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

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

June 2024

Abstract

We combine modern methods from Speech Emotion Recognition and Natural Language Processing with high-frequency financial data to precisely analyze how the vocal emotions and language of ECB President Mario Draghi affect the yields and yield spreads of major euro area economies. This novel approach to central bank communication reveals that vocal and verbal emotions significantly impact the yield curve, with effects varying in magnitude and direction. Positive signals raise German and French yields, while negative signals increase Italian yields. Our analysis of bond spreads indicates that positive communication influences the risk-free yield component, whereas negative communication affects the risk premium. Additionally, our study contributes by constructing a synchronized dataset for voice and language analysis.

JEL Classification: E50, E58, G12, G14

Keywords: Artificial Intelligence, Asset Prices, Communication, ECB, High-Frequency Data, Speech Emotion Recognition

[1]Corresponding Author: Deutsche Bundesbank, Wilhelm-Epstein-Str. 14, 60431 Frankfurt am Main, Germany; E-mail: dimitrios.kanelis@bundesbank.de

[2]Department of Economics, Wilfrid Laurier University, Balsillie School of International Affairs, and the Centre for Applied Macroeconomic Analysis (CAMA)

[‡]We thank Martin Bohl, Nicole Branger, Bernd Kempa, Joscha Beckmann, Peter Tillmann, and Jan Szczygielski for their comments and remarks. We also thank seminar and conference participants at various places for their helpful comments and remarks. We are grateful to Niklas Humann for his outstanding research assistance. Part of the research for the paper was conducted while the first author was a visiting scholar at the Balsillie School of International Affairs. He is grateful for the hospitality. This draft is a revised version of the previous one from December 2022. This paper represents the authors' personal opinions and does not necessarily reflect the views of the Deutsche Bundesbank or the Eurosystem.

"[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, former President of the European Central Bank (ECB), spoke financial markets listened. In parallel, psychologists, linguists, and, more recently, economists have begun to understand that financial markets also respond to the emotions displayed by policymakers. Nevertheless, the long-recognized role of emotion and our ability to measure its implications reliably has only been recently investigated.

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]

Over time, central bank communication has evolved to become markedly more uniform, with central bankers keenly focused on minimizing market shocks. This evolution has naturally led financial markets to pay closer attention to less conventional forms of information. Among these, vocal cues stand out as a uniquely spontaneous source of insight, often conveying messages that extend beyond the mere words spoken. Gorodnichenko et al. (2023)

[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).

have demonstrated the significant influence of the Federal Reserve chair's vocal emotions on various asset prices, highlighting the critical role of spontaneous information in U.S. monetary policy communication. While this research sheds valuable light on the subject, it is not without its limitations. In this paper, we introduce and apply an improved methodology that captures more precisely and effectively the nuances and influences of vocal emotions on euro area monetary policy.

We employ modern methods from Speech Emotion Recognition (SER) and Natural Language Processing (NLP), in conjunction with high-frequency financial data. Our aim is to estimate the impact on yields in the four largest euro area economies resulting from the interplay between 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.

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.

To measure vocal emotions, we implement the Fully Convolutional Neural Network (FCN) developed by García-Ordás et al. (2021). This model has the ability to process audio files of varying lengths, allowing us to bypass the need for prior standardization and processing of audio signals, thereby preserving their full informational content. To analyze language framing, which is predominantly unscripted during the Q&A sessions, we use the FinBERT language model. This model, derived from the Bidirectional Encoder Representations from Transformers (BERT) which converts words into numerical representations, is specifically adapted for economic and finance-related texts. Unlike traditional dictionary-based methods, FinBERT offers advanced linguistic analysis capabilities, as demonstrated by Huang et al. (2022).

The results underscore the relevance of vocal emotions in conjunction with language, highlighting significant policy implications for enhancing monetary policy communication. We estimate a significant impact of vocal emotions and language on yields and yield spreads of major euro area economies. For example, we find that for German and French bonds, positive vocal and verbal cues lead to yield increases, whereas this effect is observed only at the short end of the yield curve for German bonds. Conversely, negative vocal and verbal cues result in increased yields for Italian bonds, while no effects are estimated for the German, French, and Spanish yields.

Analyzing yield spreads helps us to understand how vocal emotions influence investor risk behavior. Vocal and verbal cues do not affect the spreads of French and Spanish bonds, while negative signals continue to increase the Italian spread. These findings suggest that positive unscripted communication affect the risk-free interest component, functioning similarly to a conventional monetary policy impulse, whereas negative cues influence the risk premium of individual countries. This finding is in line with the extensive literature that finds differential financial asset markets responses between good and bad news.

Our study makes four significant contributions to the literature. First, to the best our knowledge, we are the first to document the importance of the vocal emotions of the President of the ECB. Second, our study applies deep learning to economic analysis by introducing and using the model developed by García-Ordás et al. (2021), which is part of the FCN model class. This methodological advancement is a crucial step forward in utilizing audio data for research in central bank communication and empirical finance. The model's ability to process non-fixed audio files is essential for accurately evaluating responses of varying durations, such as those during Q&A sessions or earnings calls, and it also allows for real-time classification. Third, we add to a small but growing literature linking behavioural elements in monetary policy communication to financial outcomes. Finally, we construct a novel dataset of synchronized voice and language data, along with 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 existing literature that considers 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 represent an influential approach for analyzing the direct causal effects of monetary policy decisions. Using a narrow time window around an event and utilizing high-frequency data (daily or intra-daily), the causality is identified by disentangling and ordering the sequence of events (Ramey, 2016). The literature has grown considerably since the first seminal studies using this identification strategy (e.g., Kuttner (2001); Cochrane and Piazzesi (2002); Rigobon and Sack (2004)). 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 the "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 on the expected 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. 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 news composition in central bank statements varies considerably, with non-monetary news driving communication between 2008 - 2013 and monetary news after that. 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).

The second strand of literature relevant to our study delves into the functioning and impacts of central bank communication, frequently analyzed through machine learning techniques. Early surveys like those by Blinder et al. (2008), updated recently by Blinder et al. (2024),

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

lay foundational insights. Typical approaches to quantifying qualitative central bank communications include dictionary methods (e.g., Loughran and McDonald (2011); Apel and Blix Grimaldi (2014)) or textual indicators such as complexity or similarity (Ferrara and Angino, 2022), applied to textual data from various central bank outputs 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)), and social media contributions (Ehrmann and Wabitsch (2022)). Notable findings include Hubert and Labondance (2021) showing how the tone of introductory statements can explain monetary shocks and predict future policy decisions, and Schmelling and Wagner (2019) and Parle (2022) who show that the framing of the central bank press conference statement affects stock prices in real-time.

Since press releases following policy rate decisions change little most of the time (Ehrmann and Talmi, 2020), markets also want to look for scripted information that provide 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 investors and that positive (negative) vocal emotions of managers precede positive (negative) news about corporate performance.

In the case of central banking, Gorodnichenko et al. (2023) 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.[3] Studies by Curti and Kazinnik (2023) and Alexopoulos et al. (2023) estimate real-time effects of the chair's facial emotions on stock prices and find them to be important for our understanding of the conduct of monetary policy.

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[4] or his annoyed reaction when a journalist refers to his German critics[5]. 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.[6] One can only deduce from non-verbal behavior how satisfied 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, 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 and then

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

[4]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).

[5]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)

[6]The monetary policy accounts, which the ECB has been publishing since February 2015, provide only an overview. See: https://www.ecb.europa.eu/press/accounts/html/index.en.html (Last Access: August 1, 2022).

provides a Q&A session for journalists. By looking at median yield changes in an interval of ten minutes after the end of the press conference, we estimate an unbiased and direct causal effect (Altavilla et al., 2019).[7] 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:

$$\begin{aligned} y_t = \beta_0 + \beta_1 * Voice_t \times Positivity_t^{AN} + \beta_2 * Voice_t + \\ \beta_3 * Positivity_t^{AN} + \beta_4 * Positivity_t^{IS} + \sum_{i=5} \beta_i X_{ti} + \epsilon_t \end{aligned} \quad (1)$$

y_t represents the change in yield for German bonds of one, two, five, or ten-year maturities, or for French, Italian, and Spanish bonds of two, five, or ten-year maturities.[8] $Voice_t$ is the 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. $Positivity_t^{AN}$ is the net positivity of the individual answers that one can consider as the textual analog to vocal emotions (for the measurement, see section 3.2). Furthermore, we include an interaction term consisting of the vocal and verbal sentiment of the answers to account for potential non-linear effects arising from the interplay of voice and words.We return to this issue below. We further include $Positivity_t^{IS}$ to measure the net textual sentiment of the introductory statement. As additional control variables, we include monetary shock variables and the forecasts of the ECB/Eurosystem staff as forward-looking indicators ($\sum_{i=5} \beta_i X_{ti}$). Further

[7]The yield change happens several minutes following the end of the press conference such that the setting ensures no reverse or simultaneous causality bias (Altavilla et al., 2019).

[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.

details and the rationale for the control variables are provided in section 3.3 below. The estimation period spans from the end of the acute phase of the fiscal crisis in the euro area in July 2013 to the end of Draghi's presidency in October 2019 (see section 4).

3.1. Measuring Vocal Sentiment

3.1.1. Design of the Speech Emotion Recognition Model

Quantifying the vocal emotions of the ECB president presents significant challenges. To generate a numerical variable for our event regression estimation, we employ methods from SER, a specialized subfield of machine learning (Pérez-Espinosa et al., 2022). The primary goal of SER is to identify emotions from vocal cues independently of the spoken language. This approach has been recently adopted in economics to assess the impact of vocal sentiment expressed by figures like the Fed chair on asset prices (Gorodnichenko et al., 2023; Alexopoulos et al., 2023). The most relevant study to ours is by Gorodnichenko et al. (2023), where the authors used a Convolutional Neural Network (CNN) to classify the vocal emotions of Federal Reserve chairs during press conferences. Instead, we adopt a Fully Convolutional Neural Network (FCN), as proposed by García-Ordás et al. (2021), due to its superior out-of-sample accuracy[9] and its advantages when handling the dynamic and varied lengths of Q&A session responses.

Unlike traditional CNNs, the FCN model by García-Ordás et al. (2021) is capable of processing audio files of non-fixed length, classifying underlying vocal emotions without the need to pre-process audio into a fixed format.[10] The FCN architecture we use includes three convolutional layers, with the first two using ReLU activation and a dropout layer post the third to prevent overfitting. A pivotal feature of this architecture is the Global Average

[9]The FCN model demonstrated state-of-the-art out-of-sample accuracy as of August 2021.

[10]For a comprehensive explanation of the FCN model, see the Appendix D and the original article by García-Ordás et al. (2021).

Pooling layer, which effectively reduces data dimensionality by averaging out filter weights, thus accommodating variable audio lengths without loss of temporal dynamics - a limitation of standard CNNs. This model outputs a feature map for each emotion category, concluding with a Softmax activation layer. The model was constructed using Keras in Python, which facilitates the implementation of convolutional operations without the constraints imposed by fully connected layers.[11]

The FCN's ability to handle non-fixed length audio data is particularly advantageous. Traditional CNNs, like those used by Gorodnichenko et al. (2023), require audio data to be pre-processed into uniform lengths, often averaging the Mel-frequency cepstral coefficients (MFCC) into a single vector. This process simplifies the input but, more problematically, strips away rich temporal information critical for accurate emotion recognition. In contrast, our FCN approach retains these dynamics, thereby enhancing the detection and classification of nuanced emotional expressions in voice, which can be especially meaningful for the interpretation of monetary policy communications of the kind generated by the ECB.

3.1.2. Model Training and Validation

Following the literature on SER (García-Ordás et al., 2021), and similar to Gorodnichenko et al. (2023), 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*)[12] (Livingstone and Russo, 2018).

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

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

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). Furthermore, the audio files contain vocal cues recorded in neutral North American English such that trained models are not constrained to English with specific accents (Livingstone and Russo, 2018)[13] which is useful when applying to a proficient English-speaking European.[14]

We follow Gorodnichenko et al. (2023) and remove the emotions *Fear* and *Disgust* due to the low probability that these emotions appear during a central bank press conference.[15] 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.*[16]

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.[17] 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.[18]

---

[13] This makes the data more advantageous than vocal data in English with specific accents like in GEMEP.

[14] We do not consider the very few answers that Draghi provides in Italian.

[15] 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 leading a journalist to ask: "*[...] 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?*"

[16] 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.

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

[18] García-Ordás et al. (2021) use and compare Mel Spectrograms with MFCC and conclude that MFCC

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.[19] To ensure that our results are robust to changing stochastic sequences during the training process, we follow the deep learning literature and the literature in economics applying deep learning methods (Gu et al., 2020) and, for every training and validation process, we use a different seed that we generate randomly. To avoid overfitting, we use an "Early Stopping" - Callback.[20]

3.1.3. Using ECB Press Conference Data

Now that we have prepared a 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 the Q&A session (European Central Bank, 2022).[21] Next, we consider only the audio data[22]

---

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.

[19]Using a model ensemble ensures generalization and reduces the overall classification variance. In Appendix A, table (A.1) shows the out-of-sample accuracy of each FCN model and the average accuracy. The average accuracy is 90.7%.

[20]A common challenge in deep learning is determining the optimal training duration for a model. The 'Early Stopping - Callback' strategy helps address this by halting training before model accuracy and loss on the evaluation sample begin to deteriorate. We use an Intel Core i7 CPU at 2.3 GHz, and the entire training process for the ensemble takes approximately 16 days.

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

[22]Using the video material on the Q&A sessions, it is possible to generate emotions based on facial expressions (Curti and Kazinnik, 2023; Alexopoulos et al., 2023). In this study, we focus on vocal emotions and keep the analysis of facial expressions open for future research.

and remove the introductory statement, 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 begins to answer a question until he stops talking.[23] However, Draghi usually answers several questions in a row or uses the opportunity to summarize the debates during the governing council meetings and talks about several different topics. 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 output of this process is a novel data set consisting of synchronized voice and 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]

---

[23]This is the way Gorodnichenko et al. (2023) 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.

[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.

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

[27]Due to poor audio quality, we excluded ten press conferences from our sample: 4 October 2012, 2 May

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]

- Figure (1) -

Panel (a) of figure (1) illustrates the development of the net vocal sentiment during all press conferences 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]

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. (2023) and explicitly considers *calm* emotions.

[29] We use the first six FCN models to classify an answer, and if they disagree, the seventh FCN model is employed for the final decision.

[30] We use this quantitative measure of vocal emotions for the main analysis. In addition, we provide in Appendix E.3 a robustness check using a qualitative definition of vocal emotions. Our conclusions remain unchanged.

[31] We exclude the press conference on 20 October 2016 from our analysis due to significant model-internal disagreement during classification, classifying it as an outlier. However, our results remain robust even with the inclusion of this outlier.

[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.

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 (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).

[33] We employ the crisis dates as defined by Hartmann and Smets (2018), who identify the acute phase of the ESDC as lasting until June 2013. The authors define a period that starts around the time shortly before the Lehman Brothers' failure (September 2008) until the 'taper tantrum' as the period of financial and sovereign debt crises. The entire euro area sovereign debt crisis is dated May 2010 to June 2013.

[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.

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 introductory statement 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, this statement 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 introductory statement. While in the previous section, we describe the measurement of vocal sentiment, we also aim to capture verbal sentiment during the Q&A session.

To accurately quantify the net sentiment of the ECB president's answers and assess whether they convey a positive or negative tone, we adopt the approach by Curti and Kazinnik (2023), favoring large language modeling over traditional dictionary methods. In line with the findings of Kanelis and Siklos (2024), which highlight the effectiveness of the model in this context, we employ FinBERT, a sophisticated transformer model developed by Huang et al. (2022). FinBERT excels in discerning sentiment in economic and financial language, adeptly interpreting the tone of complex sentences within their specific context, thus surpassing the capabilities of basic word counting techniques.[37] The analysis happens at the individual sentence level, and we classify 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 in the calculation of the vocal sentiment

[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] Manela and Moreira (2017) provide evidence that automated methods are increasingly superior to lexicographic methods. Huang et al. (2022) show that FinBERT has a higher classification precision than alternative machine learning methods, for example, support vector machines.

to derive the textual sentiment for each press conference:

$$Positivity_t^{AN} = \frac{Positive_t - Negative_t}{Positive_t + Negative_t} \tag{3}$$

$Positive_t$ 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. Figure (1) displays the indicator for each press conference. In line with the findings of Alexopoulos et al. (2023) regarding the non-verbal communication of Fed chairs, our analysis also reveals a relatively low correlation between verbal and vocal cues. This underscores the additional insights gained from non-verbal communication.

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. Following the literature, we have to consider information that is revealed during the press conference.

To control for information stemming from the introductory statement, we use FinBERT on the sentence level and calculate the net sentiment using the same formula as in equation (3), just with the individual sentences instead of answers (Kanelis and Siklos, 2024).

$$Positivity_t^{IS} = \frac{Positive_{i,t} - Negative_{i,t}}{Positive_{i,t} + Negative_{i,t}} \tag{4}$$

$Positive_{i,t}$ represents the count of sentences with positive sentiment, and $Negative_{i,t}$ those with negative sentiment. This indicator quantifies the net positivity in the introductory statement. Variations in this indicator capture how ECB staff rephrase sentences to alter sentiment, introduce new sentences with positive or negative connotations, or eliminated existing sentences with contradictory sentiments.[38]

To control for possible news that Draghi may reveal in parallel 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.

Finally, 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

In our empirical analysis, we concentrate on the timeframe from July 2013 to October 2019. This period encompasses all the press conferences from the formal introduction of FG (Hartmann and Smets, 2018) up to the conclusion of Mario Draghi's tenure as the ECB president. The selection of this specific period is strategic. The formal introduction of FG, as Parle (2022) notes, marked a significant decrease in the volume of unexpected

[38]For more details on this indicator, see Kanelis and Siklos (2024).

information conveyed during press conferences, thereby diminishing the ECB's informational asymmetry. Moreover, by starting our analysis post-July 2013, we deliberately exclude the ESDC, which was characterized by heightened fiscal instability and the implementation of critical interventions like the Outright Monetary Transactions (OMT) and the European Stability Mechanism (ESM). These factors had a substantial impact on bond yields and could potentially introduce structural breaks in the data. Focusing on the period after these tumultuous events allows for a more controlled analysis, free from the distortions of an acute crisis environment.

Our study highlights that press conferences convey more than just words. As noted in the introduction, policymakers' emotions are also on display. The link between emotion and financial markets has long been known to exist (e.g., Garber (2000); Shiller (2005); Kindleberger and Aliber (2005)). Recall that the President of the ECB must convey the sense of the entire governing council, where there are likely to be differences of opinion amongst the members about the current and expected states of the euro area economy. Even if disagreements are papered over in the policy statement following a monetary policy decision, it is not surprising that financial markets seek additional, non-verbal clues, for example, about risks around the outlook. Indeed, the objectivity and credibility of the speaker are also on display at press conferences and are partly discoverable via the tone of, in this case, his voice.

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 to be the benchmark in the euro area (Altavilla et al., 2019).[39] Therefore, we estimate the following regression model, based on equation (1):

---

[39]When talking about *investors* or *capital markets*, it should be obvious that euro area government bonds are traded by market participants with all kinds of different European and non-European cultural

$$
\begin{aligned}
y_t^{DE} = \beta_0 + \beta_1 * Voice_t \times Positivity_t^{AN} + \beta_2 * Voice_t + \\
\beta_3 * Positivity_t^{AN} + \beta_4 * Positivity_t^{IS} + \sum_{i=5} \beta_i X_{ti} + \epsilon_t
\end{aligned}
\tag{5}
$$

where y_t^{DE} is the median change in the German yield of government bonds with either one-, two-, five-, or ten-year maturities within an interval of ten minutes immediately after the press conference ends. We test the following first hypothesis:

I: *The interplay of vocal emotions and language during the Q&A session influences the yield of sovereign bonds. Positive emotions raise yields as this conveys optimism about current and future expected economic conditions; conversely, negative emotions reduce yields.*

Table (1) shows the estimation results for the yields of German bonds with varying maturities:

- Table (1) around here -

We estimate a significant effect of the textual variable for the Q&A session ($Positivity_t^{AN}$) on the yields of one-year and two-year bonds (see column (1) of table (1)). Indeed, even without considering the role of vocal emotions ($Voice_t$), a more positive language used by the president to answer the questions leads to increased yields on the short-end of the yield curve. As noted above, financial markets perceive the more positive framing of the language and associate it with more optimism about underlying economic conditions. This result aligns with findings in the literature, as discussed in section 2 above. More importantly, vocal emotions amplify this relationship. Indeed, omitting a role for vocal emotions cuts the effects of positive news by roughly one half. Figure (2) provides margin plots for the interaction term such that one can measure the marginal effects of vocal emotions given a

backgrounds and not only by domestic investors of the specific country. This ensures that the effects of emotions are not driven by country-specific cultural differences in interpreting vocal emotions.

fixed level for language ($Positivity_t^{AN}$).[40]

- Figure (2) around here -

More positive vocal cues significantly amplify the impact of language, leading to higher yields. The voice serves as a crucial complement to the verbal cues and framing employed by the president. However, we observe that the effects of voice and language variables become insignificant at the longer maturities, specifically over the five and ten-year maturities. The pronounced influence of vocal and verbal cues on the shorter end of the German yield curve functions similarly to conventional monetary policy tightening.[41]

Finally, we note that FG_t and QE_t also raise German yields at all maturities. Since positive values of the monetary shock variables are scaled to be consistent with a restrictive monetary policy, the results are to be expected assuming that these instruments are intended to reduce all yields in the euro area. Forecasts exhibit an asymmetric impact on German yields. Increases in the NCY forecasts for inflation and RGDP growth across the euro area do not yield statistically significant effects. However, decreases relative to previous forecasts significantly influence some yields. Notably, declines in expected RGDP growth result in lower yields for five-year bonds, aligning with economic theory predictions regarding the direction of the effect. Conversely, the decline in the NCY inflation forecast increasing short-end yields is somewhat unexpected, as economic theory would typically also anticipate a reduction. Moreover, when statistically significant, expectations of inflation and RGDP growth exert a greater influence on yields than other determinants. This is expected, as

[40]For illustration, we choose levels of 0.25 and -0.25 for $Positivity_t^{AN}$ 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.

[41]This interpretation may be contingent on the downward deviation of the inflation rate from its target. During Draghi's tenure, the primary concern was excessively low inflation. In scenarios where inflation exceeds the target, and monetary policy tightens, positive vocal and verbal cues might have an expansionary effect, as reducing inflation is deemed as due to the successful conduct of monetary policy. However, extending this analysis beyond Draghi's term remains an avenue for future research.

forecasts, despite their imprecision, offer a numerical indication of the economic outlook.

4.2. Effects on French, Italian, and Spanish Yields

As with the German case, we estimate a significant effect of the interaction term ($Voice_t \times Positivity_t^{AN}$) on the yields of French and Italian public bonds for all maturities. Indeed, Draghi's vocal emotions have a significantly larger impact on Italian yields compared to their French or Spanish counterparts.[42] This may be because international investors are more sensitive to developments in Italy than in other parts of the euro area.[43] Table (2) provides the results.

- Table (2) around here -

Additionally, for the two-year bonds of France and Italy, we observe significant effects from the individual terms of $Voice_t$ and $Positivity_t^{AN}$, respectively. These results suggest that, under conditions of balanced vocal cues, more positive verbal cues lead to an increase in the yield of two-year French bonds, and under the conditions of balanced verbal cues, more positive vocal cues lead to a decline in the yield of two-year Italian bonds. While the results for France align with those for Germany, the Italian data offer new insights into the influence of vocal cues on asset pricing. If German yields are viewed as akin to a relatively risk-free rate, it appears that vocal cues diminish the short-term risk premium in Italy.[44]

To investigate the relevance of the complex interplay of vocal emotions and language, we

[42] We apply Wald tests to examine whether the combined effect of vocal and verbal cues on Italian bonds is similar to that on French and Spanish bonds of the same maturity. For the tests, we use average values for $Voice_t$ and $Positivity_t^{AN}$. The hypothesis is rejected for French bonds with 10-year maturity and for all Spanish maturities (results not shown).

[43] According to Eurostat, Italy's general government debt-to-GDP ratio has been persistently higher than that of France or Spain. In 2019, Italy's debt ratio was 134.2%, while France's was 97.8% and Spain's was 98.2%. (Eurostat data code is gov_10dd_edpt1 with last access on 07-06-2024.)

[44] Our analysis does not permit conclusions about whether Draghi's vocal cues uniquely impact audiences from different nations. Nonetheless, it's important to acknowledge that these bonds are traded on an international scale, transcending the national identity of the issuing state.

once again turn to marginal plots to measure the effects of vocal emotions given a specific level of positive sentiment in the language used during the Q&A session.[45] Figure (3) shows the marginal effect of vocal emotions on the bonds spreads with two-year maturity:

- Figure (3) around here -

The interplay between vocal emotions and language during the Q&A sessions significantly and asymmetrically influences French and Italian bond yields, albeit in opposite directions. For French bonds, a combination of positive vocal and verbal cues leads to higher yields, whereas in Italy, a combination of negative vocal and verbal cues results in yield increases. This phenomenon can be attributed to the information channel where unscripted communicative signals are crucial for investors as they assess potential future risks, particularly impacting the risk premium in Italy. These observations align with Mayew and Venkatachalam (2012), who found that investors place greater emphasis on negative over positive vocal emotions in risk assessment. Interestingly, no significant statistical or economic effects were observed on Spanish yields, a consistency maintained across the entire Spanish yield curve.

Figure (4) illustrates the marginal effect of vocal emotions on the spreads of government bonds with a five-year maturity:

- Figure (4) around here -

The interplay between vocal emotions and language also impacts the yields of five-year public bonds. For France, yields rise with a combination of positive vocal and verbal cues, while in Italy, they increase when cues are negatively aligned. Spanish bonds, however, remain unaffected by such cues during press conferences. Clearly, ECB communication signals influencing yields have varying impacts across the euro area's bond markets. Our estimates, however, cannot disentangle the extent to which differences across sovereign bond markets

[45] As before, results remain robust when using different values than 0.25 and -0.25 for $Positivity_t^{AN}$.

are explained by domestic versus international factors. Nevertheless, our results indicate that purely domestic factors, in addition to the euro area-wide determinants considered here, also drive yields.[46] The capital markets prioritize varying information when pricing different public bonds. We reach a similar conclusion for the yields of ten-year French, Italian, and Spanish bonds.[47]

The ECB president's vocal and verbal cues significantly influence public bond yields in the euro area. We observe that both consistently positive and negative cues during Q&A sessions cause bond yields to rise across various countries. This suggests a deeper exploration into how investors react to yield differentials. Initial evidence from Germany suggests that positive communication signals predominantly affect the risk-free yield component. This aligns with the notion that negative signals do not lead to a premium in capital investments viewed as inherently risk-free. Consequently, our next section delves into euro area yield spreads vis-a-vis German bonds.

Finally, we observe positive impacts of monetary shocks on the yields of bonds in France, Italy, and Spain. Consistent with Altavilla et al. (2019), we find that FG influences the short to medium segments of the yield curves, while QE impacts the longer end. Additionally, we find no statistically significant effects from the sentiment of the introductory statement on yields. This lack of impact is expected given the scripted nature of these statements, which allows investors to anticipate them, especially considering the pronounced relevance of FG (Parle, 2022). This result also highlights the importance of examining how unscripted information and signals impact bond yields.

---

[46]Since the ECB conducts monetary policy for the euro area, a comprehensive analysis of the role of domestic factors is beyond the scope of this paper.

[47]The margin plots for these ten-year bond estimations are provided in Appendix E.1.

4.3. Effects on Spreads

Next, we examine how vocal and verbal cues influence sovereign bond yield differentials to comprehend the effects of unscripted communication on the different components of the yields. Thus, we propose the following hypothesis:

II: *The interplay of vocal emotions and language during the Q&A session influences the spread of French, Italian, and Spanish government bonds. Positive cues reduce yield differentials while negative cues increase them.*

We subtract the yields of French, Italian, and Spanish bonds from the yield of German bonds with identical duration to obtain yield spreads:

$$\overline{y}_t^C = y_t^C - y_t^{DE},\ with\ C \in \{FR, IT, ES\} \tag{6}$$

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

$$\begin{split}\overline{y}_t^C = \beta_0 + \beta_1 * Voice_t \times Positivity_t^{AN} + \beta_2 * Voice_t + \beta_3 * Positivity_t^{AN} + \\ \beta_4 * Positivity_t^{IS} + \sum_{i=5} \beta_i X_{ti} + \epsilon_t,\ with\ C \in \{FR, IT, ES\}\end{split} \tag{7}$$

Figure (5) visualizes the impact of varying levels of vocal cues given a fixed level of verbal cues on the spread by utilizing margin plots:

- Figure (5) around here -

We estimate that consistent negative communication has a statistically and economically significant impact on the spread of Italian bonds. Consequently, bad news increases the Italian spread and such a result is usually interpreted as a rise in the risk premium. Interestingly, we find no statistically significant effect on the spreads of French and Spanish bonds, leading us to reject the hypothesis for these two countries. Based on the estimation results for yields and spreads, we infer that the increase in the French yield following consistent positive communication is due to a change in the risk-free interest rate, rather than a change in the issuer-specific risk premium.

The results reveal a significant asymmetry in the impact of unscripted communication during ECB press conferences. Positive vocal and verbal cues during the Q&A session influence the risk-free interest rate, functioning similarly to a monetary policy impulse. Conversely, negative communication affects the risk premium without impacting the risk-free interest rate. In the appendix, we present the regression results for the spreads of bonds with durations of five and ten years. These results are qualitatively similar and reinforce our interpretation of the overall yield curve.[48]

In summary, the finding of an asymmetric response to some shocks, whether verbal or economic, is not surprising. However, the differential response of yields at different maturities to ECB announcements and Draghi's verbal emotions suggests that a variable influencing yield spreads may have been omitted.

4.4. Robustness Checks

For the robustness check, we first consider the role of complexity or clarity of the language used by the ECB president. Language clarity significantly influences media attention and engagement (Ferrara and Angino, 2022). Additionally, Hayo et al. (2022) utilize high-frequency data from European stock index futures trading between 2009 and 2017, demonstrating that less clear language reduces trading activity and shifts traders' focus to the Q&A session. To ensure that communication complexity does not bias our analysis, we incorporate the Flesch-Kincaid (F-K) Grade Level indicator as an additional control variable. Furthermore, we use

[48]See Appendix E.1.

the FOG index as an alternative metric to measure complexity. More details and the estimations are provided in the appendix. Our previous conclusions remain both qualitatively and quantitatively robust when accounting for textual complexity.

Next, we replace our numerical variable for measuring vocal sentiment with a categorical variable based on a Likert scale. We distinguish between *Positive, Moderate Positive, Balanced, Moderate Negative, and Negative.* This scaling aligns more closely with how emotions are perceived in interpersonal relationships. In the appendix, we detail the derivation of this variable and provide the estimations. Once again, our main findings remain both quantitatively and qualitatively robust.

5. Conclusions

Our research provides a novel perspective of the impact of non-verbal and unscripted communication on euro area bond yields. We demonstrate that vocal emotions displayed during press conferences significantly influence the pricing of yields and yield spreads of sovereign bonds. Furthermore, we highlight the crucial interplay between vocal and verbal cues expressed by the ECB president during the Q&A sessions. Using data from the four largest euro area economies, we establish a statistically significant impact of non-verbal and verbal communication on yields, with varying effects across different maturities. Notably, we identify an asymmetric impact of vocal emotions, a previously undocumented finding. In Germany and France, unscripted communication positively affects yields, though this effect is confined to the short end of the yield curve and varies with the type of vocal emotion. Positive communication signals an increase in yields in Germany and France, while negative verbal and vocal communication leads to increases in the Italian yield curve. Further analysis of bond spreads indicates that negative communication impacts yield spreads. This likely translates into a change in the risk premium vis-a-vis German bonds.

Our results are consistent with the view that investors react to positive signals by anticipating a future monetary policy tightening. Conversely, negative signals affect the risk premium of individual countries, such as Italy, where we observe a relationship between negative communication and increasing spreads. These findings support the hypothesis that euro area bond yields can vary idiosyncratically. Therefore, while the ECB conducts a unified monetary policy, sovereign yields can react differently to the central bank's unscripted pronouncements. Nevertheless, our results also demonstrate that euro area monetary policy remains a significant factor in explaining bond spreads of euro area member states relative to German yields.

Our research contributes to the growing body of literature recognizing the influence of vocal cues on financial markets, underscoring that communication extends beyond words alone. That said, much else remains to be explored. Future research should investigate why the same emotions translate into asymmetric effects across sovereign bond markets and how financial markets perceive and process vocal cues during crises, such as the COVID-19 pandemic or rising inflation since 2021. Additionally, understanding how central bankers' emotions influence asset prices amid increasing economic and geopolitical uncertainty presents a promising avenue for further investigation.

References

Alexopoulos, M., Han, X., Kryvtsov, O., Zhang, X., 2023. *More than Words: Fed Chairs' Communication During Congressional Testimonies*. Journal of Monetary Economics .

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., 2024. *Central Bank Communication with the General Public: Promise or False Hope?* Journal of Economc Literature Forthcoming.

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., 2023. *Let's Face It: Quantifying the Impact of Nonverbal Communication in FOMC Press Conferences*. Journal of Monetary Economics 139, 110–126.

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.

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.

European Central Bank, 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.

Ferrara, F.M., Angino, S., 2022. *Does Clarity Make Central Banks More Engaging? Lessons From ECB Communications*. European Journal of Political Economy 74. 102146.

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.

Garber, P.M., 2000. *Famous First Bubbles: The Fundamentals of Early Manias*. (Cambridge, Mass.: The MIT 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.

Gorodnichenko, Y., Pham, T., Talavera, O., 2023. *The Voice of Monetary Policy*. American Economic Review Vol. 113, No. 2.

Gu, S., Kelly, B., Xiu, D., 2020. *Empirical Asset Pricing via Machine Learning*. Review of Financial Studies , 2223–2273.

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.

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.

Huang, A.H., Wang, H., Yang, Y., 2022. *FinBERT - A Large Language Model Approach to Extracting Information from Financial Text*. Contemporary Accounting Research .

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

Kanelis, D., Siklos, P.L., 2024. *The ECB Press Conference Statement: Deriving a New Sentiment Indicator for the Euro Area.* International Journal of Finance & Economics Forthcoming.

Kindleberger, C.P., Aliber, R., 2005. *Manias, Panics, and Crashes: A History of Financial Crises* (5th ed). Palgrave Macmillan.

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.

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.

Parle, C., 2022. *The Financial Market Impact of ECB Monetary Policy Press Conferences - A Text Based Approach*. European Journal of Political Economy. 74. 102230.

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.

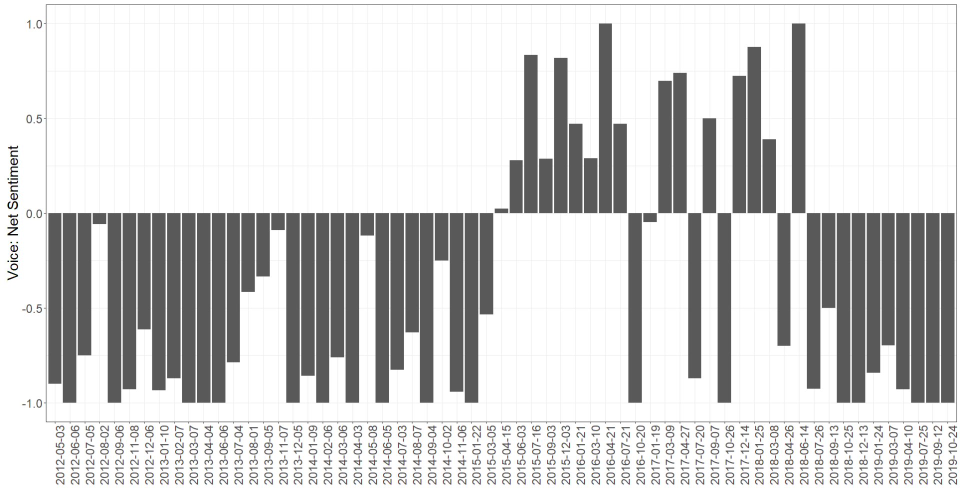
Schmelling, M., Wagner, C., 2019. *Does Central Bank Tone Move Asset Prices?* Working Paper .

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 .

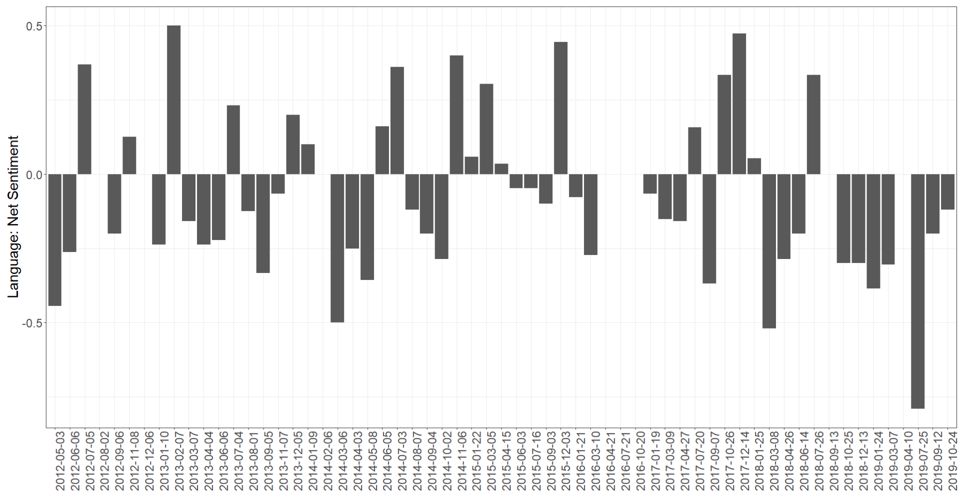
Shiller, R., 2005. *Irrational Exuberance*, Second Edition. (Princeton, N.J.: Princeton University Press).

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.

Figure 1: Net Sentiment of Vocal and Verbal Cues during Q&A Sessions



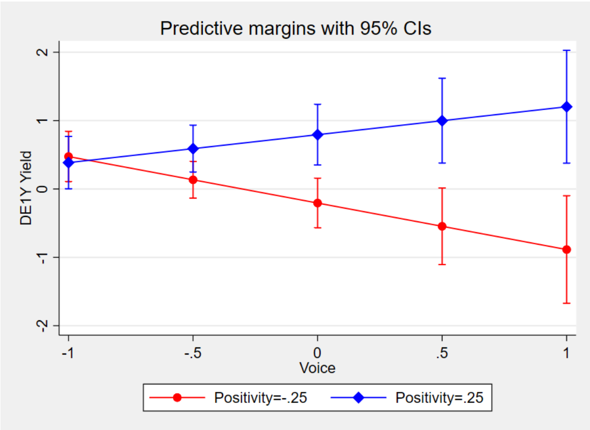
(a) Sentiment: Vocal Cues



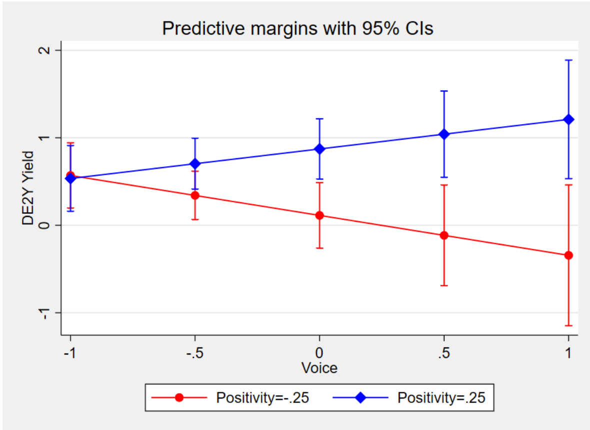
(b) Sentiment: Verbal Cues

Note: Panel (a) illustrates the net sentiment of the vocal emotions measurable in Mario Draghi's answers during the Q&A sessions of the ECB press conferences. We measure vocal sentiment using our SER model (see section 3.1) and calculate the net value using equation (2). Panel (b) illustrates the net sentiment of the verbal signals measurable in Mario Draghi's answers during the Q&A sessions of the ECB press conferences. We measure verbal sentiment using FinBERT (see section 3.2) and calculate the net value using equation (3). In the appendix, table (A.2) and table (A.3) provide the numbers and the vocal and verbal composition of positive v. negative answers.

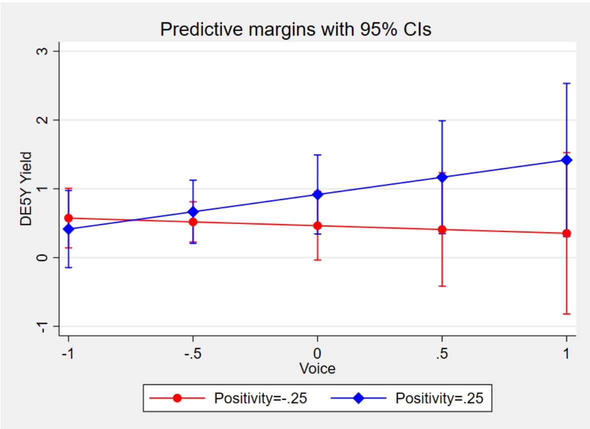
Figure 2: Marginal Effect of Vocal Emotions Given Language on German Yields



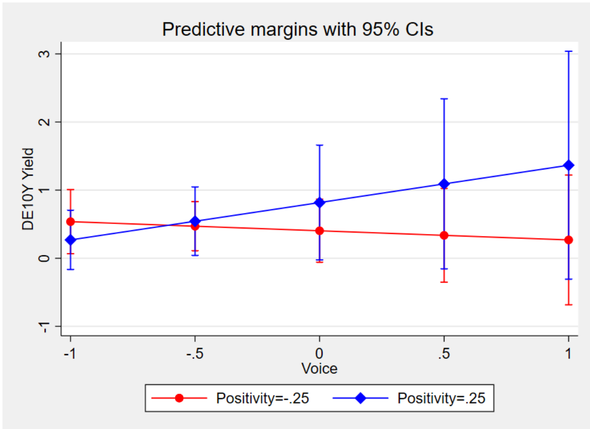
(a) One-Year German Yield



(b) Two-Year German Yield



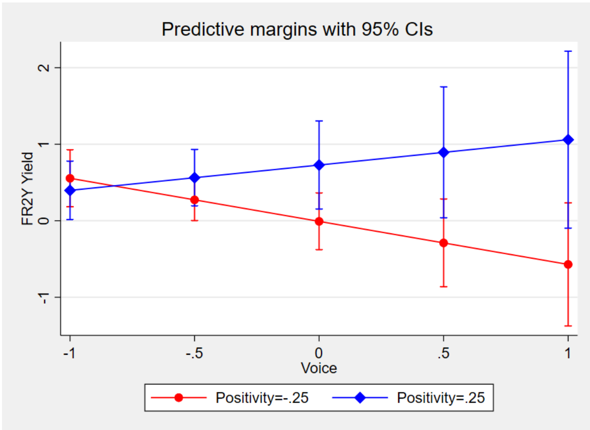
(c) Five-Year German Yield



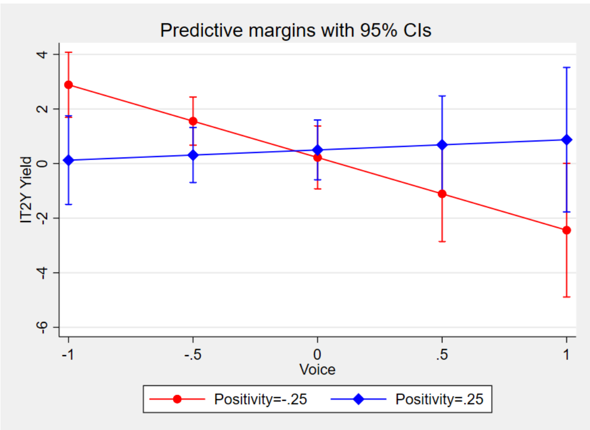
(d) Ten-Year German Yield

Note: These plots visualize the marginal effect of a change in $Voice_t$ given a specific level of $Positivity_t^{AN}$ on the yields of German government bonds. We report the change in yields in basis points. We illustrate the marginal effect of $Voice_t$ for $Positivity_t^{AN} \in \{-0.25,\ 0.25\}$. We report the estimations in table (2).

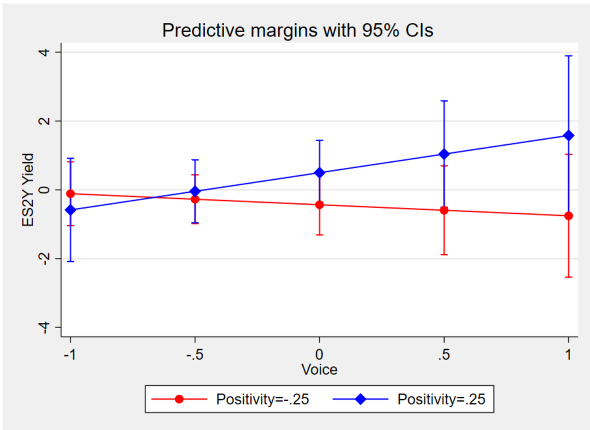
Figure 3: Marginal Effect of Vocal Emotions Given Language on Two-Year Yields



(a) Two-Year French Yield



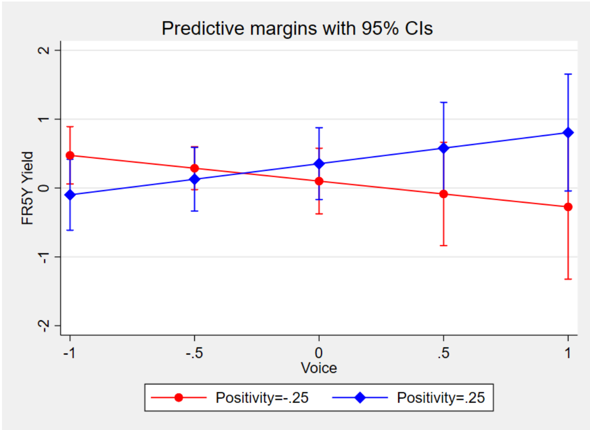
(b) Two-Year Italian Yield



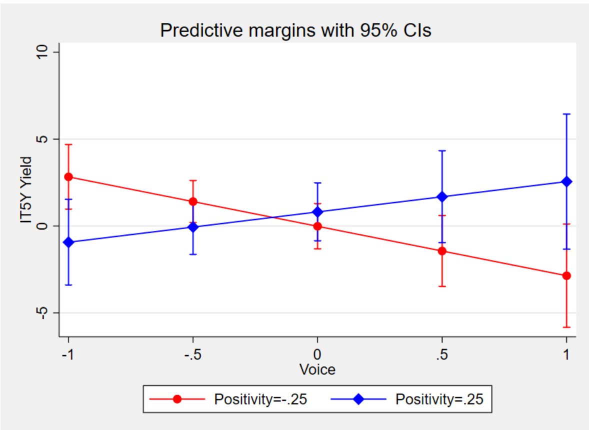
(c) Two-Year Spanish Yield

Note: These plots visualize the marginal effect of a change in $Voice_t$ given a specific level of $Positivity_t^{AN}$ on the yield of two-year government bonds from a) France, b) Italy, and c) Spain. We report the change in yields in basis points. We illustrate the marginal effect of $Voice_t$ for $Positivity_t^{AN} \in \{-0.25,\ 0.25\}$. We report the estimations in table (2).

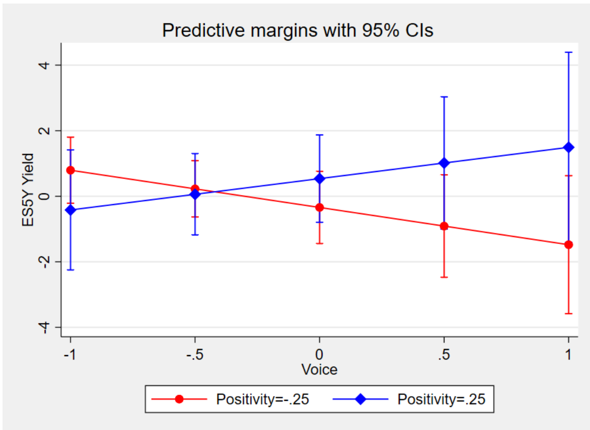
Figure 4: Marginal Effect of Vocal Emotions Given Language on Five-Year Yields



(a) Five-Year French Yield



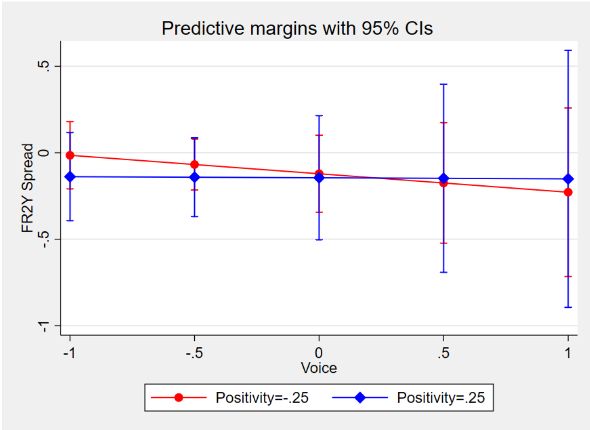
(b) Five-Year Italian Yield



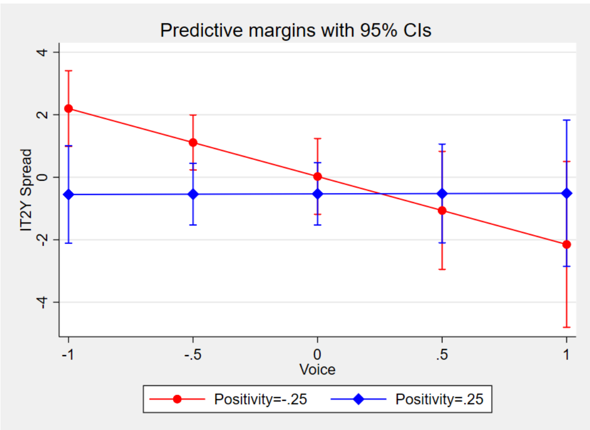
(c) Five-Year Spanish Yield

Note: These plots visualize the marginal effect of a change in $Voice_t$ given a specific level of $Positivity_t^{AN}$ on the yields of five-year government bonds from a) France, b) Italy, and c) Spain. We report the change in yields in basis points. We illustrate the marginal effect of $Voice_t$ for $Positivity_t^{AN} \in \{-0.25,\ 0.25\}$. We report the estimations in table (2).

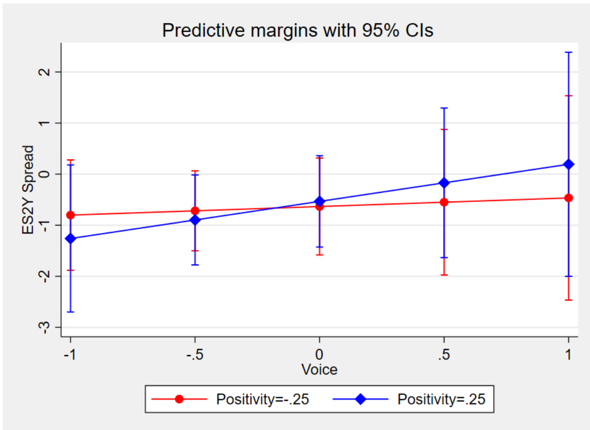
Figure 5: Marginal Effect of Vocal Emotions Given Language on Two-Year Spreads



(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 $Positivity_t^{AN}$ on the spread of two-year government bonds from a) France, b) Italy, and c) Spain. We report the change in yields in basis points. We illustrate the marginal effect of $Voice_t$ for $Positivity_t^{AN} \in \{-0.25,\ 0.25\}$.

Table 1: **German Bond Yields**

	(1) DE1Y	(2) DE2Y	(3) DE5Y	(4) DE10Y
$Voice_t \times Positivity_t^{AN}$	2.18***	1.59***	1.22	1.36
	(0.50)	(0.53)	(0.99)	(1.22)
$Voice_t$	-0.14	-0.06	0.20	0.21
	(0.20)	(0.19)	(0.24)	(0.21)
$Positivity_t^{AN}$	2.00***	1.52***	0.91	0.83
	(0.49)	(0.42)	(0.74)	(1.09)
$Positivity_t^{IS}$	0.15	0.02	-0.20	-1.04**
	(0.37)	(0.37)	(0.48)	(0.45)
$Timing$	0.83***	1.01***	0.89***	0.00
	(0.16)	(0.16)	(0.17)	(0.20)
FG	0.78***	0.96***	1.16***	0.48***
	(0.08)	(0.10)	(0.13)	(0.12)
QE	0.18***	0.34***	0.68***	1.32***
	(0.06)	(0.08)	(0.15)	(0.15)
$\Delta Inflation_t^{NCY,Positive}$	1.42	2.13	-0.13	-2.59
	(2.42)	(3.55)	(2.33)	(2.81)
$\Delta Inflation_t^{NCY,Negative}$	4.49**	3.21*	1.47	1.13
	(1.62)	(1.62)	(1.82)	(1.63)
$\Delta RGDP_t^{NCY,Positive}$	-1.44	-1.71	-2.67	0.73
	(1.88)	(2.59)	(1.94)	(2.48)
$\Delta RGDP_t^{NCY,Negative}$	0.40	-2.28	-7.31***	0.06
	(1.52)	(1.48)	(2.14)	(2.17)
$Constant$	-0.37	-0.17	0.29	0.43
	(0.24)	(0.22)	(0.30)	(0.30)
R^2	0.868	0.914	0.929	0.949
Obs	48	48	48	48

Note: We regress our variables on the change in the yield of one, 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), $Positivity_t^{AN}$ 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). $Positivity_t^{IS}$ measures the change in the framing of the introductory statement 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 2: **French, Italian, and Spanish Bond Yields**

	(1) FR2Y	(2) IT2Y	(3) ES2Y	(4) FR5Y	(5) IT5Y	(6) ES5Y	(7) FR10Y	(8) IT10Y	(9) ES10Y
$Voice_t \times Positivity_t^{AN}$	1.79**	6.08**	2.81	1.65**	9.18**	4.18	3.54**	11.27***	4.29
	(0.75)	(2.85)	(2.40)	(0.69)	(4.09)	(2.80)	(1.36)	(3.90)	(2.89)
$Voice_t$	-0.12	-1.14**	0.38	0.04	-0.55	-0.09	0.13	0.46	0.21
	(0.21)	(0.47)	(0.38)	(0.22)	(0.67)	(0.44)	(0.24)	(0.73)	(0.57)
$Positivity_t^{AN}$	1.47**	0.55	1.86	0.51	1.65	1.76	1.45	2.67	1.76
	(0.67)	(1.69)	(1.38)	(0.66)	(2.34)	(1.87)	(1.23)	(2.76)	(2.30)
$Positivity_t^{IS}$	0.41	2.26	-0.86	1.13*	1.54	0.48	0.30	0.46	-0.05
	(0.41)	(1.44)	(1.32)	(0.63)	(1.84)	(1.57)	(0.59)	(1.85)	(1.81)
$Timing_t$	0.94***	0.55	0.61**	0.91***	0.36	0.69*	0.10	-0.12	0.04
	(0.16)	(0.41)	(0.30)	(0.20)	(0.46)	(0.37)	(0.22)	(0.43)	(0.38)
FG_t	0.84***	0.57**	0.37**	1.13***	0.45	0.57**	0.63***	0.26	0.31
	(0.10)	(0.25)	(0.17)	(0.16)	(0.29)	(0.23)	(0.18)	(0.35)	(0.33)
QE_t	0.35***	0.60***	0.39***	0.86***	0.78***	0.67***	1.43***	1.41***	1.29***
	(0.08)	(0.16)	(0.14)	(0.09)	(0.22)	(0.19)	(0.17)	(0.25)	(0.21)
$\Delta Inflation_t^{NCY,Positive}$	2.43	4.50	2.97	-0.57	4.05	3.00	-0.81	4.07	1.08
	(3.45)	(5.01)	(3.36)	(2.94)	(5.60)	(3.92)	(3.50)	(5.94)	(4.98)
$\Delta Inflation_t^{NCY,Negative}$	2.54	9.42	-4.80	1.82	-2.91	-2.30	-0.88	-1.44	-4.11
	(1.59)	(7.49)	(5.61)	(1.41)	(8.49)	(5.85)	(1.70)	(8.05)	(6.12)
$\Delta RGDP_t^{NCY,Positive}$	-2.36	-7.38	-4.02	-1.33	-7.71	-4.65	0.48	-6.48	-3.39
	(2.54)	(4.93)	(3.35)	(2.19)	(5.83)	(4.12)	(2.95)	(5.81)	(4.46)
$\Delta RGDP_t^{NCY,Negative}$	0.80	0.90	-2.05	-4.76**	-0.98	3.13	2.91	1.68	1.45
	(1.37)	(6.17)	(5.24)	(2.03)	(7.65)	(5.99)	(2.06)	(7.68)	(5.90)
$Constant$	-0.43	-1.18	0.19	-0.78**	-0.37	-0.60	-0.38	0.34	0.09
	(0.26)	(0.81)	(0.74)	(0.38)	(1.04)	(0.87)	(0.37)	(1.06)	(1.01)
R^2	0.897	0.684	0.552	0.942	0.615	0.615	0.939	0.758	0.730
Obs	48	47	47	48	47	47	48	47	47

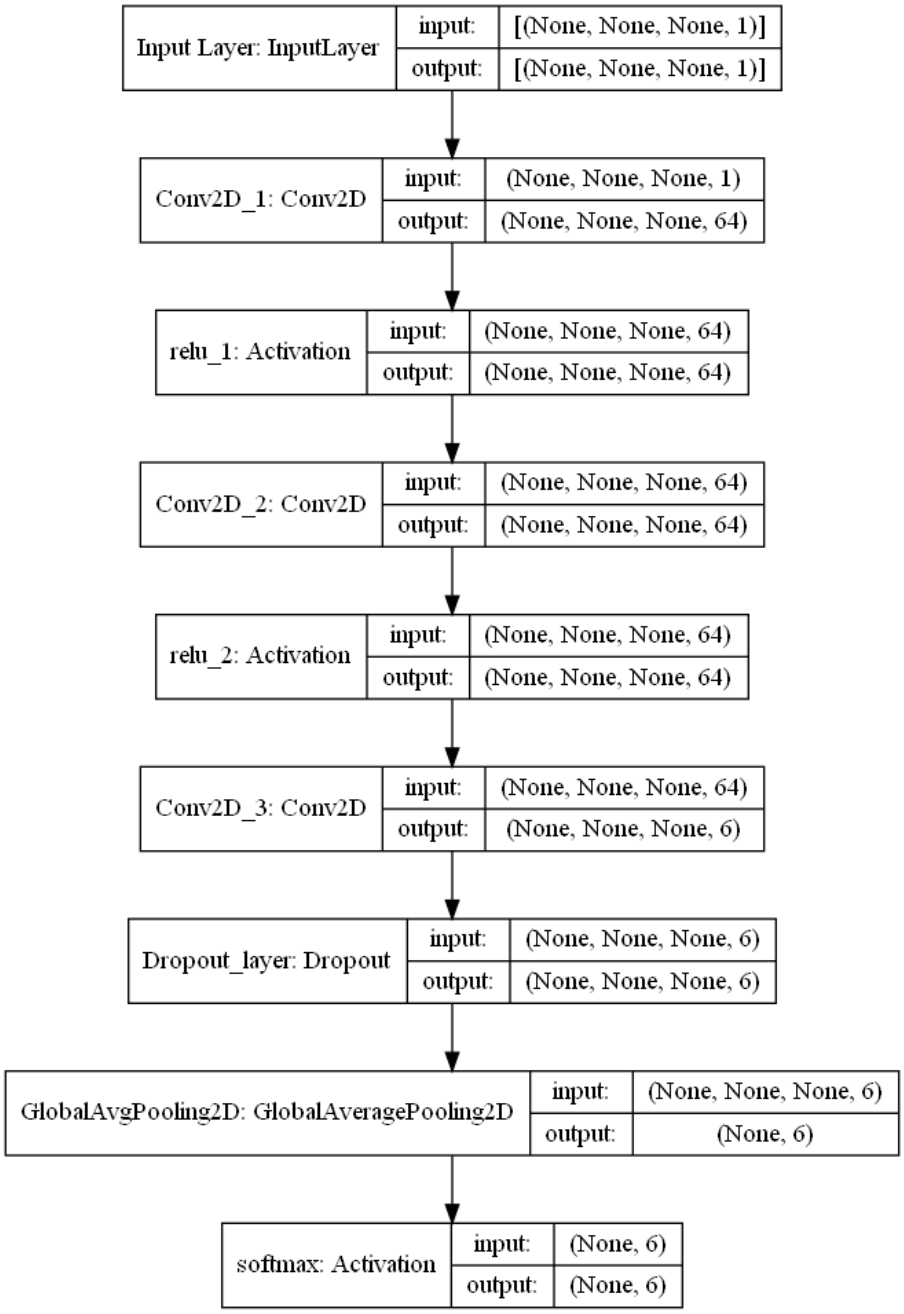
Note: We regress our variables on the change in the yield of two, five, and ten-year government bonds of France, Italy, and Spain for July 2013 until October 2019. We removed the observation on July 2013 for Italy and Spain to avoid an outlier driving the results. The dependent variable is calculated as the difference in the median price in a narrow time window before the ECB press conference and a narrow time window afterward. $Voice_t$ is Draghi's net vocal sentiment during the Q&A session (see 3.1) and $Positivity_t^{AN}$ measures the net positivity of Draghi's answers. $Positivity_t^{IS}$ measures the framing of the introductory statement (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. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Online Appendix to "Emotion in Euro Area Monetary Policy Communication and Asset Prices: The Draghi Era"

Dimitrios Kanelis   Pierre L. Siklos

Appendix A. Additional Figures and Tables

Figure A1: FCN Model Representation



Note: This figure provides a representation of the FCN model we are using for the SER.

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

1. FCN	**2. FCN**	**3. FCN**	**4. FCN**	**5. FCN**	**6. FCN**	**7. FCN**	**Average**
92.7%	90.9%	91.6%	90.5%	88.7%	90.5%	89.9%	90.7%

Note: This table displays the out-of-sample accuracy of each FCN model that achieved the highest precision in its respective training session. Identical hyperparameters were used across sessions, but the composition of training and validation sets, as well as the randomization seed, varied to enhance generalization.

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

PC Date	Voice: Positive	Voice: Negative	Voice: Neutral	Voice: Net Sentiment	Text: Net Sentiment
03-05-2012	1	19	3	-0.90	-0.44
06-06-2012	0	30	0	-1.00	-0.26
05-07-2012	4	28	0	-0.75	0.37
02-08-2012	16	18	0	-0.06	0
06-09-2012	0	30	0	-1.00	-0.2
08-11-2012	1	27	0	-0.93	0.13
06-12-2012	6	25	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	26	1	-1.00	-0.22
04-07-2013	3	25	1	-0.79	0.23
01-08-2013	7	17	0	-0.42	-0.13
05-09-2013	10	20	0	-0.33	-0.33
07-11-2013	10	12	1	-0.09	-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	3	22	0	-0.76	-0.5
03-04-2014	0	23	0	-1.00	-0.25
08-05-2014	15	19	0	-0.12	-0.36
05-06-2014	0	47	0	-1.00	0.16
03-07-2014	6	63	0	-0.83	0.36
07-08-2014	13	57	0	-0.63	-0.12
04-09-2014	0	42	0	-1.00	-0.20
02-10-2014	9	15	0	-0.25	-0.29
06-11-2014	1	33	0	-0.94	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 A.3: **Vocal and Textual Sentiment (Part II)**

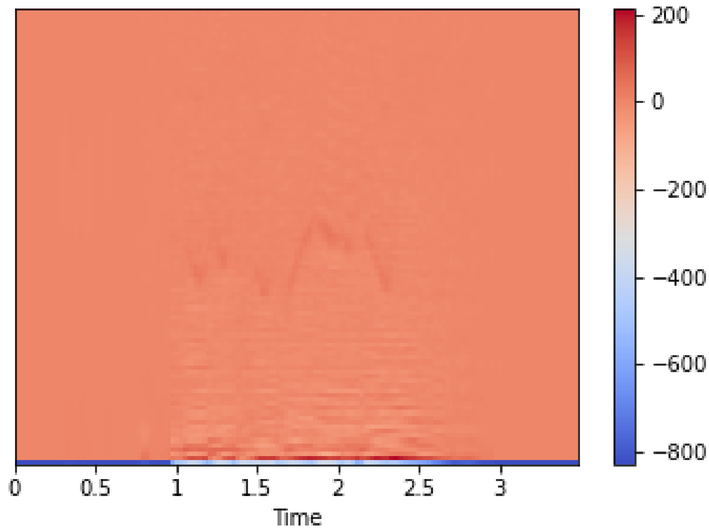
PC Date	**Voice: Positive**	**Voice: Negative**	**Voice: Neutral**	**Voice: Net Sentiment**	**Text: Net Sentiment**
22-01-2015	0	39	6	-1.00	0.06
05-03-2015	7	23	0	-0.53	0.30
15-04-2015	23	22	0	0.02	0.03
03-06-2015	23	13	3	0.28	-0.05
16-07-2015	33	3	3	0.83	-0.05
03-09-2015	18	10	0	0.29	-0.10
03-12-2015	30	3	0	0.82	0.44
21-01-2016	25	9	0	0.47	-0.08
10-03-2016	20	11	4	0.29	-0.27
21-04-2016	23	0	3	1.00	0.00
21-07-2016	25	9	1	0.47	0.00
19-01-2017	10	11	0	-0.05	-0.07
09-03-2017	39	7	1	0.70	-0.15
27-04-2017	20	3	11	0.74	-0.16
20-07-2017	2	29	0	-0.87	0.16
07-09-2017	18	6	13	0.50	-0.37
26-10-2017	0	13	17	-1.00	0.33
14-12-2017	25	4	0	0.72	0.47
25-01-2018	30	2	2	0.88	0.05
08-03-2018	25	11	1	0.39	-0.52
26-04-2018	3	17	1	-0.70	-0.29
14-06-2018	39	0	0	1.00	-0.2
26-07-2018	1	26	0	-0.93	0.33
13-09-2018	7	21	5	-0.50	0
25-10-2018	0	12	17	-1.00	-0.3
13-12-2018	0	27	1	-1.00	-0.3
24-01-2019	3	35	1	-0.84	-0.38
07-03-2019	5	28	2	-0.70	-0.30
10-04-2019	1	27	0	-0.93	0.00
25-07-2019	0	31	1	-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).

Appendix B. Audio Processing

We convert all audio files to a 22050 Hz sample rate and mono channel. To extract MFCCs, we use the default options of the Librosa package: the Fast Fourier Transform window length is 2048, and the number of samples between successive frames is 512. We apply an orthonormal discrete cosine transformation and extract the first 100 MFCCs for each audio file. The audio files are neither cut nor preprocessed. Figure (B1) visualizes the MFCCs for a happy vocal emotion with normal intensity from RAVDESS:

Figure B1: **Visual Illustration of MFCCs**

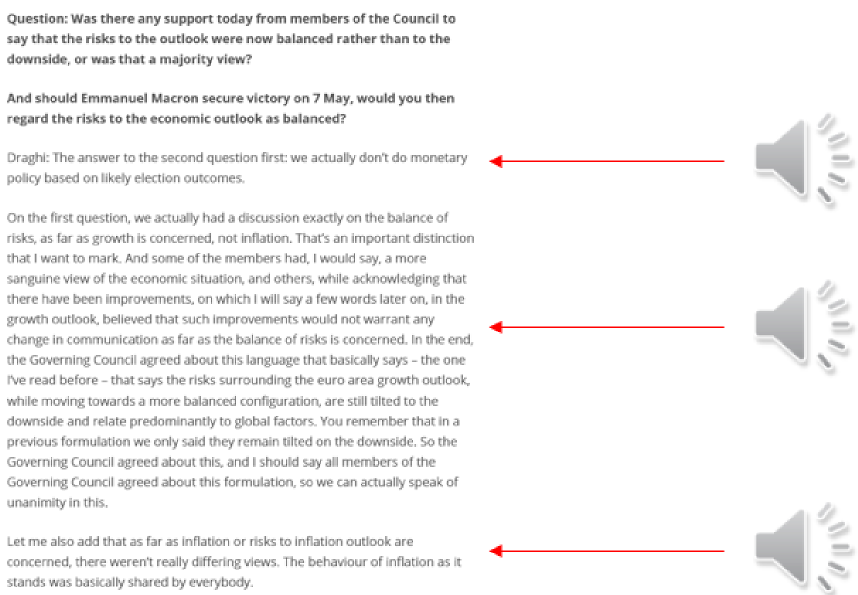


Appendix C. Construction of a synchronized Audio-Language Data Set

For our analysis, we introduce a novel dataset that synchronizes audio and textual data of Mario Draghi's voice and language. We download all audio data from the ECB Webcasts and convert them into WAV files. We exclude the introductory statements, moderator's interventions, journalists' questions, and the vice president's answers.

Typically, journalists at the ECB press conference ask two or three questions at once, including intermediate questions. Draghi usually responds to all questions in one sequence, even though the questions and answers often cover fundamentally different topics. Additionally, Draghi uses specific questions to share information about discussions during the ECB governing council meetings. To identify all individual answers within Draghi's contributions, we utilize a distinctive feature of the ECB press conference transcripts: the ECB staff separates the president's answers into distinct paragraphs. We adhere to this structure and edit the audio files to ensure synchronization with the text, maintaining the integrity of the message. Figure (C1) illustrates our approach:

Figure C1: **Illustration of Data Set Construction and Synchronization**



We apply this identification strategy to all press conferences held between May 2012 and October 2019. By manually screening all press conferences, we identify interjections by journalists that the ECB staff did not account for. In such cases, we remove the interjections and split the response into two separate answers: one before the interjection and one after. Our dataset comprises 71 press conferences, resulting in 2,336 individual answers as both audio and textual data.

Appendix D. FCN Model Structure Information

- **Convolutional Layers:** Convolutional layers perform a linear operation that involves multiplying a set of weights with the input data. This input can be raw audio data represented as an image or a feature map output from a previous convolutional layer. In our model framework, we utilize three convolutional layers. Each layer applies filters to the input data to detect various features. The first layer contains 64 filters with a kernel size of (11, 7), meaning it uses 64 sets of weights with a window size of 11x7 to process the input. The second layer also has 64 filters but with a kernel size of (7, 11). The third layer's number of filters corresponds to the number of emotion classes in our classification task, which is 6. These layers help the network learn to recognize patterns and features associated with different vocal emotions.

- **Activation Function:** Activation functions are used to propagate the output of one layer's nodes forward to the next layer. In our model, we use rectified linear units (ReLU), which are the most commonly used activation function in convolutional neural networks. The ReLU function outputs zero for any input value below zero, and for any input above zero, it outputs the input value itself, thus establishing a linear relationship:

$$f(x) = \max(0, x) \tag{D.1}$$

  This non-linear transformation allows the network to learn complex patterns. For the final classification layer, we use a Softmax activation function, which is recommended for neural networks solving classification problems. Softmax converts the output into a probability distribution over the possible classes, ensuring the sum of the probabilities is one.

- **Dropout Layer:** The Dropout layer helps prevent overfitting by randomly setting a fraction of input units to zero during each training step. This process forces the network to learn more robust features by not relying too heavily on any single neuron. In our model, we use a dropout rate of 20%, meaning 20% of the input units are set to zero at each training step.

- **GlobalAveragePooling Layer:** The GlobalAveragePooling layer calculates the average output of each feature map from the previous layer. This operation reduces the data dimensions, preparing the model for the final classification step using a Softmax activation function. The GlobalAveragePooling layer extracts a single value from each filter, corresponding to the average of all filter weights. This approach allows for the analysis of non-fixed-length audio files, ensuring the model can handle varying input sizes effectively (García-Ordás et al., 2021).

- **Model Compilation:** We use "Adam" as the optimizer for model compilation, which is a stochastic gradient descent method known for its efficiency and adaptive learning rate. For model evaluation, we focus on accuracy, defined as:

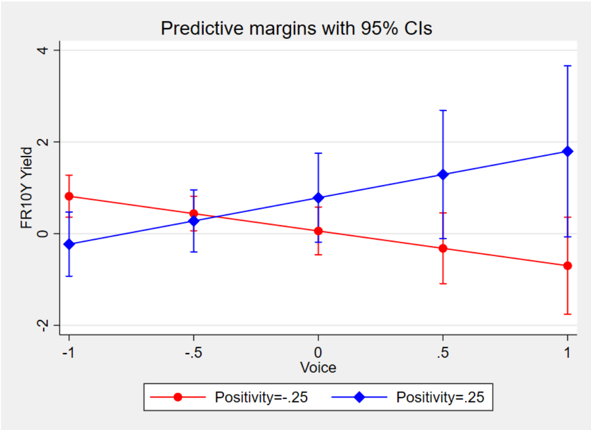
$$Accuracy = \frac{1}{N} \sum^{N} i = 1\mathbb{1}(y_{i,\text{Pred}} = y_{i,\text{True}}) \tag{D.2}$$

N is the number of observations in our validation set and $\mathbb{1}(y_{i,Pred} = y_{i,True})$ is an indicator variable that equals 1 if our classification for observation i is correct, and 0 otherwise. We evaluate accuracy metrics exclusively on data that are not part of the training set. Our model uses a batch size of 80 and runs the training process for 2500 epochs. To avoid overfitting and reduce training time, we include an EarlyStopping callback with a patience of 100 epochs. If the model does not improve in out-of-sample accuracy after 100 epochs, the training stops, and the model with the highest accuracy is saved.

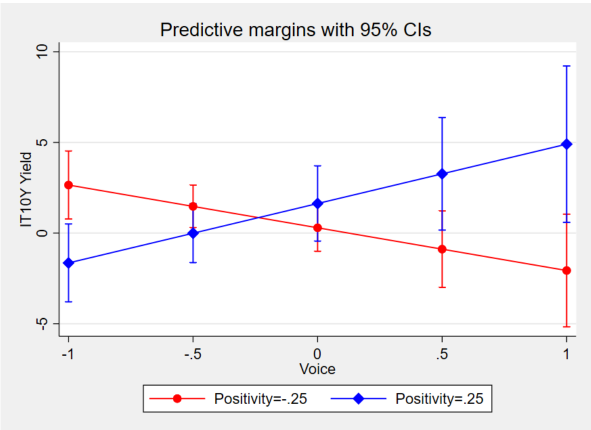
Appendix E. Additional Results

Appendix E.1. Additional Results for Yields and Spreads

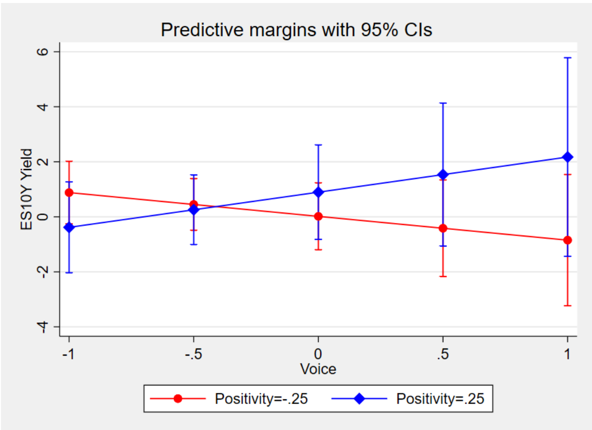
Figure E1: Marginal Effect of Vocal Emotions Given Language on Ten-Year Yields



(a) Ten-Year French Yield



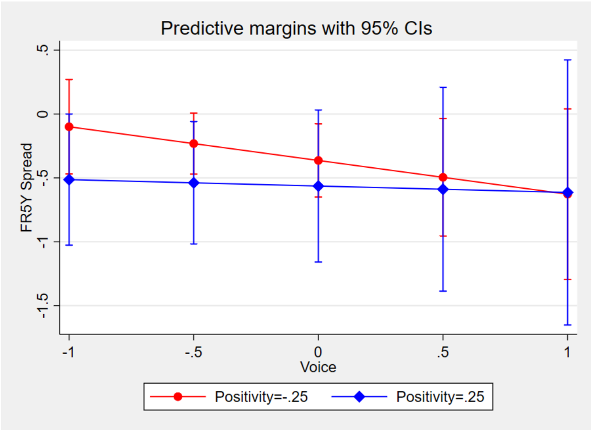
(b) Ten-Year Italian Yield



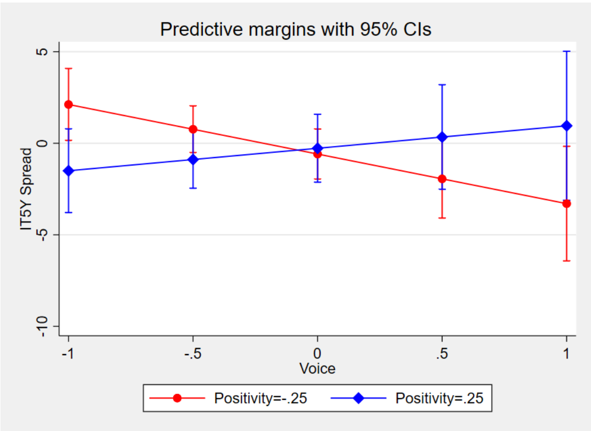
(c) Ten-Year Spanish Yield

Note: These plots visualize the marginal effect of a change in $Voice_t$ given a specific level of $Positivity_t^{AN}$ on the yields of ten-year government bonds from a) France, b) Italy, and c) Spain. We report the change in yields in basis points. We illustrate the marginal effect of $Voice_t$ for $Positivity_t^{AN} \in \{-0.25,\ 0.25\}$. We report the estimations in table (2).

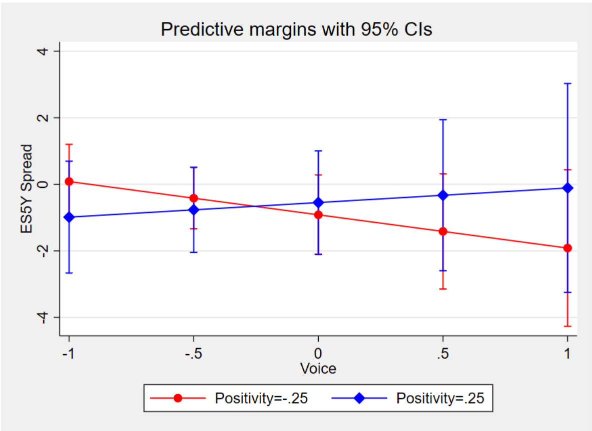
Figure E2: Marginal Effect of Vocal Emotions Given Language on Five-Year Spreads



(a) Ten-Year French Yield



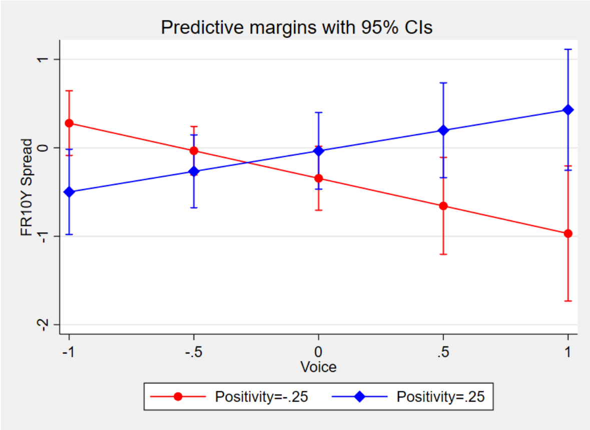
(b) Ten-Year Italian Yield



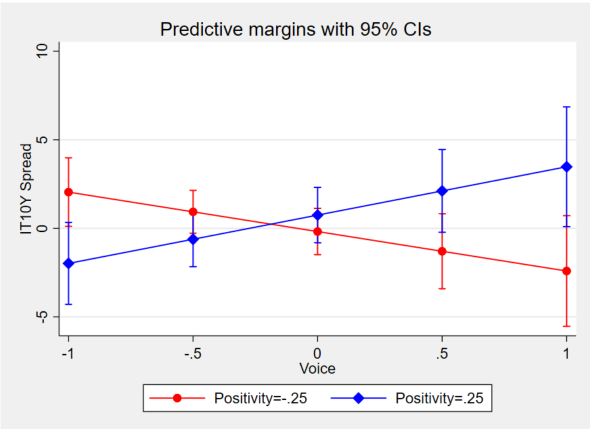
(c) Ten-Year Spanish Yield

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

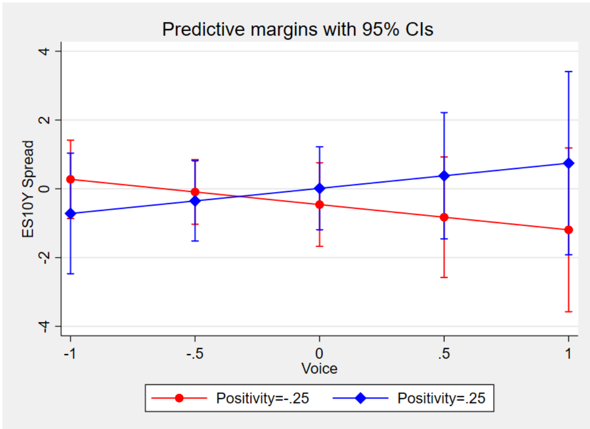
Figure E3: Marginal Effect of Vocal Emotions Given Language on Ten-Year Spreads



(a) Ten-Year French Spreads



(b) Ten-Year Italian Spreads



(c) Ten-Year Spanish Spreads

Note: These plots visualize the marginal effect of a change in $Voice_t$ given a specific level of $Positivity_t^{AN}$ on the yields of ten-year government bonds from a) France, b) Italy, and c) Spain. We report the change in yields in basis points. We illustrate the marginal effect of $Voice_t$ for $Positivity_t^{AN} \in \{-0.25,\ 0.25\}$.

Appendix E.2. Robustness Check: Complexity of the Language

First, we use the Flesch-Kincaid (F-K) Grade Level to estimate text difficulty. This score approximates the number of years of education a person needs to understand the content:

$$F-K\ Grade\ Level = 0.39\times\left(\frac{Total\ Words}{Total\ Sentences}\right)+11.8\times\left(\frac{Total\ Syllables}{Total\ Words}\right)-15.59 \quad (E.1)$$

We calculate the F-K Grade Level for each answer i individually and then average it over all N answers of the Q&A session in t. To prevent short answers from biasing this indicator, we exclude all answers that consist of only one sentence:

$$TextComplexity_t = \frac{1}{N} * \sum_{i=1}^{N} F-K\ Grade\ Level_i \quad (E.2)$$

Second, we replace the F-K Grade Level with the Gunning FOG Index, which explicitly accounts for the use of complex words when calculating text complexity:

$$FOG\ Index = 0.4\times\left(\left[\frac{Total\ Words}{Total\ Sentences}\right]+100\times\left[\frac{Complex\ Words}{Total\ Words}\right]\right) \quad (E.3)$$

We calculate the FOG Index for each answer individually and then average it over all N answers for each Q&A session in t:

$$TextComplexity_t = \frac{1}{N} * \sum_{i=1}^{N} FOG\ Index_i \quad (E.4)$$

Our conclusions remain consistent regardless of which text complexity indicator we use.

Table E.1: **Robustness Check: F-K Grade**

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)
	DE1Y	DE2Y	DE5Y	DE10Y	FR2Y	FR5Y	FR10Y	IT2Y	IT5Y	IT10Y	ES2Y	ES5Y	ES10Y
$Voice_t \times Positivity_t^{AN}$	2.35***	1.76***	1.31	1.31	2.03**	1.85**	3.48**	6.14*	8.58**	10.59***	2.34	3.98	3.80
	(0.53)	(0.58)	(1.02)	(1.23)	(0.81)	(0.72)	(1.34)	(3.06)	(3.92)	(3.50)	(2.29)	(2.84)	(2.64)
$voice_t$	-0.16	-0.08	0.19	0.21	-0.14	0.02	0.13	-1.15**	-0.47	0.54	0.44	-0.06	0.27
	(0.19)	(0.17)	(0.24)	(0.21)	(0.20)	(0.22)	(0.24)	(0.48)	(0.62)	(0.66)	(0.33)	(0.44)	(0.52)
$Positivity_t^{AN}$	2.18***	1.70***	0.99	0.77	1.73**	0.71	1.39	0.61	1.08	2.03	1.42	1.57	1.29
	(0.51)	(0.45)	(0.75)	(1.09)	(0.74)	(0.64)	(1.21)	(1.81)	(2.30)	(2.48)	(1.37)	(1.90)	(2.09)
$Positivity_t^{IS}$	0.18	0.04	-0.19	-1.05**	0.44	1.16*	0.30	2.28	1.28	0.16	-1.06	0.39	-0.27
	(0.36)	(0.34)	(0.48)	(0.45)	(0.37)	(0.62)	(0.60)	(1.48)	(1.71)	(1.67)	(1.19)	(1.56)	(1.68)
$TextComplexity_t^{FK-Grade}$	0.21*	0.22**	0.10	-0.06	0.30***	0.24*	-0.07	0.09	-0.90	-1.01**	-0.70**	-0.30	-0.73*
	(0.12)	(0.08)	(0.08)	(0.13)	(0.07)	(0.13)	(0.14)	(0.51)	(0.56)	(0.48)	(0.33)	(0.45)	(0.39)
$Timing_t$	0.76***	0.94***	0.86***	0.02	0.84***	0.83***	0.12	0.52	0.62	0.17	0.81**	0.78*	0.25
	(0.16)	(0.15)	(0.19)	(0.20)	(0.15)	(0.19)	(0.23)	(0.42)	(0.52)	(0.45)	(0.33)	(0.41)	(0.41)
FG_t	0.75***	0.94***	1.15***	0.49***	0.80***	1.11***	0.64***	0.57**	0.51	0.33	0.42**	0.59**	0.36
	(0.08)	(0.09)	(0.14)	(0.13)	(0.09)	(0.16)	(0.18)	(0.25)	(0.31)	(0.34)	(0.18)	(0.25)	(0.32)
QE_t	0.19***	0.35***	0.69***	1.31***	0.36***	0.87***	1.43***	0.60***	0.73***	1.35***	0.35***	0.65***	1.25***
	(0.06)	(0.08)	(0.15)	(0.15)	(0.08)	(0.09)	(0.17)	(0.16)	(0.20)	(0.25)	(0.13)	(0.18)	(0.20)
$\Delta Inflation_t^{NCY,Positive}$	1.49	2.20	-0.10	-2.61	2.52	-0.49	-0.83	4.49	4.19	4.22	3.07	3.05	1.19
	(2.44)	(3.56)	(2.37)	(2.82)	(3.45)	(2.93)	(3.52)	(5.01)	(5.94)	(6.07)	(3.61)	(4.07)	(5.10)
$\Delta Inflation_t^{NCY,Negative}$	3.55**	2.23*	1.00	1.43	1.19	0.74	-0.55	9.01	1.15	3.11	-1.68	-0.94	-0.84
	(1.58)	(1.30)	(1.89)	(1.97)	(1.05)	(1.55)	(1.90)	(8.59)	(8.99)	(7.84)	(5.31)	(6.97)	(5.51)
$\Delta RGDP_t^{NCY,Positive}$	-1.66	-1.94	-2.78	0.80	-2.67	-1.59	0.55	-7.44	-7.09	-5.78	-3.55	-4.44	-2.89
	(1.87)	(2.57)	(1.98)	(2.50)	(2.53)	(2.17)	(2.95)	(5.09)	(6.06)	(5.68)	(3.56)	(4.30)	(4.40)
$\Delta RGDP_t^{NCY,Negative}$	0.42	-2.27	-7.31***	0.06	0.82	-4.75**	2.90	0.88	-0.79	1.89	-1.91	3.19	1.60
	(1.48)	(1.44)	(2.17)	(2.17)	(1.41)	(2.01)	(2.06)	(6.25)	(7.47)	(7.32)	(5.16)	(6.04)	(5.71)
Constant	-2.42*	-2.29**	-0.73	1.07	-3.37***	-3.13**	0.33	-2.07	8.57	10.37**	7.07**	2.39	7.30*
	(1.23)	(0.85)	(0.93)	(1.37)	(0.85)	(1.42)	(1.49)	(5.03)	(5.65)	(4.89)	(3.40)	(4.44)	(3.98)
R^2	0.876	0.919	0.930	0.949	0.910	0.945	0.939	0.685	0.647	0.784	0.610	0.621	0.751
Obs	48	48	48	48	48	48	48	47	47	47	47	47	47

Note: We regress our variables on the change in the yields of government bonds of Germany, 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. $Voice_t$ is Draghi's net vocal sentiment during the Q&A session (see 3.1) and $Positivity_t^{AN}$ measures the net positivity of Draghi's answers. $Positivity_t^{IS}$ measures the framing of the introductory statement (see 3.3). We include a complexity indicator to measure to clarity of the language used during the Q&A session based on the F-K Grade Score. 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 E.2: **Robustness Check: FOG Index**

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)
	DE1Y	DE2Y	DE5Y	DE10Y	FR2Y	FR5Y	FR10Y	IT2Y	IT5Y	IT10Y	ES2Y	ES5Y	ES10Y
$Voice_t \times Positivity_t^{AN}$	2.31***	1.71***	1.29	1.36	1.95**	1.78**	3.53**	6.15**	8.90**	10.94***	2.56	4.09	4.06
	(0.51)	(0.55)	(1.01)	(1.24)	(0.78)	(0.70)	(1.37)	(2.99)	(4.01)	(3.71)	(2.33)	(2.83)	(2.80)
$Voice_t$	-0.17	-0.09	0.18	0.21	-0.15	0.01	0.13	-1.16**	-0.47	0.55	0.45	-0.06	0.27
	(0.18)	(0.17)	(0.24)	(0.22)	(0.19)	(0.22)	(0.25)	(0.47)	(0.65)	(0.69)	(0.35)	(0.45)	(0.55)
$Positivity_t^{AN}$	2.15***	1.67***	0.99	0.83	1.67**	0.65	1.44	0.63	1.34	2.31	1.58	1.65	1.51
	(0.49)	(0.43)	(0.75)	(1.11)	(0.71)	(0.63)	(1.24)	(1.75)	(2.34)	(2.65)	(1.39)	(1.90)	(2.23)
$Positivity_t^{IS}$	0.16	0.03	-0.19	-1.04**	0.43	1.15*	0.30	2.29	1.40	0.30	-0.98	0.43	-0.16
	(0.36)	(0.34)	(0.47)	(0.46)	(0.36)	(0.62)	(0.60)	(1.47)	(1.76)	(1.72)	(1.21)	(1.57)	(1.73)
$TextComplexity_t^{FOG}$	0.25**	0.23***	0.13*	-0.00	0.31***	0.24*	-0.02	0.17	-0.68	-0.78	-0.61**	-0.24	-0.55
	(0.11)	(0.08)	(0.07)	(0.11)	(0.07)	(0.12)	(0.13)	(0.44)	(0.52)	(0.47)	(0.30)	(0.40)	(0.36)
$Timing_t$	0.74***	0.92***	0.84***	0.00	0.82***	0.82***	0.11	0.49	0.60	0.16	0.83**	0.77*	0.24
	(0.16)	(0.15)	(0.18)	(0.20)	(0.15)	(0.19)	(0.22)	(0.41)	(0.51)	(0.46)	(0.33)	(0.41)	(0.41)
FG_t	0.75***	0.93***	1.15***	0.48***	0.80***	1.10***	0.63***	0.56**	0.51	0.33	0.43**	0.60**	0.36
	(0.07)	(0.09)	(0.14)	(0.13)	(0.09)	(0.16)	(0.18)	(0.24)	(0.31)	(0.35)	(0.18)	(0.25)	(0.33)
QE_t	0.20***	0.36***	0.69***	1.32***	0.37***	0.88***	1.43***	0.61***	0.72***	1.34***	0.34***	0.65***	1.25***
	(0.06)	(0.08)	(0.15)	(0.15)	(0.08)	(0.09)	(0.17)	(0.16)	(0.21)	(0.25)	(0.12)	(0.18)	(0.21)
$\Delta Inflation_t^{NCY,Positive}$	1.39	2.10	-0.15	-2.59	2.39	-0.60	-0.80	4.42	4.40	4.47	3.28	3.13	1.36
	(2.39)	(3.53)	(2.35)	(2.85)	(3.40)	(2.91)	(3.55)	(4.96)	(5.80)	(6.10)	(3.50)	(4.03)	(5.12)
$\Delta Inflation_t^{NCY,Negative}$	3.33**	2.11*	0.86	1.15	1.10	0.72	-0.80	8.64	0.28	2.23	-1.93	-1.19	-1.55
	(1.39)	(1.18)	(1.80)	(1.88)	(0.91)	(1.55)	(1.87)	(8.31)	(8.93)	(7.92)	(5.30)	(6.86)	(5.52)
$\Delta RGDP_t^{NCY,Positive}$	-1.60	-1.87	-2.76	0.73	-2.56	-1.49	0.49	-7.44	-7.46	-6.19	-3.80	-4.56	-3.19
	(1.82)	(2.55)	(1.97)	(2.53)	(2.49)	(2.16)	(2.98)	(5.01)	(5.96)	(5.77)	(3.47)	(4.23)	(4.47)
$\Delta RGDP_t^{NCY,Negative}$	0.39	-2.29	-7.32***	0.06	0.80	-4.77**	2.91	0.86	-0.80	1.88	-1.90	3.19	1.58
	(1.41)	(1.42)	(2.14)	(2.20)	(1.32)	(1.98)	(2.08)	(6.26)	(7.63)	(7.64)	(5.13)	(6.07)	(5.86)
$Constant$	-3.64**	-3.28***	-1.43	0.47	-4.52***	-3.91**	-0.14	-3.40	8.70	10.79	8.35*	2.55	7.39
	(1.51)	(1.04)	(1.05)	(1.58)	(1.03)	(1.72)	(1.77)	(5.96)	(7.14)	(6.43)	(4.17)	(5.37)	(5.04)
R^2	0.881	0.921	0.930	0.949	0.912	0.945	0.939	0.686	0.635	0.776	0.602	0.619	0.744
Obs	48	48	48	48	48	48	48	47	47	47	47	47	47

Note: We regress our variables on the change in the yields of government bonds of Germany, 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. $Voice_t$ is Draghi's net vocal sentiment during the Q&A session (see 3.1) and $Positivity_t^{AN}$ measures the net positivity of Draghi's answers. $Positivity_t^{IS}$ measures the framing of the introductory statement (see 3.3). We include a complexity indicator to measure to clarity of the language used during the Q&A session based on the FOG index. 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. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Appendix E.3. Robustness Check: Likert Scale Definition for Vocal Emotions

To ensure that our conclusions are robust with a more qualitative definition of vocal emotions, we define $Voice_t$ as follows:

$$Voice_t = \begin{cases} +2 & \text{if } \frac{Positive_t - Negative_t}{Positive_t + Negative_t} \in [+1,\ +0.6) \\ +1 & \text{if } \frac{Positive_t - Negative_t}{Positive_t + Negative_t} \in [+0.6,\ +0.2) \\ 0 & \text{if } \frac{Positive_t - Negative_t}{Positive_t + Negative_t} \in [+0.2,\ -0.2] \\ -1 & \text{if } \frac{Positive_t - Negative_t}{Positive_t + Negative_t} \in (-0.2,\ -0.6] \\ -2 & \text{if } \frac{Positive_t - Negative_t}{Positive_t + Negative_t} \in (-0.6,\ -1] \end{cases} \tag{E.5}$$

Our conclusions remain consistent even when using a qualitative classification of vocal emotions instead of a precise quantitative measurement.

Table E.3: **Robustness Check: Likert Scale Vocal Emotion Regression**

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)
	DE1Y	DE2Y	DE5Y	DE10Y	FR2Y	FR5Y	FR10Y	IT2Y	IT5Y	IT10Y	ES2Y	ES5Y	ES10Y
$Voice_t \times Positivity_t^{AN}$	0.86***	0.61**	0.26	0.32	0.71**	0.63**	1.49**	2.27*	3.02	4.38**	0.67	1.46	1.35
	(0.25)	(0.24)	(0.41)	(0.52)	(0.33)	(0.26)	(0.61)	(1.33)	(1.98)	(1.87)	(1.19)	(1.35)	(1.33)
$Voice_t$	-0.01	-0.06	0.02	0.18	-0.06	0.01	0.21	-0.52*	-0.30	0.36	0.10	-0.19	0.08
	(0.09)	(0.10)	(0.14)	(0.13)	(0.11)	(0.11)	(0.17)	(0.30)	(0.44)	(0.44)	(0.25)	(0.28)	(0.32)
$Positivity_t^{AN}$	1.13**	0.95**	0.35	0.06	0.82*	-0.16	-0.11	-1.60	-2.05	-2.15	0.53	0.19	-0.14
	(0.46)	(0.36)	(0.54)	(0.69)	(0.47)	(0.54)	(0.81)	(1.24)	(1.97)	(2.13)	(1.10)	(1.51)	(1.66)
$Positivity_t^{IS}$	0.18	0.14	-0.02	-1.04**	0.50	1.24*	0.34	2.47*	2.10	1.05	-0.54	0.90	0.27
	(0.38)	(0.35)	(0.50)	(0.48)	(0.38)	(0.61)	(0.64)	(1.33)	(1.76)	(1.80)	(1.26)	(1.46)	(1.71)
$Timing_t$	0.89***	1.04***	0.91***	0.05	0.98***	0.94***	0.20	0.72*	0.61	0.17	0.67**	0.77*	0.15
	(0.16)	(0.16)	(0.17)	(0.21)	(0.16)	(0.19)	(0.22)	(0.39)	(0.47)	(0.45)	(0.32)	(0.39)	(0.41)
FG_t	0.82***	0.98***	1.18***	0.53***	0.86***	1.16***	0.72***	0.66***	0.62*	0.52	0.44**	0.63**	0.41
	(0.08)	(0.10)	(0.13)	(0.13)	(0.10)	(0.16)	(0.18)	(0.24)	(0.33)	(0.38)	(0.19)	(0.26)	(0.35)
QE_t	0.16**	0.34***	0.70***	1.32***	0.34***	0.86***	1.40***	0.56***	0.78***	1.36***	0.42**	0.69***	1.30***
	(0.06)	(0.08)	(0.15)	(0.14)	(0.08)	(0.09)	(0.16)	(0.19)	(0.25)	(0.26)	(0.16)	(0.20)	(0.21)
$\Delta Inflation_t^{NCY,Positive}$	0.41	1.76	-0.51	-3.80	1.87	-1.11	-2.59	1.92	-0.04	-0.63	1.68	1.97	-0.87
	(2.28)	(3.23)	(2.22)	(2.57)	(3.11)	(2.63)	(3.10)	(4.45)	(4.95)	(5.19)	(2.90)	(3.26)	(4.43)
$\Delta Inflation_t^{NCY,Negative}$	4.82**	3.41**	1.59	1.37	2.78*	2.07	-0.19	10.00	-1.88	0.45	-4.46	-1.84	-3.53
	(2.06)	(1.49)	(1.58)	(1.57)	(1.48)	(1.42)	(1.55)	(7.07)	(7.95)	(8.05)	(5.40)	(5.23)	(6.05)
$\Delta RGDP_t^{NCY,Positive}$	-0.30	-1.15	-2.10	1.94	-1.62	-0.61	2.47	-4.46	-2.94	-0.81	-2.43	-3.08	-1.07
	(1.82)	(2.33)	(1.80)	(2.07)	(2.26)	(1.91)	(2.48)	(3.80)	(4.55)	(4.74)	(2.66)	(3.10)	(3.66)
$\Delta RGDP_t^{NCY,Negative}$	-0.58	-2.88**	-7.77***	-0.70	0.08	-5.42***	1.32	-1.92	-5.12	-3.10	-3.25	1.53	-0.43
	(1.49)	(1.38)	(1.95)	(1.79)	(1.18)	(1.88)	(1.60)	(5.63)	(6.64)	(6.44)	(4.99)	(5.46)	(5.42)
$Constant$	-0.40**	-0.25	0.14	0.34	-0.49**	-0.88**	-0.54*	-1.12	-0.71	-0.38	-0.13	-0.87	-0.22
	(0.19)	(0.18)	(0.24)	(0.24)	(0.20)	(0.33)	(0.31)	(0.74)	(1.00)	(0.99)	(0.69)	(0.83)	(0.91)
R^2	0.862	0.913	0.926	0.949	0.896	0.941	0.939	0.669	0.582	0.740	0.530	0.612	0.721
Obs	48	48	48	48	48	48	48	47	47	47	47	47	47

Note: We regress our variables on the change in the yield of government bonds of Germany, 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. $Voice_t$ is Draghi's net vocal sentiment during the Q&A session using a Likert scale definition and $Positivity_t^{AN}$ measures the net positivity of Draghi's answers. $Positivity_t^{IS}$ measures the framing of the introductory statement (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. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$