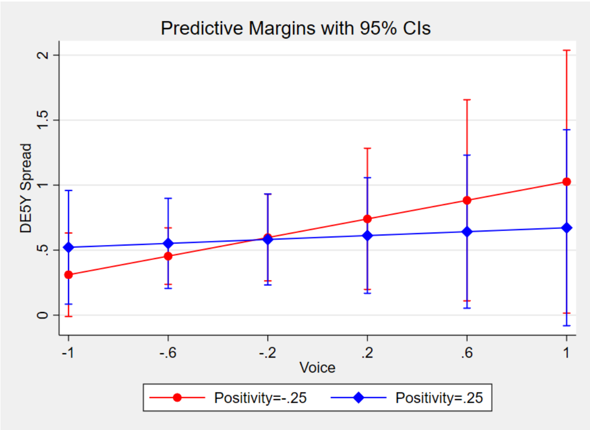
Wolf, T., Debut, L., Sanh, V., Chaumond, J., Delangue, C., Moi, A., Cistac, P., Rault, T., Louf, R., et al., 2020. *Transformers: State-of-the-Art Natural Language Processing*. Proceesings of the 2020 conference on empirical methods in natural language processing: system demonstrations , 38–45.

Yang, Y., Huang, A.H., Wang, H., 2022. *FinBERT - A Large Language Model Approach to Extracting Information from Financial Text*. Working Paper Draft: June 2020 .

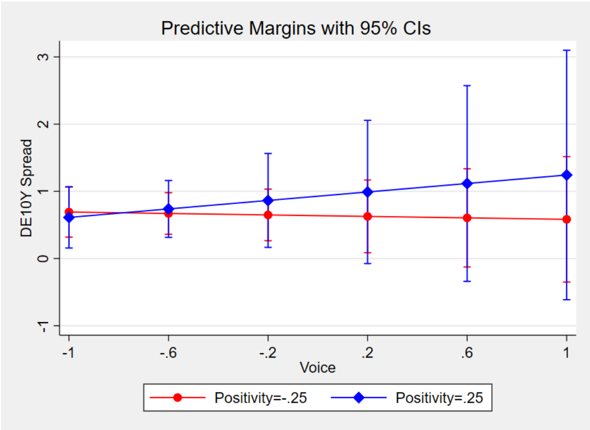
Figure 1: Marginal Effect of Vocal Emotions Given Language on German Yields



(a) Two-Year German Yield



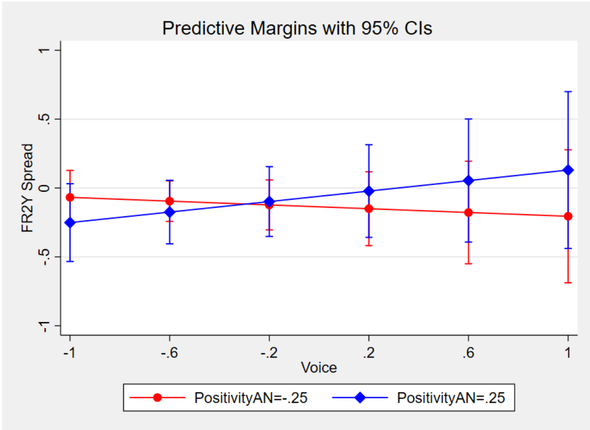
(b) Five-Year German Yield



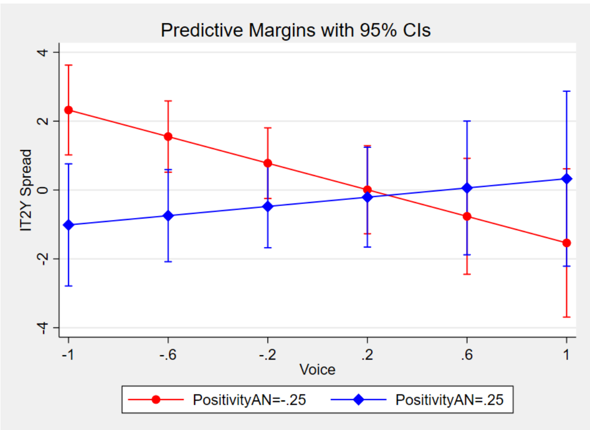
(c) Ten-Year German Yield

Note: These plots visualize the marginal effect of a change in $Voice_t$ given a specific level of $PositivityAN_t$ on the yields of German government bonds. We report the change in spreads in basis points. We illustrate the marginal effect of $Voice_t$ for $PositivityAN_t \in \{-0.25,\ 0.25\}$. We report the estimations in table (5).

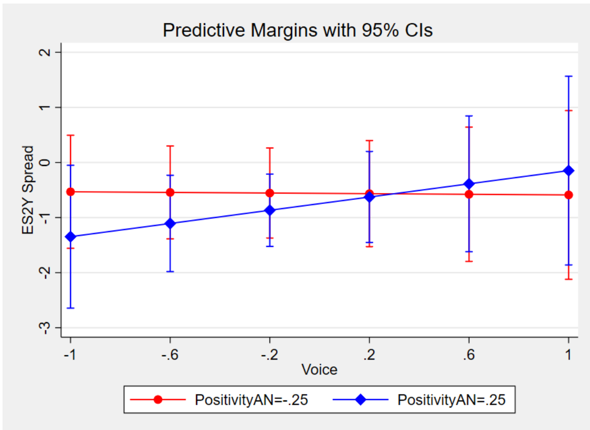
Figure 2: Marginal Effect of Vocal Emotions Given Language on Two-Year Spreads



(a) Two-Year French Spread



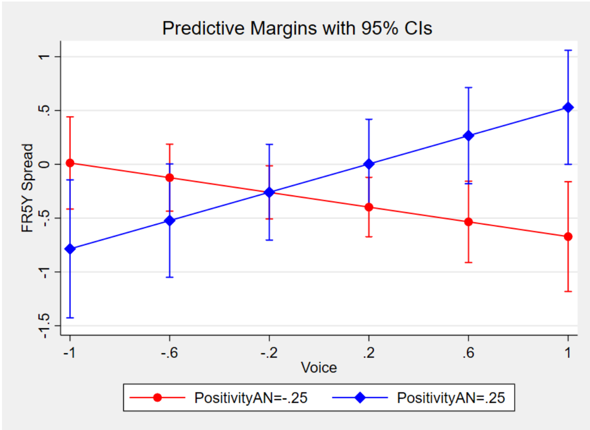
(b) Two-Year Italian Spread



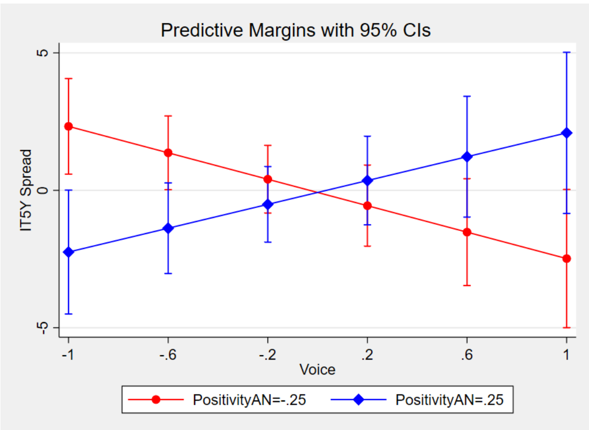
(c) Two-Year Spanish Spread

Note: These plots visualize the marginal effect of a change in $Voice_t$ given a specific level of $PositivityAN_t$ on the spread of two-year government bonds from a) France, b) Italy, and c) Spain. We report the change in spreads in basis points. We illustrate the marginal effect of $Voice_t$ for $PositivityAN_t \in \{-0.25,\ 0.25\}$. We report the estimations in table (5).

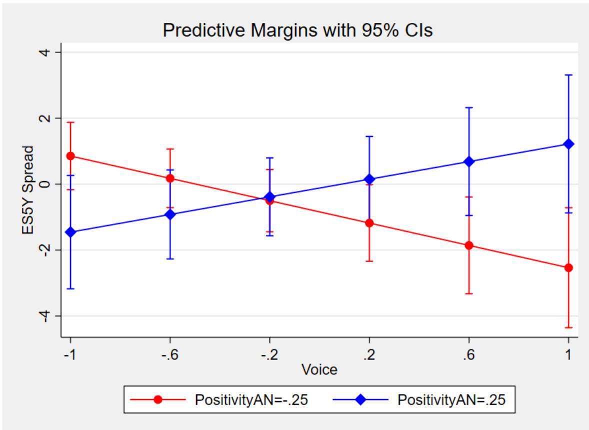
Figure 3: Marginal Effect of Vocal Emotions Given Language on Five-Year Spreads



(a) Five-Year French Spread



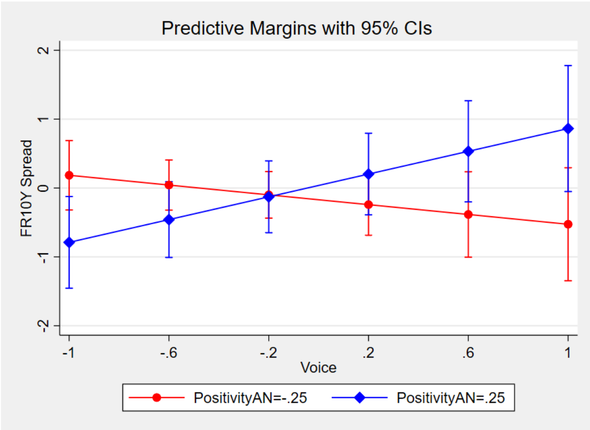
(b) Five-Year Italian Spread



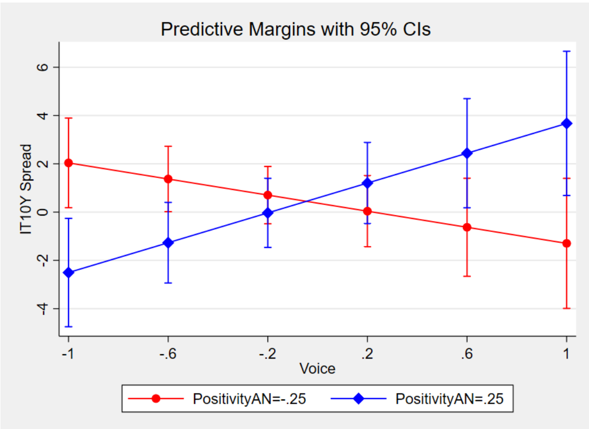
(c) Five-Year Spanish Spread

Note: These plots visualize the marginal effect of a change in $Voice_t$ given a specific level of $PositivityAN_t$ on the spread of five-year government bonds from a) France, b) Italy, and c) Spain. We report the change in spreads in basis points. We illustrate the marginal effect of $Voice_t$ for $PositivityAN_t \in \{-0.25,\ 0.25\}$. We report the estimations in table (5).

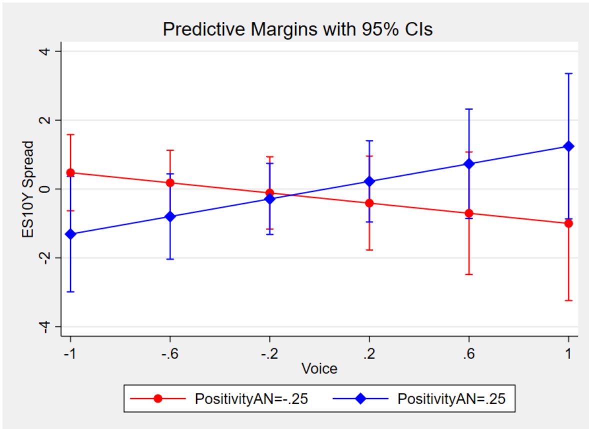
Figure 4: Marginal Effect of Vocal Emotions Given Language on Ten-Year Spreads



(a) Ten-Year French Spread



(b) Ten-Year Italian Spread



(c) Ten-Year Spanish Spread

Note: These plots visualize the marginal effect of a change in $Voice_t$ given a specific level of $PositivityAN_t$ on the spread of ten-year government bonds from a) France, b) Italy, and c) Spain. We report the change in spreads in basis points. We illustrate the marginal effect of $Voice_t$ for $PositivityAN_t \in \{-0.25,\ 0.25\}$. We report the estimations in table (5).

Table 1: **FCN Out-of-Sample Classification Precision**

1. FCN	2. FCN	3. FCN	4. FCN	5. FCN	6. FCN	7. FCN	FCN Ensemble
94.4%	92.9%	92.2%	90.7%	93.1%	93.8%	91.9%	92.7%

Note: This table shows the Out-of-Sample accuracy of each FCN model with the highest precision during the individual model training. The hyperparameters for each training session are identical, while the training and validation set composition is randomly changed to ensure generalization. We calculate the accuracy of the FCN Ensemble as the average accuracy of all individual FCN models.

Table 2: **Vocal and Textual Sentiment (Part I)**

PC Date	Voice: Positive	Voice: Negative	Voice: Neutral	Voice: Net Sentiment	Text: Net Sentiment
03-05-2012	0	25	0	-1.00	-0.44
06-06-2012	1	29	0	-0.93	-0.26
05-07-2012	3	29	0	-0.82	0.37
02-08-2012	13	21	0	-0.24	0
06-09-2012	0	30	0	-1.00	-0.2
09-11-2012	1	27	0	-0.93	0.13
06-12-2012	8	23	0	-0.48	0.00
10-01-2013	1	30	0	-0.94	-0.24
07-02-2013	2	29	0	-0.87	0.5
07-03-2013	0	28	0	-1.00	-0.16
04-04-2013	0	38	0	-1.00	-0.24
06-06-2013	0	27	0	-1.00	-0.22
04-07-2013	3	26	0	-0.79	0.23
01-08-2013	7	17	0	-0.42	-0.13
05-09-2013	15	15	0	0.00	-0.33
07-11-2013	5	17	1	-0.55	-0.07
05-12-2013	0	23	0	-1.00	0.20
09-01-2014	2	26	0	-0.86	0.10
06-02-2014	0	40	0	-1.00	0.00
06-03-2014	2	23	0	-0.84	-0.5
03-04-2014	0	23	0	-1.00	-0.25
08-05-2014	30	6	0	0.67	-0.36
05-06-2014	0	47	0	-1.00	0.16
03-07-2014	7	62	0	-0.80	0.36
07-08-2014	7	63	0	-0.80	-0.12
04-09-2014	0	42	0	-1.00	-0.20
02-10-2014	15	17	0	-0.06	-0.29
06-11-2014	0	34	0	-1.00	0.40

Note: This table presents the number of answers with a positive, negative, and neutral vocal sentiment and the net sentiment, which we calculate using equation (2). For comparison, we provide the net sentiment of the textual sentiment (see 3.2).

Table 3: **Vocal and Textual Sentiment (Part II)**

PC Date	Voice: Positive	Voice: Negative	Voice: Neutral	Voice: Net Sentiment	Text: Net Sentiment
22-01-2015	0	1	46	-1.00	0.06
05-03-2015	23	7	0	0.53	0.30
15-04-2015	42	3	0	0.87	0.03
03-06-2015	31	6	2	0.68	-0.05
16-07-2015	31	2	6	0.88	-0.05
03-09-2015	25	3	0	0.79	-0.10
03-12-2015	29	0	4	1.00	0.44
21-01-2016	31	3	0	0.82	-0.08
10-03-2016	30	1	4	0.94	-0.27
21-04-2016	22	3	1	0.76	0.00
21-07-2016	25	0	1	1.00	0.00
20-10-2016	0	30	1	-1.00	0.00
19-01-2017	15	6	0	0.43	-0.07
09-03-2017	41	4	2	0.82	-0.15
27-04-2017	19	7	8	0.46	-0.16
20-07-2017	1	30	0	-0.94	0.16
07-09-2017	26	5	6	0.68	-0.37
26-10-2017	0	22	8	-1.00	0.33
14-12-2017	20	9	0	0.38	0.47
25-01-2018	31	1	2	0.94	0.05
08-03-2018	33	3	1	0.83	-0.52
26-04-2018	4	15	1	-0.58	-0.29
14-06-2018	36	3	0	0.85	-0.2
26-07-2018	1	26	0	-0.93	0.33
13-09-2018	11	15	7	-0.15	0
25-10-2018	0	16	13	-1.00	-0.3
13-12-2018	0	27	0	-1.00	-0.3
24-01-2019	10	28	1	-0.47	-0.38
07-03-2019	11	20	4	-0.29	-0.30
10-04-2019	1	27	0	-0.93	0.00
25-07-2019	0	32	0	-1.00	-0.79
12-09-2019	0	17	0	-1.00	-0.2
24-10-2019	0	41	0	-1.00	-0.12

Note: This table presents the number of answers with a positive, negative, and neutral vocal sentiment and the net sentiment, which we calculate using equation (2). For comparison, we provide the net sentiment of the textual sentiment (see 3.2).

Table 4: **German Bond Yields**

	(1) DE2Y	(2) DE5Y	(3) DE10Y
$Voice_t \times PositivityAN_t$	1.46**	-0.57	0.74
	(0.68)	(0.85)	(1.30)
$Voice_t$	0.06	0.22	0.13
	(0.13)	(0.17)	(0.24)
$PositivityAN_t$	1.26**	-0.14	0.58
	(0.53)	(0.63)	(1.15)
$\Delta PositivityIS_t$	1.01*	-0.35	-0.59
	(0.58)	(0.73)	(0.79)
$TextComplexity_t$	0.11	0.10	-0.03
	(0.08)	(0.08)	(0.14)
$Timing_t$	0.93***	0.84***	0.13
	(0.18)	(0.17)	(0.19)
FG_t	0.93***	1.11***	0.61***
	(0.12)	(0.13)	(0.12)
QE_t	0.35***	0.89***	1.19***
	(0.10)	(0.10)	(0.08)
$\Delta Inflation_t^{NCY,Positive}$	1.64	-1.67	-4.50*
	(3.71)	(2.46)	(2.65)
$\Delta Inflation_t^{NCY,Negative}$	3.39**	1.04	1.27
	(1.37)	(2.06)	(1.99)
$\Delta RGDP_t^{NCY,Positive}$	-1.77	-0.36	1.35
	(2.51)	(1.91)	(2.43)
$\Delta RGDP_t^{NCY,Negative}$	-1.85	-7.75***	-1.91
	(1.84)	(2.34)	(2.21)
$Constant$	-1.21	-1.03	0.49
	(0.80)	(0.79)	(1.40)
R^2	0.917	0.954	0.947
Obs	48	48	48

Note: We regress our variables on the change in the yield of two, five, and ten-year government bonds of Germany for July 2013 until October 2019. The dependent variable is calculated as the difference in the median price in a narrow time window before the ECB press conference and a narrow time window afterward and is reported in basis points. $Voice_t$ is Draghi's net vocal sentiment during the Q&A session (see 3.1), $PositivityAN_t$ measures the net positivity of Draghi's answers, and $TextComplexity_t$ the average complexity of the answers to journalists during the Q&A session (see 3.2). $PositivityIS_t$ measures the change in the framing of the IS since the last press conference (see 3.3). We use the monetary shocks identified by Altavilla et al. (2019) as additional control variables. Furthermore, $\Delta Inflation^{NCY,Positive}$ ($\Delta RGDP^{NCY,Positive}$) controls for the change in the NCY forecast for inflation (real GDP) from the ECB/Eurosystem staff projections. Analogously, $\Delta Inflation^{NCY,Negative}$ ($\Delta RGDP^{NCY,Negative}$) controls for the effects of negative changes. Robust standard errors in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 5: **Government Bonds Spreads**

	(1) FR2Y	(2) IT2Y	(3) ES2Y	(4) FR5Y	(5) IT5Y	(6) ES5Y	(7) FR10Y	(8) IT10Y	(9) ES10Y
$Voice_t \times PositivityAN_t$	0.52	5.20**	1.26	2.00***	9.15***	6.06***	2.36***	9.51***	4.03*
	(0.51)	(2.55)	(1.89)	(0.55)	(3.26)	(1.99)	(0.85)	(3.32)	(2.28)
$Voice_t$	0.06	-0.63	0.29	0.16	-0.12	-0.18	0.24	0.71	0.27
	(0.09)	(0.45)	(0.33)	(0.15)	(0.52)	(0.39)	(0.19)	(0.58)	(0.43)
$PositivityAN_t$	0.15	-1.47	-0.37	0.40	-0.00	1.45	0.42	0.43	0.46
	(0.42)	(1.92)	(1.24)	(0.54)	(2.24)	(1.87)	(0.73)	(2.21)	(1.84)
$\Delta PositivityIS_t$	-0.20	4.54	2.97	1.60***	4.65	6.92**	1.82**	3.79	5.85**
	(0.40)	(3.12)	(1.96)	(0.55)	(3.12)	(2.62)	(0.71)	(2.88)	(2.36)
$TextComplexity_t$	0.10	-0.46	-1.00***	-0.03	-1.32**	-0.83**	-0.18*	-1.27***	-1.07***
	(0.06)	(0.45)	(0.33)	(0.09)	(0.49)	(0.35)	(0.11)	(0.45)	(0.36)
$Timing_t$	-0.09	-0.64	-0.21	0.03	-0.44	-0.33	0.11	0.05	-0.01
	(0.07)	(0.47)	(0.34)	(0.12)	(0.50)	(0.40)	(0.18)	(0.49)	(0.42)
FG_t	-0.13**	-0.58*	-0.49**	-0.01	-0.78**	-0.67***	0.13	-0.24	-0.30
	(0.06)	(0.29)	(0.19)	(0.13)	(0.30)	(0.22)	(0.16)	(0.34)	(0.29)
QE_t	-0.01	0.22	-0.06	-0.02	-0.10	-0.31**	0.06	-0.03	-0.08
	(0.04)	(0.20)	(0.17)	(0.06)	(0.25)	(0.14)	(0.08)	(0.25)	(0.21)
$\Delta Inflation_t^{NCY,Positive}$	0.23	5.03	1.54	0.30	5.71	7.39*	1.82	5.84	6.40
	(0.86)	(6.92)	(4.96)	(1.48)	(6.05)	(3.92)	(2.81)	(6.28)	(5.18)
$\Delta Inflation_t^{NCY,Negative}$	-1.59	10.71	-0.22	0.53	4.23	5.04	-0.77	4.84	4.05
	(1.18)	(8.08)	(5.10)	(1.50)	(8.55)	(5.11)	(1.83)	(8.30)	(5.45)
$\Delta RGDP_t^{NCY,Positive}$	-0.80	-8.16	-3.27	-0.79	-7.91	-7.97*	-0.88	-8.32	-7.02
	(0.84)	(6.37)	(4.09)	(1.22)	(6.32)	(4.42)	(2.44)	(6.47)	(4.96)
$\Delta RGDP_t^{NCY,Negative}$	3.41**	8.22	0.64	3.72***	11.69	15.05***	4.43**	5.62	5.41
	(1.30)	(6.16)	(4.91)	(1.12)	(7.38)	(5.44)	(1.80)	(7.77)	(5.76)
$Constant$	-1.06*	4.44	9.82***	-0.06	13.35***	7.81**	1.56	13.18***	10.51***
	(0.62)	(4.29)	(3.25)	(0.94)	(4.70)	(3.27)	(1.05)	(4.35)	(3.44)
R^2	0.464	0.410	0.487	0.374	0.468	0.484	0.485	0.441	0.341
Obs	48	47	47	48	47	47	48	47	47

Note: We regress our variables on the change in the spread of two, five, and ten-year government bonds of France, Italy, and Spain for July 2013 until October 2019. We removed the observation on July 2013 for Italy and Spain to avoid an outlier driving the results. The dependent variable is calculated as the difference in the median price in a narrow time window before the ECB press conference and a narrow time window afterward and then subtracted by the yield from a German bond with the same duration and is reported in basis points. $Voice_t$ is Draghi's net vocal sentiment during the Q&A session (see 3.1), $PositivityAN_t$ measures the net positivity of Draghi's answers, and $TextComplexity_t$ the average complexity of the answers to journalists during the Q&A session (see 3.2). $PositivityIS_t$ measures the change in the framing of the IS since the last press conference (see 3.3). We use the monetary shocks identified by Altavilla et al. (2019) as additional control variables. Furthermore, $\Delta Inflation^{NCY,Positive}$ ($\Delta RGDP^{NCY,Positive}$) controls for the change in the NCY forecast for inflation (real GDP) from the ECB/Eurosystem staff projections. Analogously, $\Delta Inflation^{NCY,Negative}$ ($\Delta RGDP^{NCY,Negative}$) controls for the effects of negative changes. Robust standard errors in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$