

The Voice of Monetary Policy[†]

By YURIY GORODNICHENKO, THO PHAM, AND OLEKSANDR TALAVERA*

We develop a deep learning model to detect emotions embedded in press conferences after the Federal Open Market Committee meetings and examine the influence of the detected emotions on financial markets. We find that, after controlling for the Federal Reserve's actions and the sentiment in policy texts, a positive tone in the voices of Federal Reserve chairs leads to significant increases in share prices. Other financial variables also respond to vocal cues from the chairs. Hence, how policy messages are communicated can move the financial market. Our results provide implications for improving the effectiveness of central bank communications. (JEL D83, E31, E44, E52, E58, F31, G14)

How can a president not be an actor?

—Ronald Reagan (1980)

As Chairman, I hope to foster a public conversation about what the Fed is doing to support a strong and resilient economy. And one practical step in doing so is to have a press conference like this after every one of our scheduled FOMC meetings. ... [This] is only about improving communications.

—Jerome Powell (2018)¹

Monetary policy is 98 percent talk and 2 percent action, and communication is a big part.

—Ben Bernanke (2022)²

*Gorodnichenko: Department of Economics, University of California-Berkeley (email: ygorodni@econ.berkeley.edu); Pham: Department of Economics, University of Reading (email: t.pham@reading.ac.uk); Talavera: Department of Economics, Birmingham Business School, University of Birmingham (email: o.talavera@bham.ac.uk). Isaiah Andrews was the coeditor for this article. We thank Annette Vissing-Jorgensen, Michael Ehrmann, William English, anonymous referees, and seminar and conference participants at the University of Reading, the University of Birmingham, the University of East Anglia, UC Davis, FED Cleveland, FED Richmond, Economic Policy Working Group—Hoover Institution, NBER SI 2021, ESCoE Economic Measurement 2021 conference, Advanced Analytics: New Methods and Applications for Macroeconomic Policy conference, Non-traditional Data, Machine Learning and Natural Language Processing in Macroeconomics conference, and the Workshop on Empirical Monetary Economics for their comments. We are grateful to Eric Swanson for sharing his series of monetary policy shocks.

[†]Go to <https://doi.org/10.1257/aer.20220129> to visit the article page for additional materials and author disclosure statements.

¹<https://www.federalreserve.gov/mediacenter/files/FOMCpresconf20180613.pdf> (accessed on July 21, 2021).

²<https://www.brookings.edu/events/ben-bernanke-the-fed-from-the-great-inflation-to-covid-19/> (accessed on May 24, 2022).

In a famous analysis, Mehrabian (1971) posited a 7–55–38 rule of communication: the words convey 7 percent of a message, the body language (gestures, facial expressions, etc.) accounts for 55 percent, and the tone delivers 38 percent. While the debate on exact percentages for each channel is open, it is clear that effective communication has to involve more than just words. Central banks have been relying increasingly on communication-based tools (e.g., forward guidance) to manage the public's expectations, but do central bankers utilize communication to its full potential?

Textual analyses of policy statements, minutes, and transcripts (e.g., Rozkrut et al. 2007; Hansen and McMahon 2016; Hansen, McMahon, and Prat 2018; Cieslak, Morse, and Vissing-Jorgensen 2019; Ehrmann and Talmi 2020) suggest that central bankers' words carry considerable weight,³ but little is known about the effects of their nonverbal communication. To shed further light on this issue, we use deep learning methods to quantify tone (vocal emotions) embedded in the answers given by Federal Reserve (Fed) chairs during press conferences. To the best of our knowledge, our study is the first to examine the effects of central bank communications through the vocal dimension. In other words, we move beyond text analysis and study *how* policy messages are voiced and whether emotions channeled through voice tone can move financial markets. This offers a new tool for communicating with the public and for managing expectations.

We focus on policy communication during press conferences for several reasons. First, press conferences have been commonly used as an important communication tool. As suggested by Ehrmann and Fratzche (2007) and emphasized by Jerome Powell,⁴ press conferences, particularly the Q&A sessions, play a key role in helping financial markets and the general public to understand policy outlook and the interpretation of current economic conditions. Especially during times of high uncertainty, market participants tend to seek further guidance and clarification through press conferences' Q&A sessions. Second, press conferences allow policymakers to go off script and to communicate soft information via nonverbal channels, thus potentially influencing investors' decision-making.⁵ Finally, because video-audio recordings of press conferences are available in a consistent format, we can measure the tone of communication in a consistent manner and provide a systematic analysis of how voice tone can influence economic outcomes.

Specifically, we split a given Federal Open Market Committee (FOMC) press conference during the April 2011–June 2019 period into audio segments corresponding to the response of the speaker to each question raised during the event. The split audios are then run through a machine learning model, which is trained to recognize emotions from voice variations. Each answer is rated as being one of three emotion classes: positive (happy or pleasantly surprised emotions), negative (sad or angry emotions), and neutral. After aggregating the tone of the answers for a given

³More generally, central banks have significant power to influence the macroeconomy and expectations. For example, a large number of studies have documented the effectiveness of policy announcements in moving financial markets (e.g., Kuttner 2001; Gürkaynak, Sack, and Swanson 2005a) or shaping firm and household inflation expectations (e.g., Coibion, Gorodnichenko, and Weber 2019; Enders, Hünnekes, and Müller 2019; Lamla and Vinogradov 2019).

⁴<https://www.federalreserve.gov/mediacenter/files/FOMCpresconf20180613.pdf> (accessed on July 21, 2021).

⁵Caffi and Janney (1994); Visschedijk et al. (2013); Dricu et al. (2017); and others document that voice conveys information beyond the verbal content and that information contained in voice can affect decision-making.

press conference, we examine how the tone affects a variety of financial variables at high frequencies. We find that the tone can materially move financial markets. For example, making the voice tone more positive by one standard deviation could raise S&P 500 returns by approximately 75 basis points. This order of magnitude is comparable to what one can observe after a one standard deviation shock to forward guidance. In other words, the voice component can generate economically significant effects on the stock market. We also find that inflation expectations and exchange rates respond to variations in voice tone, e.g., a more positive tone leads to a decrease in expected inflation. At the same time, the evidence for the bond market is more mixed in our sample. These results suggest that policy communication is more nuanced than reading and posting prepared remarks and speeches. It appears that a certain level of acting skill may be helpful for ensuring that the public receives the policy message fully and accurately.

In addition to the vast literature on policy communication (see Blinder et al. 2008 and Coenen et al. 2017 for comprehensive surveys) and high-frequency analyses of monetary policy shocks (Kuttner 2001; Gürkaynak, Sack and Swanson 2005a; and many others), our study relates to research investigating the economic impacts of vocal cues. Using a sample of CEO speeches made during earnings conference calls, Hobson, Mayew, and Venkatachalam (2012) find that the vocal markers of cognitive dissonance can predict the likelihood of irregular restatements of earnings reports. In a related study, Mayew and Venkatachalam (2012) show that market participants and analysts react to the affective states of managers expressed through vocal cues, such as happy or unhappy voices. For example, a positive affect is positively related to changes in stock recommendations and future unexpected earnings. These results suggest that the affective states contain useful information about a firm's fundamentals. In a more recent study, Hu and Ma (2020) find that positivity about a startup, shown through visual, verbal, and vocal dimensions, increases the likelihood of being funded, even if the startup's quality is low. Apart from having a different focus (central banking communication versus CEO/manager communication), our study differs in terms of the tools employed to quantify the variation in tone. Earlier studies use commercial software or pretrained machine learning algorithms for voice analysis, while we develop a customized deep learning model for detecting speech emotion. Our approach offers several advantages in terms of flexibility and the potential for further development and implementation. For example, we can fine-tune the model's parameters to achieve a higher accuracy rate, which is not a feature available in commercial software. Similarly, the customized model also allows us to adjust the number and class of emotions, which cannot be done with commercial software and pretrained algorithms.

Curti and Kazinnik (2021), a concurrent paper that is closest in spirit to our work, examine the responses of the financial market to variations in the chair's facial expressions during post-FOMC press conferences. Using intraday data, they find that negative facial expressions lead to lower stock-market returns. We view their results as reinforcing our message that nonverbal communication can move the financial markets, and hence, the nonverbal component is a potentially important channel of communication for economic players.

The rest of this paper is organized as follows. In the next section, we provide anecdotal evidence to justify the importance of the nonverbal channels in monetary

policy communication. Section II describes the deep learning algorithms used to analyze the tone of voice and the sentiment of policy texts. In Section III, we discuss our main results and their robustness to a series of robustness checks. We also provide some tentative interpretations and explanations of our results in this section. Finally, Section IV concludes and discusses the implications of the results.

I. Nonverbal Communication and Monetary Policy

Why would nonverbal communication matter? One explanation is that, due to asymmetric information between the public and the central bank, market participants tend to seek additional information through the aspects which are not explicitly “scripted,” such as the choice of words, tone of voice, or the body language of the Fed chair. As with words, the nonverbal elements of communication can signal the Fed’s perspective on the current and future economic outlook and the future course of monetary policy. The Fed chairs understand that press conferences are high-stake interactions with the public and the media and that communication is a complex process. For example, in her closing remarks for the FOMC meeting on December 16, 2015, Janet Yellen said:

Okay. Boxed lunches will be available. If anybody wants to watch TV in the Special Library and see me get skewered at the press conference, please feel free. I will do my best to communicate the points that have been made here. END OF MEETING.⁶

Unsurprisingly, the Fed invests significant resources into preparation for press conferences, as well as postconference analysis.

We also know that investors and the media watch and listen to FOMC press conferences, analyze the chair’s voice, and attempt to interpret what it (i.e., the voice tone, emotion, etc.) implies. For example, in the *International Quest Means Business* program aired on June 22, 2011, Felicia Taylor, a business reporter for CNN, said while covering an FOMC press conference:

The press conference, though, that is coming up in just a few minutes is where traders are really going to be looking for every little nuance. They want to see how [Ben Bernanke] is going to read into everything. The tone of his voice, his body language, his inflection, for any clue about the direction the markets are still looking for.

This quote suggests that the press and financial market investors appear to pay attention to nonverbal communication.⁷ Indeed, it is not uncommon for media reports to assess the nonverbal elements of press conferences. For example, the *Wall Street Journal*⁸ reported Ben Bernanke’s voice as either shaky or quavering during

⁶ <https://www.federalreserve.gov/monetarypolicy/files/FOMC20151216meeting.pdf> (accessed on December 16, 2022)

⁷ See, e.g., <https://cutt.ly/FzbnCBo>; <https://cutt.ly/fzbn1le> (accessed on July 21, 2021)

⁸ <https://www.wsj.com/articles/BL-REB-13815> (accessed on July 21, 2021)

the first FOMC press conference on April 27, 2011. The tone of Powell's voice at the press conference on September 16, 2020 was perceived to be consistent with previous press conferences, which was interpreted as a signal of downplaying his dovish position.⁹

These anecdotes suggest that nonverbal communication could be an important channel. To study this channel systematically—in an objective, reproducible fashion—we build on recent advances in voice recognition technology and classify the voice tone of the Fed chairs into a spectrum of emotions. We, then, study how variations in voice tone (emotions) can affect financial variables.

II. FOMC Speeches: Voice and Linguistic Analysis

Our sample runs from April 2011 (when the first FOMC press conference was held) to June 2019. During this period, 68 meetings and 36 press conferences were held. For each meeting, the FOMC statement and the transcript of the press conference are obtained from the Fed website. The press conference videos are downloaded from the Fed's official YouTube channel. We use only the audio component of these videos.¹⁰ As the Q&A session is the only part of the press conference that is not scripted, our analysis focuses on the answers of the chair during the Q&A.

A. Voice Tone

In this section, we describe how we train a neural network model (a deep learning algorithm) to recognize emotions and refer the readers to online Appendix A for more details. Conceptually, it is necessary to create a mapping from various measures of sound waves into emotions. We also present descriptive statistics for variations in voice tone during post-FOMC press conferences.

Emotion Detection Using Neural Networks.—Voice can be characterized by various parameters such as pitch (indicating the level of highness/lowness of a tone) and frequency (indicating the variation in the pitch) which are useful for determining the emotion of a speaker. Building on earlier research on voice recognition (e.g., Pan, Shen, and Shen 2012; Gao et al. 2017; Likitha et al. 2017; Bhavan, Chauhan, and Shah 2019), we use Librosa, a Python package, to extract the following vocal features. First, we extract 128 mel spectrogram frequencies which allows us to determine the level of loudness of a particular frequency at a particular time. Second, a chromagram with 12 chroma coefficients is extracted. The chromagram reflects the distribution of energy along 12 chroma bands (i.e., C, C#, D, D#, E, F, F#, G, G#, A, A#, and B) over time and, hence, can capture melodic and harmonic characteristics of audio. Finally, we extract 40 mel-frequency cepstral coefficients (MFCCs), which are discrete cosine transformations of the mel frequency spectrogram. Although MFCCs can be derived from mel spectrogram frequencies, we find

⁹<https://cutt.ly/RzbnAm5> (accessed on July 21, 2021)

¹⁰Video recordings can be used to study body language (e.g., facial expressions, gestures). However, videos are harder to analyze because cameras are moving (different speakers, different angles). In this respect, audio tracks offer a more consistent method of measurement.

that using both types of features helps to improve the accuracy of the model. Note that the number of mel spectrogram coefficients, MFCCs, and chroma coefficients can be adjusted to achieve more accurate predictions.

Once the inputs from an audio track are constructed, we need measures of emotions corresponding to the audio track so that we can build a model to classify sounds into emotions. We use two datasets where emotions can be matched to audio tracks. The first (Livingstone and Russo 2018) is the Ryson Audio-Visual Database of Emotional Speech and Song (RAVDESS). To create these data, 12 actors and 12 actresses spoke two statements in a neutral North American accent using 8 different emotions (calm, happy, sad, angry, fearful, surprised, disgusted, and neutral). The second dataset (Pichora-Fuller and Dupuis 2010) is the Toronto Emotional Speech Set (TESS). To create these data, two actresses spoke a set of 200 words using 7 emotions (happy, sad, angry, fearful, pleasantly surprised, disgusted, and neutral). These datasets are widely used in the computer science literature to build speech emotion/expression systems (see, e.g., Verma and Mukhopadhyay 2016; Gao et al. 2017; Choudhury et al. 2018; Bhavan, Chauhan, and Shah 2019; Zisad, Hossain, and Andersson 2020). As emotions of fear and disgust are unlikely to arise during the Q&A sessions, we only use audios for five emotions: happy, (pleasantly) surprised, neutral, sad, and angry. We use 80 percent of RAVDESS and TESS as the training sample and the remaining 20 percent are used for testing.

After extracting the vocal features from each recording in RAVDESS and TESS, we use Keras, a deep learning application programming interface (API) run on top of Google's machine learning platform, TensorFlow, to build a neural network model, i.e., a computing system consisting of different layers where each layer is a collection of different neurons (nodes). In this study, we build a fully connected network with an architecture specified as follows. The first layer takes 180 vocal features (128 mel coefficients, 40 MFCCs, and 12 chroma coefficients) as inputs to produce 200 nodes as outputs. The second layer has 200 nodes which are connected with 200 nodes in the first layer through the linear activation function. The third layer has 200 nodes which are connected with 200 nodes in the previous layer through the linear activation function. The output layer has five nodes representing five emotions (happy, pleasantly surprised, neutral, sad, and angry). Given that our task is a multigroup classification task, we use the softmax activation function (i.e., normalized exponential function) to connect the nodes in this layer with 200 nodes in the second hidden layer.¹¹ To minimize overfitting, we add three dropout layers with a dropout rate of 0.3 after each layer prior to the output layer. This means that 30 percent of the inputs are randomly set to 0 at each step during the training time (hence, only 70 percent of the inputs are retained for training). After training the model, we use the accuracy score to evaluate the model's performance:

$$(1) \quad Accuracy(y, \hat{y}) = \frac{1}{n} \sum_{i=1}^n \mathbf{1}\{\hat{y}_i = y_i\},$$

¹¹In other words, a linear model is applied to all layers except the output layer where a multinomial logistic model is applied. Thus, the output in the output layer is a probability distribution over five emotion classes and the sum of all probabilities is equal to one. The predicted emotion is the emotion class which has the highest probability.

where y and \hat{y} are the true emotion and the predicted emotion, respectively, and n is the number of audio files in the testing dataset. The trained model gives us an overall accuracy score of 84 percent. Applying the same formula for each emotion, we obtain accuracy levels for predicting angry, sad, neutral, surprised, and happy emotions of 87 percent, 84 percent, 74 percent, 87 percent, and 80 percent, respectively.

A key advantage of our approach to classifying voice tone into emotions is its objectivity and reproducibility. Indeed, any interested researcher can apply this tool—or its variations—to trace every step of the approach and to measure sensitivity to various assumptions and data points. In contrast, using emotions obtained from human classification is not only much more costly but is also likely to involve judgment, biases, conflicting interpretations, and potentially too great a reliance on “reading between the lines.” However, humans may be better at detecting subtle tone variations.

FOMC Audio Input.—We manually processed and split each FOMC press conference’s audio into smaller audio segments where each segment is a chair’s answer to a question. As an illustration of our approach, consider the first press conference given by Ben Bernanke on April 27, 2011. After giving opening remarks, he invited the audience to ask questions (11:23 mark of the press conference). The first question (about a weak GDP forecast) ended at 11:49. Bernanke’s response lasted until 12:59. Hence, the first audio segment runs for 1 minute and 10 seconds from 11:49 to 12:59. We process other segments in the same manner. In the end, our audio sample contains 692 answers from three speakers (Bernanke, Yellen, and Powell). The lengths of the audio segments vary from 10 seconds to more than 5 minutes, but most answers are between 1 and 3 minutes long. The number of audio segments per press conference ranges from 12 to 26, with an average of 19.

Emotion Detection Output.—After training the neural network to recognize emotions from variations in vocal features (MFCCs, chromagram, and mel spectrogram), we feed the audio tracks of the policymakers’ answers into the neural network. Each audio segment receives one predicted emotion (happy, pleasantly surprised, neutral, sad, or angry). We then classify each audio segment into positive (predicted emotion is “happy” or “pleasantly surprised”), negative (predicted emotion is “sad” or “angry”), or neutral. We aggregate the tone to the press conference level as follows:

$$(2) \quad \text{VoiceTone} = \frac{\text{Positive answers} - \text{Negative answers}}{\text{Positive answers} + \text{Negative answers}},$$

where *VoiceTone* ranges from -1 (negative emotions) to $+1$ (positive emotions).¹²

We find that Ben Bernanke, on average, had more positive emotions in his voice than Janet Yellen, who, in turn, had more positive emotions in her voice than Jerome Powell. Bernanke had five Q&A sessions with only positive emotions in his voice. In contrast, Jerome Powell had five Q&A sessions with only negative emotions. Janet Yellen’s sessions always had a mix of positive and negative emotions. The

¹²We report descriptive statistics in panel A of Table 1 and scores for each Q&A session in online Appendix Table A1.

average tone across these central bankers is close to zero. There is considerable within-speaker variation in tone, with Jerome Powell exhibiting the largest variation.

Although we do not have a standard benchmark to validate *VoiceTone*, we can use reports in the media and other external information to get a sense of how *VoiceTone* aligns with other sources. For example, Ben Bernanke tends to get positive score values of *VoiceTone* for most of the meetings, but his score is unusually low (close to -1) for the press conference on September 18, 2013. We can recall that the preceding press conference (June 1, 2013) resulted in the “taper tantrum” when Bernanke communicated that the Fed planned to scale back quantitative easing and surprised the markets. Between this press conference and the next (September 18), Bernanke and other Fed officials walked back on that message. In the September 18 press conference, Bernanke was on the defensive when he was repeatedly asked about the “mistake” made in the previous press conference and his personal responsibility for it.¹³ He also had to say that the Fed was working with untested policies. In addition, Bernanke had to answer questions about his own future (i.e., Janet Yellen was expected to be officially nominated to chair the Fed) and his regrets. Thus, it is natural that he sounded relatively negative.

Consider now the press conference on September 17, 2015, when Janet Yellen had an unusually positive score (0.83). During this press conference, Yellen voiced her view that the economy has a lot more space to grow despite a relatively low unemployment rate of 6.3 percent, close to the natural rate perceived at the time, and a spike in inflation. The media reported that “... Yellen seemed slightly unconvinced by the Fed’s insistence that the economy is improving”¹⁴, “[Yellen] expressed skepticism that the official headline unemployment rate (6.3 percent) was an adequate reflection of labor market tightness”¹⁵, and “Yellen struck a dovish tone”¹⁶. Perhaps in her attempts to signal her more positive outlook for the economy, Yellen sounded more positive than usual.

Powell’s first press conference on March 21, 2018 had a low score (-1) when, according to the *Financial Times*¹⁷, “Mr Powell seemed anxious to underline the uncertainty hanging over the outlook rather than sending up too many hawkish warning flares.” A year later (March 20, 2019), Powell had one of his most positive scores in our sample (0.92). The media coverage suggested then that Powell “at every turn managed to out dovish expectations”¹⁸.

Obviously, these cross-checks should be interpreted as tentative. However, they suggest that there is some consistency between our scores on the one hand and the nuances detected by the media on the other hand.

¹³For example, Binyamin Appelbaum (*New York Times*) asked, “To what extent do you regard yourself as responsible for the tightening in financial conditions that you noted? Was it a mistake to talk about tapering in the way that you did in June and do you stand by your guidance that it will be appropriate?” See <https://www.federalreserve.gov/mediacenter/files/FOMCpresconf20130918.pdf> (accessed December 16, 2022).

¹⁴<https://www.theguardian.com/commentisfree/2014/jun/18/fed-janet-yellen-economy-inflation> (accessed on January 16, 2022).

¹⁵<https://www.businessinsider.com/janet-yellen-dove-2014-6> (accessed on January 16, 2022).

¹⁶<https://www.ftadvise.com/2015/10/06/investments/fixed-income/markets-no-longer-listening-to-fed-vGD6k9R3rbm3YQYknBjsTN/article.html> (accessed on January 16, 2022).

¹⁷<https://www.ft.com/content/1047a2a8-6f7c-11e8-92d3-6c13e5c92914> (accessed on January 16, 2022).

¹⁸<https://www.afr.com/markets/equity-markets/asx-advances-on-dovish-fed-20190322-p516gq> (accessed on January 16, 2022).

B. Textual Analysis

Successful policy communication should utilize multiple channels to guide the public in a desired direction. To avoid a cacophony and, thereby, confusion, verbal and nonverbal messages should be congruent and reinforce each other. However, if policy communication is a concerted effort, how can one hear the voice (tone) of monetary policy? In the absence of exogenous variation in how policy messages are telegraphed to the public, identifying the effects of voice tone has to rely on controlling for the text sentiment of policy messages as well as policy actions. To this end, we employ state-of-the-art natural language processing tools to quantify the sentiment of FOMC texts.

Sentiment Analysis Using BERT Embeddings.— To extract the word embeddings¹⁹ of policy texts, we adopt Bidirectional Encoder Representations Transformer (BERT), the natural language processing algorithm developed by Google AI, which has several advantages over other tools. First, unlike context-free models (e.g., Word2Vec, GloVe), BERT generates an embedding representation of a word based on its surrounding context (contextual representation). Second, in contrast to unidirectional contextual models (e.g., ELMo, ULMFit), which create a word's representation based on previous words in the text, BERT is a bidirectional model which takes into account both preceding and subsequent context to generate the embeddings of a word.²⁰ As a result of these features, BERT has very high accuracy in interpreting texts (Devlin et al. 2018) and has been increasingly applied in economic research (see, e.g., Chava, Du, and Paradkar 2020; Chava, Du, and Malakar 2021).

Although pretrained BERT models can assign certain interpretations to texts (e.g., positive or negative), previous works on the textual analysis of policy communication have focused on the dovish-hawkish spectrum. To bridge our work to earlier studies, we need to undertake an additional step. Specifically, we create a customized training dataset that includes all FOMC statements released between 1997 and 2010 and the individual sentences in these statements. Each text in the training data is independently scored by several research assistants. The scores run from -10 (very hawkish) to $+10$ (very dovish). We calculate the average score for each text and classify a text as dovish (the average score ≥ 0.5), hawkish (the average score ≤ -0.5), or neutral (the average score is between -0.5 and $+0.5$). We then use the word embeddings obtained from BERT as the inputs for a neural network²¹ tailored to identify the sentiment of monetary policy texts. In other

¹⁹Simply put, word embeddings (mappings of words to vectors of real numbers) capture the semantics of the words (i.e., the meanings of the words) and the syntactic relationships between them (i.e., the grammatical structure).

²⁰There are different versions of BERT embeddings, depending on the training data and the architecture. In this study, we use word embeddings generated from the BERT base model (12 layers, 768 hidden states, 12 heads, and 110 million parameters). As a robustness check, we also use the embeddings obtained from the Robustly Optimized BERT Pretraining Approach (RoBERTa) model (12 layers, 768 hidden states, 12 heads, and 125 million parameters).

²¹The architecture of this neural network is as follows. The input layer is the output of the last layer of the BERT embedding model, which has the input dimension of 512×768 . In the first hidden layer, a long short-term memory (LSTM) layer is wrapped with a bidirectional layer (bidirectional LSTM). The LSTM layer has 512 units and is connected with the input layer through the hyperbolic tangent activation function. The output of this bidirectional LSTM layer is passed on a global average pooling 1D layer to transform into one-dimensional data. The third hid-

words, we map texts processed with BERT into three categories: dovish, hawkish, and neutral. Applying formula (1), we obtain an overall accuracy score of 81 percent and the accuracy scores in predicting hawkish, neutral, and dovish sentiments are 85 percent, 79 percent, and 77 percent, respectively.

After applying this procedure (i.e., the BERT embedding model and the trained neural network for sentiment classification) to the text in the 2011–2019 sample, we aggregate the sentiment of the text from an FOMC statement, remarks, and Q&A as follows:

$$(3) \quad \text{TextSentiment} = \frac{\text{Dovish text} - \text{Hawkish text}}{\text{Dovish text} + \text{Hawkish text}}$$

where *Dovish text* and *Hawkish text* are the counts of respective paragraphs in the FOMC statements and opening remarks, as well as the counts of respective answers when a press conference is held. By construction, *TextSentiment* is in the range of $[-1; 1]$ and a positive value indicates that an expansionary monetary policy is expected or is being implemented.

Sentiment Analysis Output.— We find that the sentiment of texts during the terms of Yellen and Bernanke was generally more dovish than Powell’s (panel B of Table 1).²² This pattern likely reflects that policy rate increases dominated during Powell’s period in our sample. The within-speaker variation in the text sentiment is broadly similar across the Fed chairs. The correlation between sentiment for statements and remarks is discernibly positive ($\rho = 0.37$), while the correlation between the text sentiment of responses during Q&A sessions and the text sentiment of statements is slightly lower ($\rho = 0.21$). At the same time, the text sentiments for Q&A and remarks are correlated at $\rho = 0.13$. To measure the totality of the sentiment, we compute *TextSentiment* using all three sources (statement, remarks, and Q&A responses). Given that we have a limited number of events in our sample, this approach allows us to save degrees of freedom, but our results are robust to using *TextSentiment* measured for each source separately or in other combinations.

C. Comovement in Policy Actions, Words, and Tone

To what extent do text sentiment and voice tone comove? Although one might think that text and voice should be highly congruent, Figure 1 demonstrates that the relationship between these two channels of communication is more nuanced. Specifically, the positive messages conveyed in the tonality of voice are associated with more dovish statements in the accompanying texts. Hence, we observe congruence in words and tone, but this relationship is not perfect. For example, the correlation between the text sentiment in statements and the voice tone in the corresponding Q&A sessions is $\rho = 0.37$ (Spearman correlation is equal to 0.30).

den layer is a dense layer with 512 nodes and the fourth hidden layer is a dense layer with 128 nodes. A dropout layer (the dropout rate of 0.1) is added after the first, second, and third hidden layers. The output layer has three nodes representing three sentiment categories (hawkish, neutral, and dovish). Similar to the neural network used for voice emotion classification, we use the softmax activation function to connect the output layer with the third hidden layer. See online Appendix C for more details.

²²The text sentiment scores for each meeting are listed in online Appendix Table A2.

TABLE 1—FOMC MEETING STATISTICS

	All (1)	Bernanke (2)	Yellen (3)	Powell (4)
Meetings	68	25	32	11
Press conferences	36	12	16	8
<i>Panel A. Voice analysis of responses in Q&A during press conferences</i>				
Answers (count)				
Positive	377	200	109	68
Negative	285	43	131	111
Neutral	30	0	28	2
Voice tone				
Mean	0.09	0.64	−0.13	−0.30
Standard deviation	0.75	0.58	0.61	0.82
<i>Panel B. Textual analysis</i>				
Statement				
Hawkish paragraphs	37	8	20	9
Dovish paragraphs	223	105	108	10
Neutral paragraphs	64	6	37	21
Text sentiment				
Mean	0.66	0.86	0.71	0.09
Standard deviation	0.53	0.27	0.34	0.94
Remarks				
Hawkish paragraphs	83	15	47	21
Dovish paragraphs	244	97	106	41
Neutral paragraphs	119	41	50	28
Text sentiment				
Mean	0.50	0.75	0.38	0.35
Standard deviation	0.37	0.24	0.33	0.47
Q&A				
Hawkish answers	233	77	95	61
Dovish answers	339	137	120	82
Neutral answers	120	29	53	38
Text sentiment				
Mean	0.18	0.29	0.11	0.14
Standard deviation	0.30	0.30	0.26	0.35
Statement, Remarks, Q&A				
Text sentiment				
Mean	0.52	0.71	0.50	0.15
Standard deviation	0.40	0.30	0.35	0.52

Notes: This table shows the statistics related to the text and voice data of FOMC meetings and press conferences. Column 1 shows statistics for all FOMC meetings during the 2011–July 2019 period. Columns 2–4 show the statistics for the FOMC meetings chaired by Ben Bernanke, Janet Yellen, and Jerome Powell, respectively. Positive, negative, and neutral indicate the number of answers expressed in the positive, negative, and neutral emotion, respectively. Voice tone is the average emotion for a given FOMC press conference (as measured by equation (2)). Hawkish and Dovish are the number of hawkish and dovish answers/sentences in the text, respectively. The average text sentiment is measured by equation (3).

Similarly, the correlation between the Q&A tone and the text sentiment in remarks or Q&A is 0.48 and 0.29, respectively. Figure 1 shows that it is not uncommon to observe dovish texts and negative tonality. These results suggest that the tone of Q&A responses may generate variation in policy communication that is unrelated to the content of the texts of policymakers' statements, remarks, or even the Q&A responses themselves.

In a similar spirit, the variation in tone appears to be only weakly correlated with actual policy shocks (panels A–C in Figure 2) as identified in Swanson (2021): a shock to the policy rate (FFR shock), a forward guidance (FG) shock, or an asset

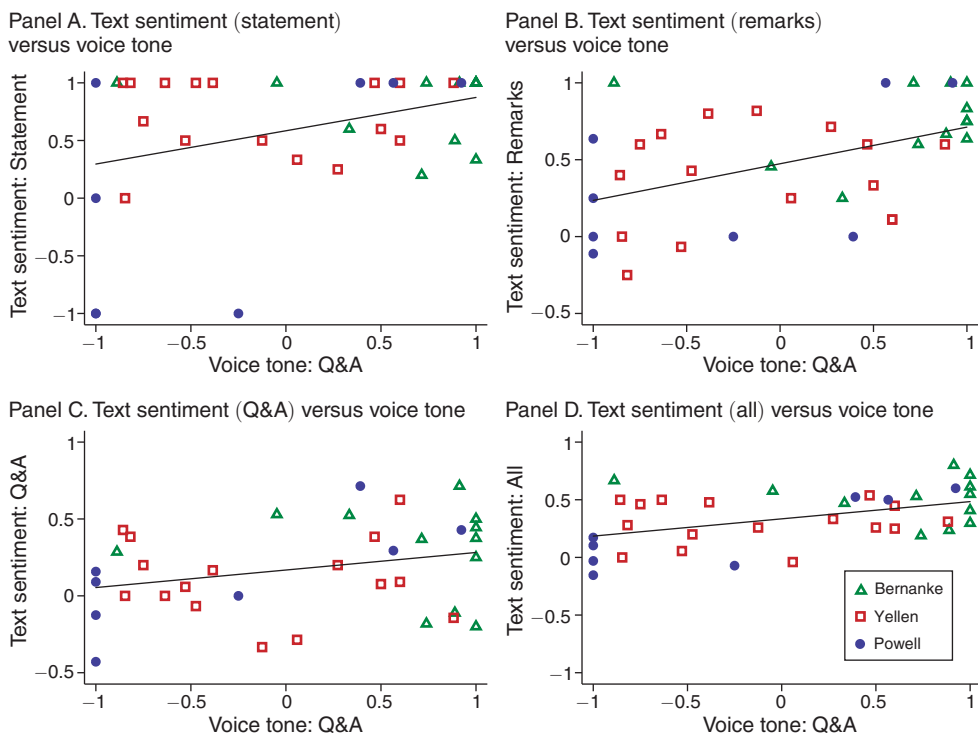


FIGURE 1. VOICE TONE VERSUS TEXT SENTIMENT

Note: This figure shows the joint distribution of voice tone and text sentiment across FOMC meetings.

purchase (AP) shock. There is a slightly stronger correlation between voice tone and the stage of the policy cycle. Specifically, the correlation between the shadow rate (as measured in Wu and Xia 2016)²³ and voice tone is -0.23 (for comparison, the correlation with FFR shocks is -0.19), i.e., the tone of voice becomes more negative as the policy rate increases.

These results suggest that communication is done via multiple channels, and it is important to control for these channels if we are to isolate the effect of voice tone. However, these correlations are far from perfect, and thus we have independent variation in voice tone. In part, these imperfect correlations can reflect the nature of the sample period. For example, one would expect that Fed funds rate movements (a standard measure of policy) and voice tone variations should be congruent. But the Fed funds rate was at the zero lower bound for most of the sample period and, hence, the correlation is naturally lower than one would expect in normal conditions. On the other hand, this period can offer us a greater chance of detecting the effects of

²³The updated series of the shadow rate are available at: <https://sites.google.com/view/jingcynthiawu/shadow-rates> (accessed on July 21, 2021).

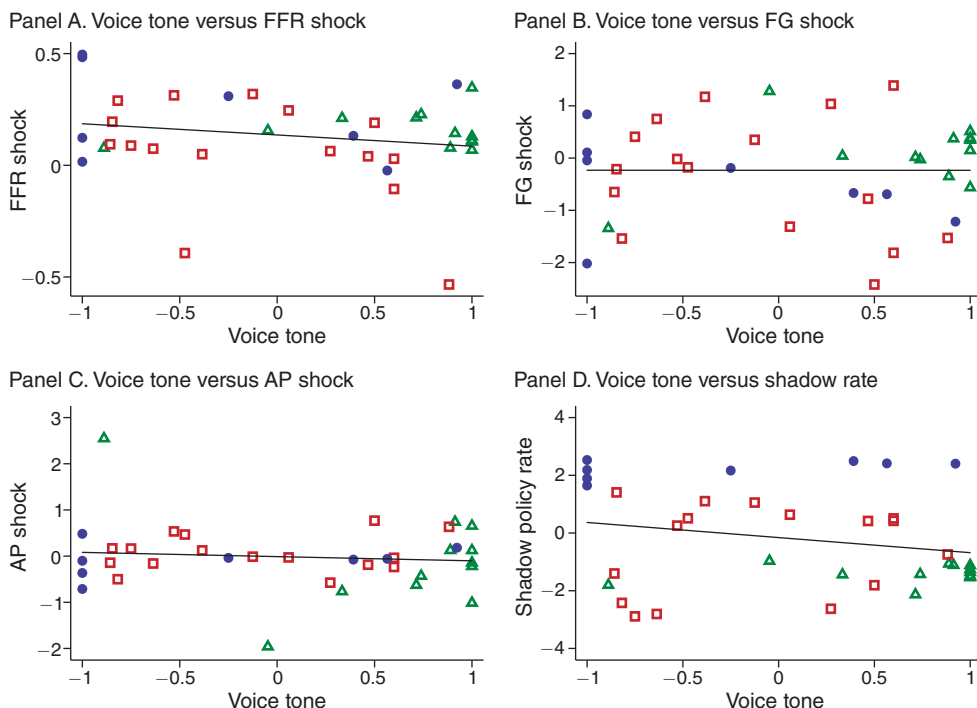


FIGURE 2. POLICY WORDS VERSUS ACTIONS

Notes: This figure shows the joint distribution of voice tone and policy actions/stance. Federal funds rate (FFR), forward guidance (FG), and asset purchase (AP) shocks are from Swanson (2021). The shadow policy rate is from Wu and Xia (2016).

policy communication because it is less clouded by potentially confounding factors such as changes in the Fed funds rate.²⁴

III. Empirical Analysis

In this section, we investigate whether voice tone can move various financial indicators. In particular, we estimate the following specification in the spirit of Jordà (2005):

$$\begin{aligned}
 (4) \quad Outcome_{t,t+h} = & b_0^{(h)} + b_1^{(h)} VoiceTone_t + b_2^{(h)} TextSentiment_t \\
 & + b_3^{(h)} FFRShock_t + b_4^{(h)} FGShock_t + b_5^{(h)} APShock_t \\
 & + b_6^{(h)} ShadowRate_t + b_7^{(h)} \mathbf{1}\{NoPressConference_t\} \\
 & + error_t^{(h)},
 \end{aligned}$$

²⁴ We do not make any normative statements about whether the observed variation in voice was helpful or not because we do not have a complete picture about the objectives of the Fed chairs in their press conferences. In other words, we do not know if a certain variation in voice was intentional or made in error. If the latter is the case, one may be concerned that voice control was inadequate, and that unnecessary volatility has been introduced.

where t dates FOMC meetings; $VoiceTone_t$ measures the voice tone of the Q&A session at date t ; $TextSentiment_t$ indicates the sentiment in the policy statement, remarks, and Q&A responses; and $FFRShock_t$, $FGShock_t$, and $APShock_t$ are policy shocks identified using intraday data with a three-factor model by Swanson (2021). These policy shocks are normalized to have unit variance over a “typical” period (e.g., the FFR shock is normalized to have unit variance for the period that excludes zero lower bound). The variable $ShadowRate_t$ is the shadow policy rate from Wu and Xia (2016). Policy shocks and the shadow rate control for “actions” of the Fed so that we can more cleanly identify the effects of voice tone on outcome variables. Note that we code $VoiceTone$ as equal to zero for FOMC meetings without Q&A sessions but our results are robust to focusing only on meetings with press conferences. We include the indicator variable $\mathbf{1}\{NoPressConference_t\}$ which is equal to 1 when an FOMC meeting did not have a press conference.

We estimate specification (4) for each horizon $h = [0; 15]$ separately and plot the estimated coefficients, e.g., $\{\hat{b}_1^{(h)}\}_{h=0}^{15}$, to illustrate the dynamics of the response to a form of policy action or communication. Note that while high-frequency analyses tend to find clear responses to policy announcements at the intraday frequencies (e.g., Kuttner 2001; Swanson 2021), we use the daily frequency which, given the dramatic volatility of some financial indicators, often yields statistically insignificant estimates (see, e.g., Gorodnichenko and Weber 2016). However, one could expect that the response may build over time, consistent with the notion of “slow-moving” capital proposed by Duffie (2010) and Fleckenstein, Longstaff, and Lustig (2014). Using daily series allows us to examine responses at longer horizons, which may be important for identifying policy actions and communication tools with durable effects. We will use estimate responses to policy shocks ($\{\hat{b}_3^{(h)}\}_{h=0}^{15}$, $\{\hat{b}_4^{(h)}\}_{h=0}^{15}$, $\{\hat{b}_5^{(h)}\}_{h=0}^{15}$) and text sentiment ($\{\hat{b}_2^{(h)}\}_{h=0}^{15}$) to benchmark responses to voice tone variations.

The outcome variables are daily financial indicators available from Thomson Reuters and other popular sources, including Yahoo Finance and Tiingo. We generally use prices of exchange-traded funds (ETFs) that track popular indices. For example, we use the SPY, an ETF that tracks the S&P 500 index, to measure the reactions of the stock market to policy shocks. We measure returns on ETF funds or similar securities as the log close price at date $t + h$ minus log open price at date t , e.g., $Outcome_{t,t+h}^{SPY} = \log(SPY_{t+h}^{close}) - \log(SPY_t^{open})$. Hence, the return on the day of an FOMC meeting is the log difference between close and open prices.

As the sample is relatively small (68 FOMC meetings), we estimate specification (4) using nonparametric (accelerated) bootstrap methods to correct for possible biases in the estimates, as well as to construct confidence intervals with good coverage. Specifically, 90 percent bias-corrected confidence intervals are reported. As a robustness check, we estimate specification (4) with $VoiceTone$ as the only regressor. In further robustness checks, we will also explore the sensitivity of the estimates to including additional controls and other variations in the specification.

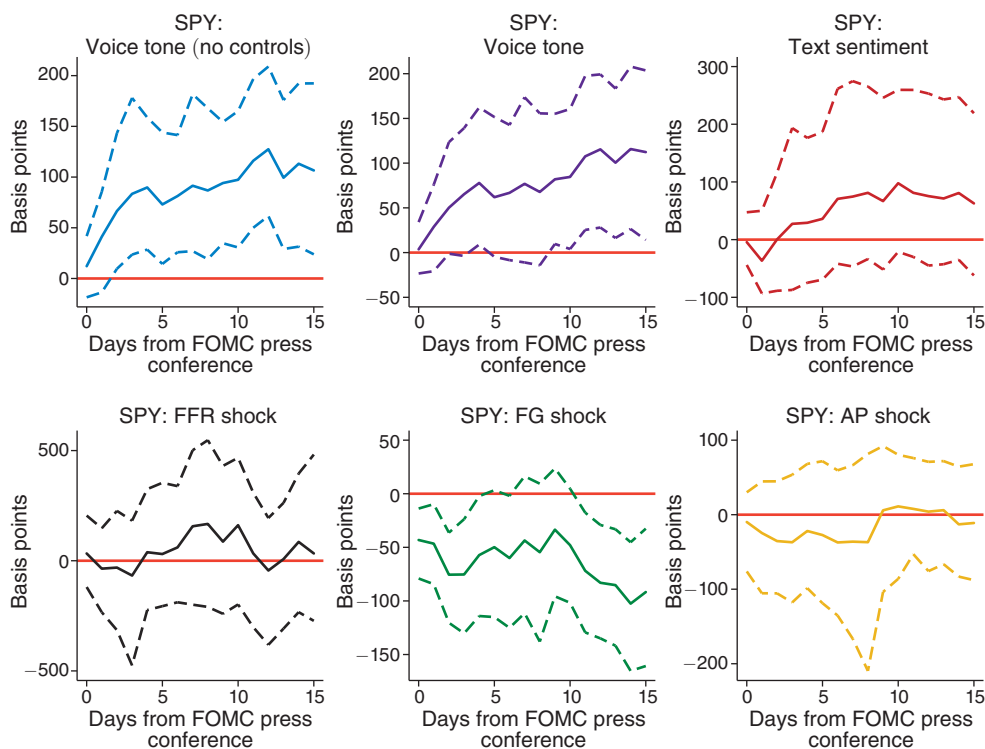


FIGURE 3. RESPONSE OF SPY ETF (S&P 500) TO POLICY ACTIONS AND MESSAGES

Notes: This figure reports the estimated slope coefficients b (specification (4)) for policy communication/actions. Dashed lines show 90 percent bias-corrected and accelerated bootstrap confidence intervals.

A. Stock Market Reactions

When we use the SPY to measure the reactions of the stock market to policy actions and communications, we find that a more positive voice tone leads to an increase in share prices (panel B of Figure 3). Specifically, the impact response (i.e., $h = 0$) of the stock market is weak and not statistically significant. Over time, the response builds up and after five days, the return on SPY reaches approximately 100 basis points for a unit increase in voice tone. The response levels off after the first few days and stays statistically significant at 10 percent. We observe this pattern irrespective of whether we include controls (panel B) or not (panel A) in specification (4).

The sentiment of the policy texts does not appear to have a statistically significant effect on the SPY in our sample, although the point estimates are positive, suggesting that a more dovish sentiment leads to a boom in the stock market. This finding is qualitatively in line with the results documented in the literature. For example, employing the high-frequency event study approach, Rosa (2011b) shows that surprise hawkish FOMC statements lead to a reduction in equity returns. However, using monthly data over the 1998–2014 period, Hansen and McMahon (2016) find a statistically insignificant reaction of stock markets to FOMC statements that focus on strong economic conditions.

The FFR shock does not have a statistically significant effect on the stock market, which likely reflects the fact that the sample period is dominated by the zero lower bound and that changes in the short-term policy rate may have provided a relatively limited outlook for the stance of monetary policy. Changes in the pace of asset purchases by the central bank (AP shock) also do not have a clear effect on the stock market, a finding consistent with Swanson (2021). Note that our sample does not include the first round of quantitative easing in 2009, which led to a strong stock market reaction (see, e.g., Chodorow-Reich 2014; Krishnamurthy and Vissing-Jorgensen 2011). For this sample period, however, an FG shock leads to a persistent decrease in stock prices, in line with the intraday responses estimated by Swanson (2021). This response is consistent with the signaling effect suggested by Campbell et al. (2012): an FG shock reveals that the Fed could be pessimistic about the state of the economy. The magnitude of the stock market response to a unit decrease in voice tone is approximately equal to the response we observe after a one standard deviation FG shock. When we use the Shapley decomposition of the R^2 , we find that the absolute contribution of voice tone to R^2 is around 20 percent, which is slightly larger than that of the text sentiment indicator and is similar to the contribution of the FG shock.²⁵ Thus, the variation in voice tone has economically significant effects.

To understand the reaction of the stock market to policy actions and messages, we examine the response of the Chicago Board Options Exchange Volatility Index (CBOE VIX) (Figure 4), a popular measure of the stock market's expectations about future volatility. We also study the responses of futures on the VIX to provide us with a more refined sense of how policy can influence the outlook for volatility. Specifically, we use VIXY (Figure 5; VIX Short-Term Futures) and VIXM (online Appendix Figure A7; VIX Mid-term Futures) ETFs. We find that Fed actions (FFR/FG/AP shocks) tend to raise the volatility in the stock market. Consistent with this result, Ehrmann et al. (2019) document the greater responsiveness of treasury yields to macro news during weak forms of FG which can be interpreted as evidence of greater uncertainty in relation to rate paths. Both a more positive tone of voice and a more dovish text sentiment could lead to a decrease in current and anticipated volatility. This result is in line with the notion that central bank communication can shape uncertainty about future economic conditions (Hansen, McMahon, and Tong 2019). The variation in voice tone has economically significant effects: a unit decrease in the tone increases the volatility by an amount that is roughly equal to the increase after a one standard deviation shock to FG.

Relatedly, monetary policy can convey information about the path of interest rates and thus reduce the interest rate risk (Hattori, Schrimpf, and Sushko 2016). To quantify the importance of this channel, we measure the interest rate risk with the following spread: $Outcome_{t,t+h} = \log\left(\frac{P_{t+h,close}^{LQD}}{P_{t,open}^{LQD}}\right) - \log\left(\frac{P_{t+h,close}^{LQDH}}{P_{t,open}^{LQDH}}\right)$, where P^{LQD} is the price of LQD (investment grade corporate bonds) and P^{LQDH} is the price of interest rate hedged corporate bond LQDH. A decrease in this measure indicates a decline in the perceived interest rate risk. Our results (Figure 6) suggest that a

²⁵ Online Appendix Figure A9 plots the contribution by horizon.

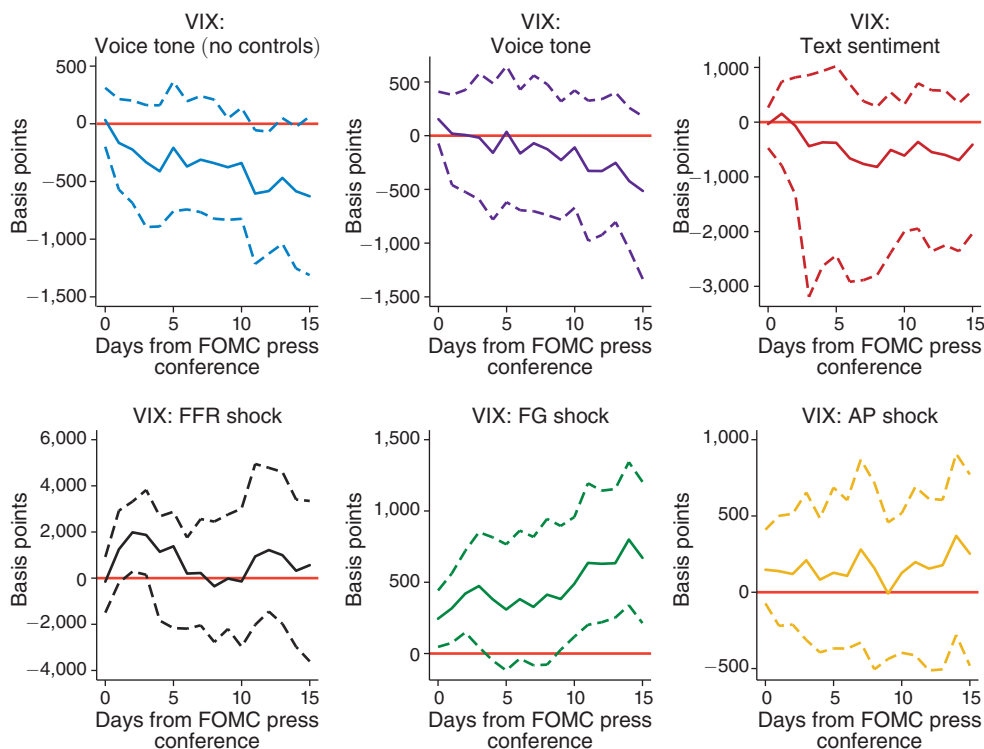


FIGURE 4. RESPONSE OF VIX (CBOE VOLATILITY INDEX) TO POLICY ACTIONS AND MESSAGES

Notes: This figure reports the estimated slope coefficients b (specification (4)) for policy communication/actions. Dashed lines show 90 percent bias-corrected and accelerated bootstrap confidence intervals.

more positive tone leads to a reduction in investor expectations about interest rate risk. Consistent with this interpretation, and in line with the existing studies which document the impact of policy actions on bond risk premia (e.g., Hattori, Schrimpf, and Sushko 2016), we find that an FG shock reduces uncertainty about the future path of interest rates. A unit decrease in voice tone and a one standard deviation increase in the FG shock generate similar responses of the spread, again pointing to the economic significance of the tone of voice. In contrast, an AP shock could signal a lower amount of interest rate risk in investor portfolios in the future, and thus, increase the perceived current interest rate risk, which is consistent with the analysis in Gorodnichenko and Ray (2017). Intuitively, asset purchases are a form of discretionary policy and the deployment of such a tool increases uncertainty about the future path of policy.

B. Bond Market Reactions

Kuttner (2001), Swanson (2021), and a number of others document a strong reaction of the bond market to monetary policy shocks. Consistent with these earlier works, we find (Figure 7) that the price of GOVT (a fund covering US government nominal debt) decreases in response to a forward guidance shock (i.e., yields

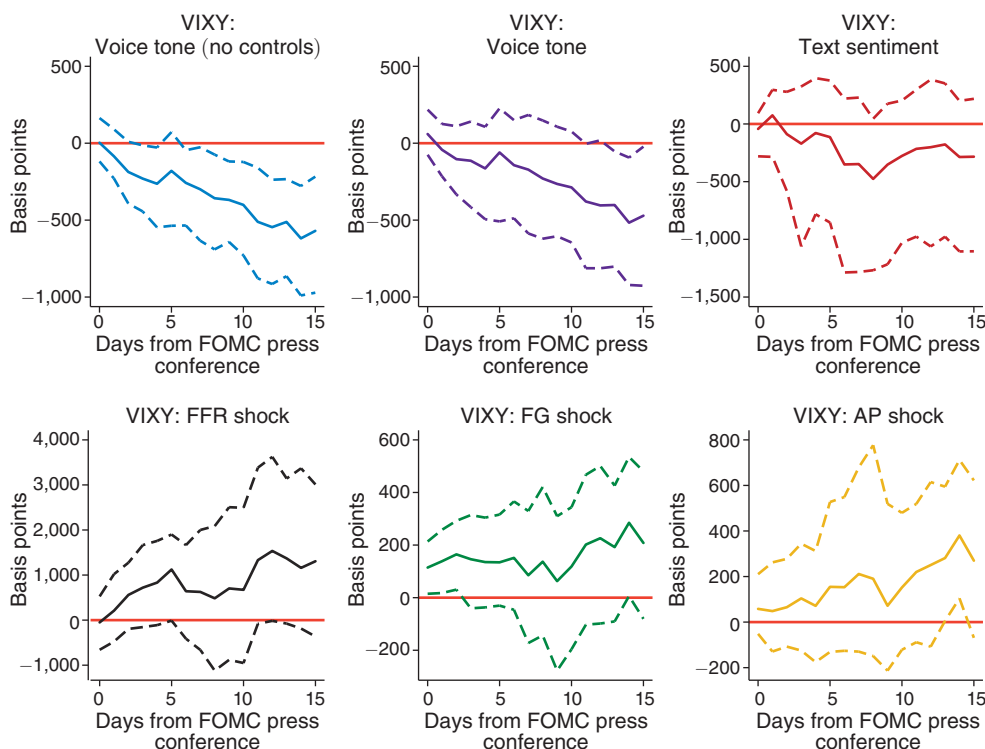


FIGURE 5. RESPONSE OF VIXY ETF (VIX SHORT-TERM FUTURES) TO POLICY ACTIONS AND MESSAGES

Notes: This figure reports the estimated slope coefficients b (specification (4)) for policy communication/actions. Dashed lines show 90 percent bias-corrected and accelerated bootstrap confidence intervals.

rise) and increases in response to an asset-purchase shock (i.e., yields fall). FFR shocks do not lead to a statistically significant response in GOVT prices, which likely reflects the prominence of the zero lower bound in our sample. In contrast to the strong responses of the stock market, the responses of the bond market to voice tone are not statistically significant (although we later document that voice tone can move spreads between nominal and real bonds). Similarly, text sentiment does not move GOVT prices materially. These findings are consistent with Cieslak and Pang (2020) and Ehrmann and Talmi (2020), who document that the bond market reaction to Fed communications is weak. Using ETFs for government debt with different maturities, we also examine if there could be a differential response across the maturity space. We find qualitatively similar responses for all maturities²⁶, although the magnitudes of the responses to FFR/FG/AP shocks tend to be smaller for shorter maturities. While the responses of the bond market appear to be somewhat decoupled from the responses of the stock market, differentiated responses have been documented in the previous literature. For example, Lucca and Moench (2015) find that there is a pre-FOMC announcement drift in the stock market but a similar effect is not found for US Treasuries.

²⁶See online Appendix Figures A1–A6.

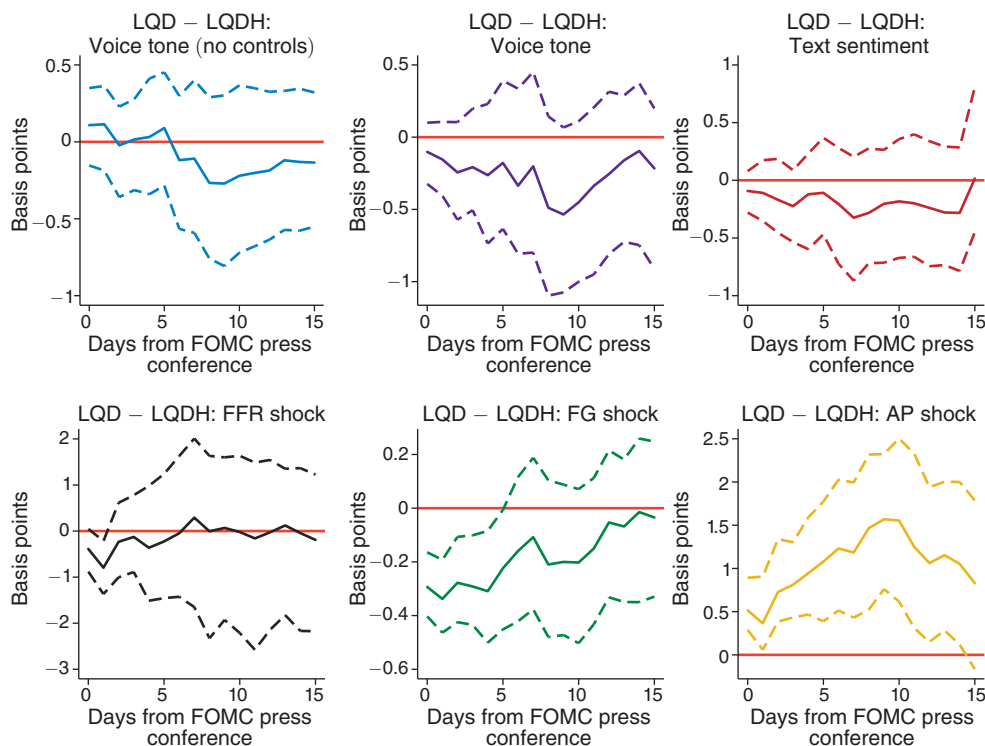


FIGURE 6. RESPONSE OF LQD (INVESTMENT GRADE CORPORATE BOND) MINUS LQDH (INTEREST RATE HEDGED CORPORATE BOND) TO POLICY

Notes: This figure reports the estimated slope coefficients b (specification (4)) for policy communication/actions. Dashed lines show 90 percent bias-corrected and accelerated bootstrap confidence intervals.

An important dimension of monetary policy transmission is how policy can influence the interest rates faced by the corporate sector. While the bond market is highly integrated, the pass-through from US government debt to corporate debt may be limited and nuanced. In our first attempt to address this question, we use the LQD (a fund covering investment grade corporate bonds) and find that policy actions (FFR/FG/AP shocks) tend to move yields in the same direction as they move yields for US government debt (Figure 8). Text sentiment does not have a statistically significant effect on LQD prices. A positive voice tone appears to elevate LQD prices (i.e., yields fall) for a few days after an FOMC meeting, but this effect is short-lived and statistically insignificant. The results are broadly similar when we use the IVR (real estate investment trust; Figure 9) to gauge the responses of the real estate sector.

C. Inflation Expectations

Management of inflation expectations is a key element of monetary policy (see Coibion et al. 2020 for a survey). To evaluate the success of policymakers in this matter, we use two popular metrics. The first one is the spread between nominal and inflation-protected US government bonds. Specifically, we use $Outcome_{t,t+h} =$

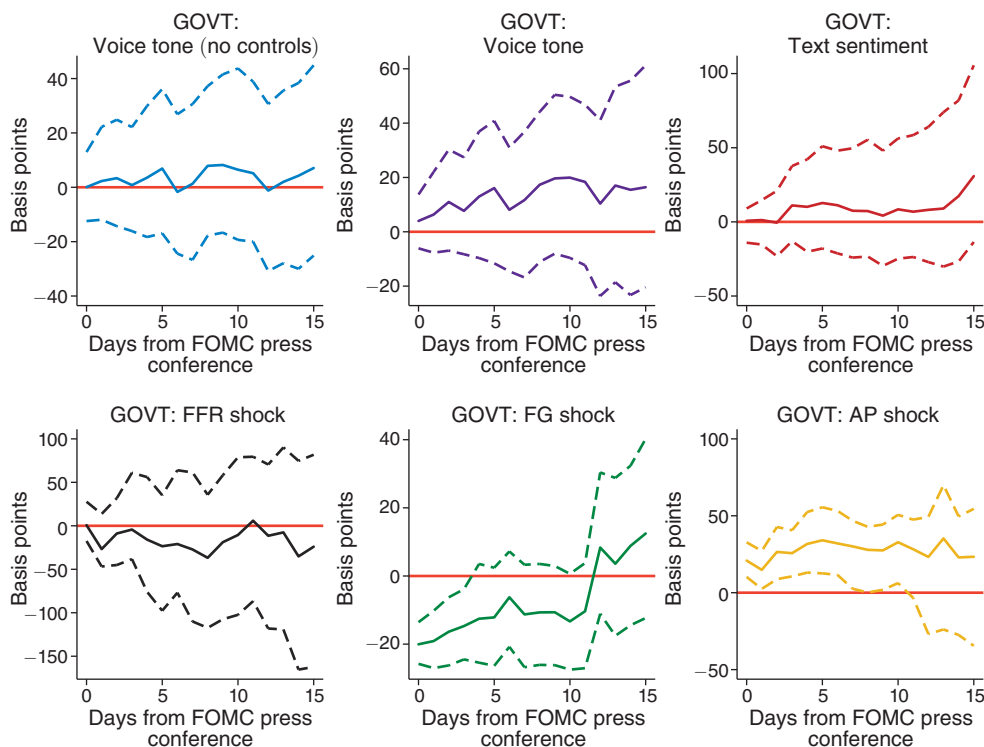


FIGURE 7. RESPONSE OF GOVT (US GOVERNMENT DEBT) TO POLICY ACTIONS AND MESSAGES

Notes: This figure reports the estimated slope coefficients b (specification (4)) for policy communication/actions. Dashed lines show 90 percent bias-corrected and accelerated bootstrap confidence intervals.

$\log\left(\frac{P_{t+h,close}^{GOVT}}{P_{t,open}^{GOVT}}\right) - \log\left(\frac{P_{t+h,close}^{TIP}}{P_{t,open}^{TIP}}\right)$ as a measure of the spread, where P^{GOVT} is the price of GOVT (nominal bonds) and P^{TIP} is the price of TIP (inflation-protected bonds). An increase in this spread can be interpreted as a decrease in expected inflation. The second is the GLD, a fund that tracks the gold spot price. This ETF is used as a hedge against inflation: an increase in the price of GLD signals higher expected inflation. Although neither of these measures is perfect (e.g., the spread varies not only due to changes in inflation expectations but also with changes in liquidity conditions; gold prices can move for reasons unrelated to inflation), these two measures are consistently available and are based on reasonably deep markets.

We find that the responses of the GOVT–TIP spread (Figure 10) and GLD (Figure 11) paint a similar picture. As before, the FFR shock does not have a clear impact. The FG shock lowers inflation expectations, while the AP shock raises inflation expectations. More dovish text sentiment appears to raise inflation expectations (the GLD price increases), but this response is not statistically significant. Moreover, it does not seem to have support from the GOVT–TIP spread, which appears to increase (i.e., expected inflation is lower) and the effect is statistically significant. The impact response of the GOVT–TIP spread to a positive tone of voice is close to zero, but the spread gradually increases (thus, signaling lower expected inflation)

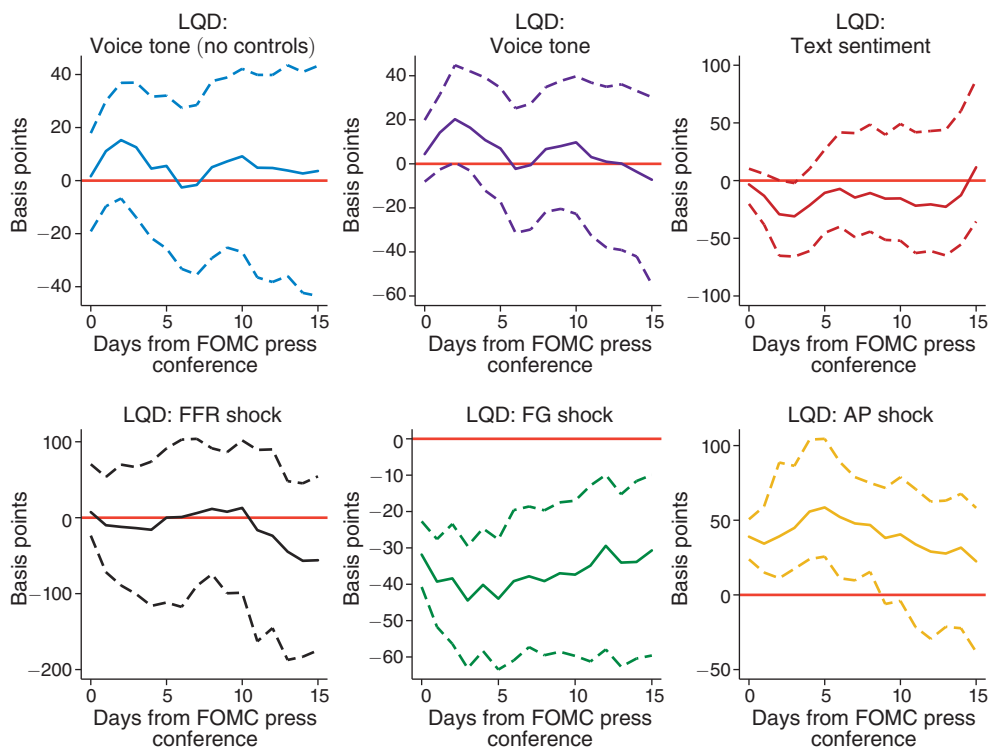


FIGURE 8. RESPONSE OF LQD (CORPORATE DEBT) TO POLICY ACTIONS AND MESSAGES

Notes: This figure reports the estimated slope coefficients b (specification (4)) for policy communication/actions. Dashed lines show 90 percent bias-corrected and accelerated bootstrap confidence intervals.

and peaks after about ten days. The GLD price has similar dynamics (i.e., lower expected inflation), but the estimates are less precise. Hence, voice tone seems to have an independent effect on inflation expectations. One may conjecture that a positive tone plays a signaling role: a happy tone of the Fed chair indicates satisfaction with future inflation dynamics.

D. Exchange Rate

The exchange rate is an important channel for monetary policy transmission in the increasingly globalized economy. To shed further light on how policy actions and communication can work via this channel, we examine the responses of two key exchange rates: dollar/yen (JPY; Figure 12) and dollar/euro (EUR; Figure 13).²⁷ We find that policy actions generally lead to mixed reactions across currencies in our sample. For example, after an FFR shock (monetary tightening), the dollar depreciates against the euro (although the effect is not statistically significant) and appreciates against the yen (for the first five days after the FOMC meetings).²⁸

²⁷ We also report results for the pound/dollar (GBP) exchange rate in online Appendix Figure 8.

²⁸ Analyses using intraday data (e.g., Swanson 2021) find that the dollar appreciates against the euro and yen after a FFR shock.

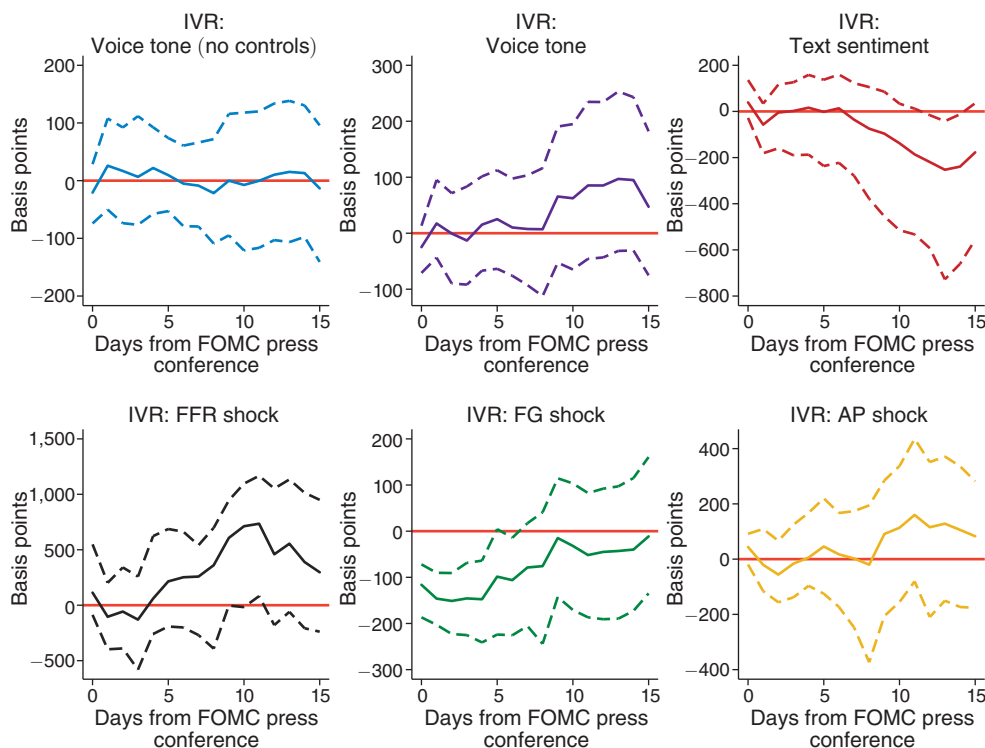


FIGURE 9. RESPONSE OF IVR (DEBT FOR THE REAL ESTATE SECTOR) TO POLICY ACTIONS AND MESSAGES

Notes: This figure reports the estimated slope coefficients b (specification (4)) for policy communication/actions. Dashed lines show 90 percent bias-corrected and accelerated bootstrap confidence intervals.

After a more dovish text sentiment (the opposite of monetary tightening), the dollar appreciates against the euro while the response of the dollar/yen exchange is close to zero. Similarly, a more positive tone of voice leads to an appreciation of the dollar against the euro, but there is neither a statistically nor economically significant response for the dollar/yen exchange rate. While somewhat unexpected, the relatively lower level of reaction of the dollar/yen exchange rate to the monetary policy shocks has also been observed in other studies (e.g., Fatum and Scholnick 2008; Rosa 2011a).

E. Robustness Checks

To assess the sensitivity of our findings to additional measurements and assumptions, we perform a series of robustness checks that may be grouped into two categories. To isolate the effect of policy shocks on financial and macroeconomic variables, it is important to control for the information set available to economic agents at the time when policy changes are announced. To this end, the first category explores the robustness of our results to additional controls or alternative specifications. As we have a small sample size, we typically include one additional control at a time. Given the importance of verbal communication, the second category focuses on alternative measures of sentiment in policy texts. To save

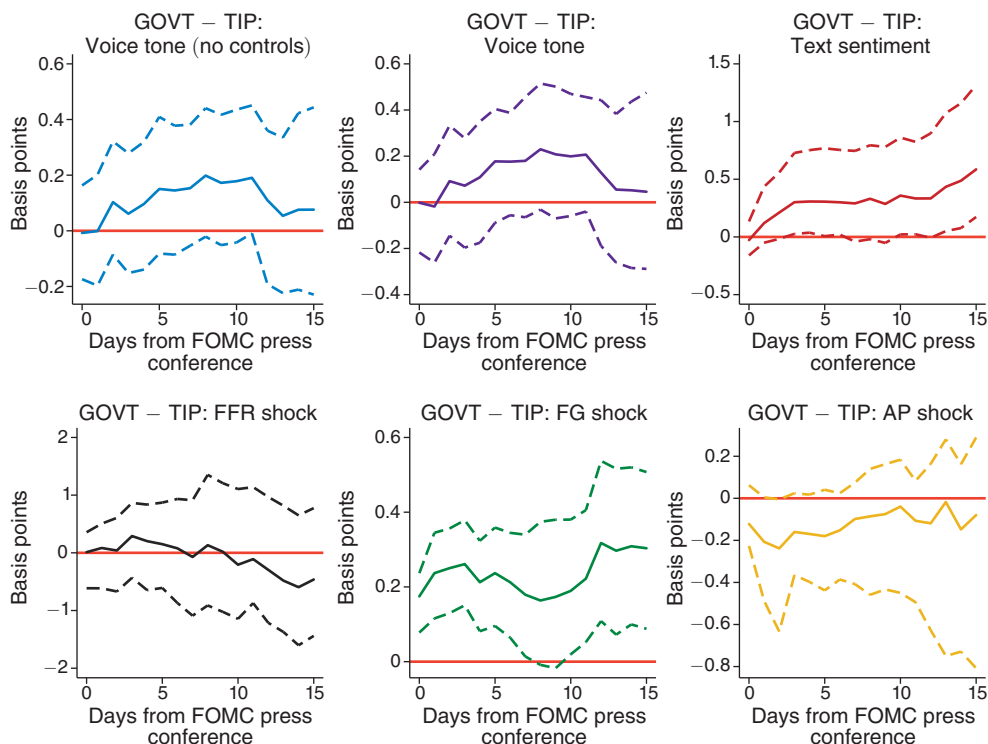


FIGURE 10. RESPONSE OF GOVT (NOMINAL US GOVERNMENT DEBT) MINUS TIP EFT (INFLATION-PROTECTED US GOVERNMENT DEBT) TO POLICY ACTIONS AND MESSAGES

Notes: This figure reports the estimated slope coefficients b (specification (4)) for policy communication/actions. Dashed lines show 90 percent bias-corrected and accelerated bootstrap confidence intervals.

space, we present results for the stock market responses, but we reach similar conclusions for other variables.

Additional Controls.—While most of the FOMC announcements did not overlap with the releases of other macrodata over the examined period, there are certain exceptions. In particular, five FOMC announcements (none of which was accompanied by a press conference) were released on the day when GDP data were unveiled. In addition, seven FOMC announcements (six of which were followed by a press conference) were made on the same day as consumer price index data releases. As macroeconomic news can move financial markets (Gürkaynak, Sack, and Swanson 2005b), we introduce an additional control in specification (4), the Citigroup Economic Surprise Index, which aggregates macroeconomic surprises in released data into one indicator. We find that voice tone continues to move the stock market (panel A of Figure 14). We also do not observe important changes in the estimates when we control for the volume of corporate earnings data releases (panel B, Figure 14).²⁹

²⁹ Data on the number of corporate earnings announcements of the US listed firms during the 2011–2019 period were scrapped from Yahoo Finance’s earnings calendar (<https://finance.yahoo.com/calendar/earnings>).

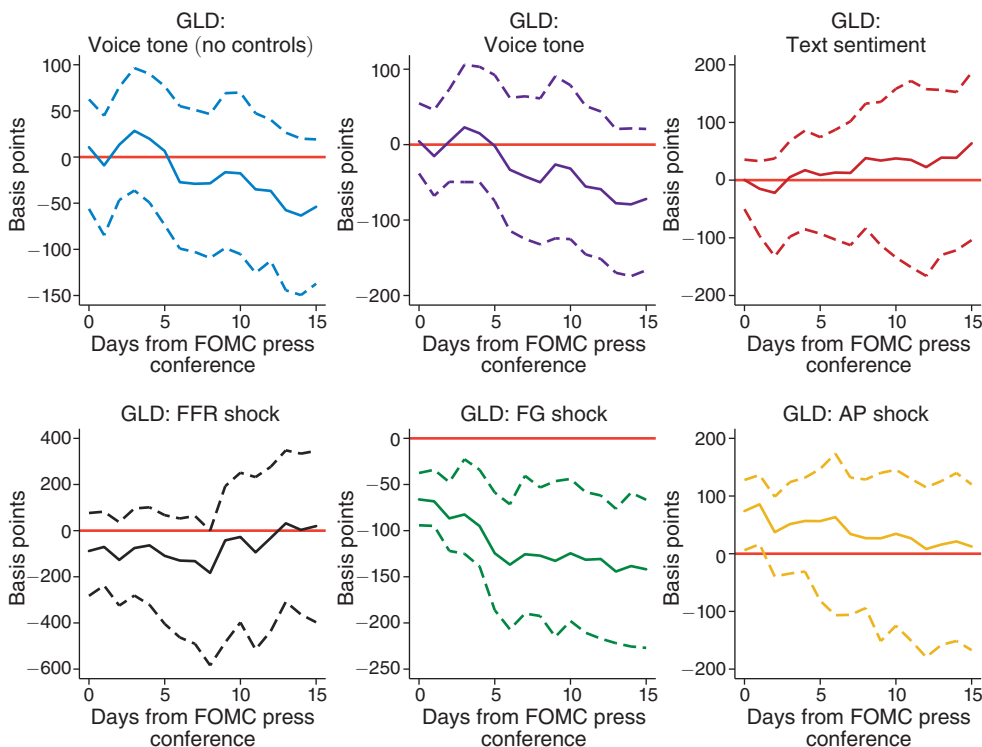


FIGURE 11. RESPONSE OF GLD (GOLD) TO POLICY ACTIONS AND MESSAGES

Notes: This figure reports the estimated slope coefficients b (specification (4)) for policy communication/actions. Dashed lines show 90 percent bias-corrected and accelerated bootstrap confidence intervals.

One may be concerned that important variations in voice tone coincide with the media cycles that can affect the stock market. To address this potential confounder, we collect FOMC-related news coverage from Nexis Uni, a popular news database. We search for all news items which contain “FOMC” and one of the following keywords: “interest rate,” “monetary,” “federal funds rate,” or “fed funds rate.” The search results are restricted to English news published in the United States. We further exclude government sources that simply announce the events and the press conference transcripts such as press releases. After screening and cleaning, the final dataset consists of 23,275 news articles covering the January 1, 2011 to July 15, 2019 period. The news dataset provides us with information on the publication date of the news articles, as well as their content. We then use our BERT-based sentiment algorithm to process the news and construct a measure of media sentiment about monetary policy for each day t :

$$(5) \quad \text{MediaSentiment}_t = \frac{\text{Dovish articles}_t - \text{Hawkish articles}_t}{\text{Articles}_t}.$$

To smooth out noise, we compute five-day averages of *MediaSentiment* before each FOMC meeting. Including this control does not materially change our results (panel C, Figure 14).

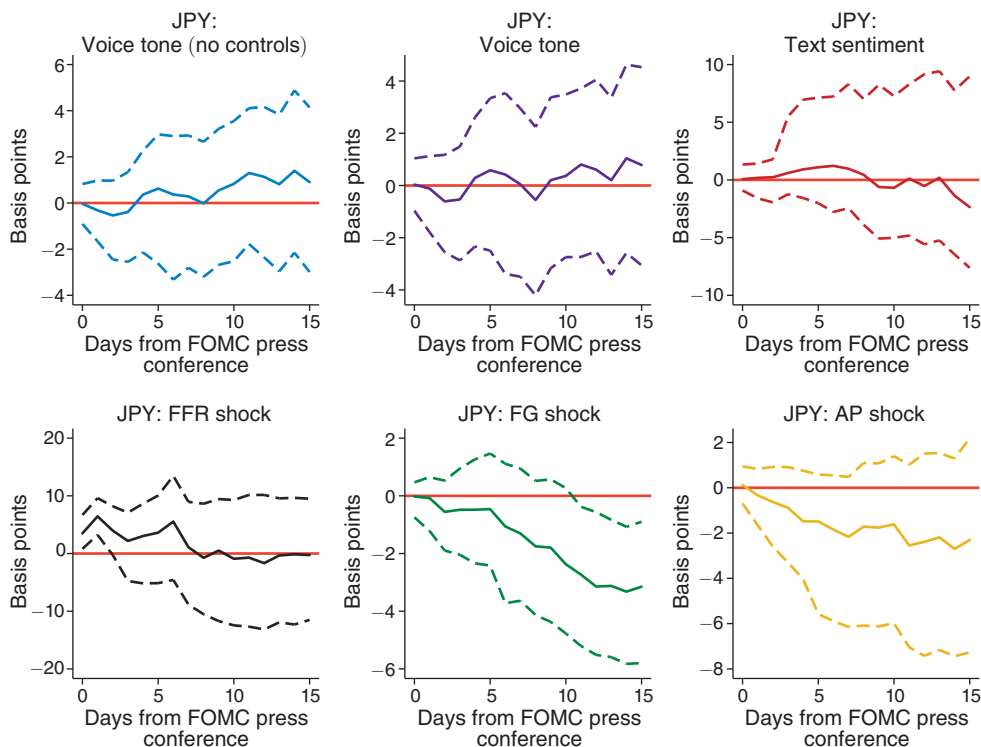


FIGURE 12. RESPONSE OF THE JAPANESE YEN TO ONE US DOLLAR (DOLLAR/YEN) EXCHANGE RATE TO POLICY ACTIONS AND MESSAGES

Notes: This figure reports the estimated slope coefficients b (specification (4)) for policy communication/actions. Dashed lines show 90 percent bias-corrected and accelerated bootstrap confidence intervals.

As we indicated above, Fed chairs vary in their voice tones. As their tenures coincided with different phases of the policy cycle (recall that we use the shadow policy rate to control for the phase of the policy cycle), one may be concerned that the phases may confound the voice tone variation. To address this potential issue, we use fixed effects for the Fed chairs. Although these fixed effects are very demanding on the data given our small sample size, we find that using within-chair variation in voice tone continues to result in voice tone moving the market (panel D). Relatedly, given the negative media bias in covering monetary policy (Berger, Ehrmann, and Fratzscher 2011), the financial market responses may be driven by negative news, which may also correlate with the policy cycle (e.g., raising interest rates may be interpreted as bad news for the economy). To assess the quantitative importance of this potential asymmetry, we modify specification (4) to give us separate regressors for positive and negative voice tone measures. While the estimates are noisier, we find that the responses to positive and negative voice tones are similar in the first ten or so days, but the effects then appear to be stronger for negative news (panel E).

Alternative Textual Sentiment.—Although our baseline analysis uses BERT, a highly accurate natural language processing tool, we want to explore if using alter-

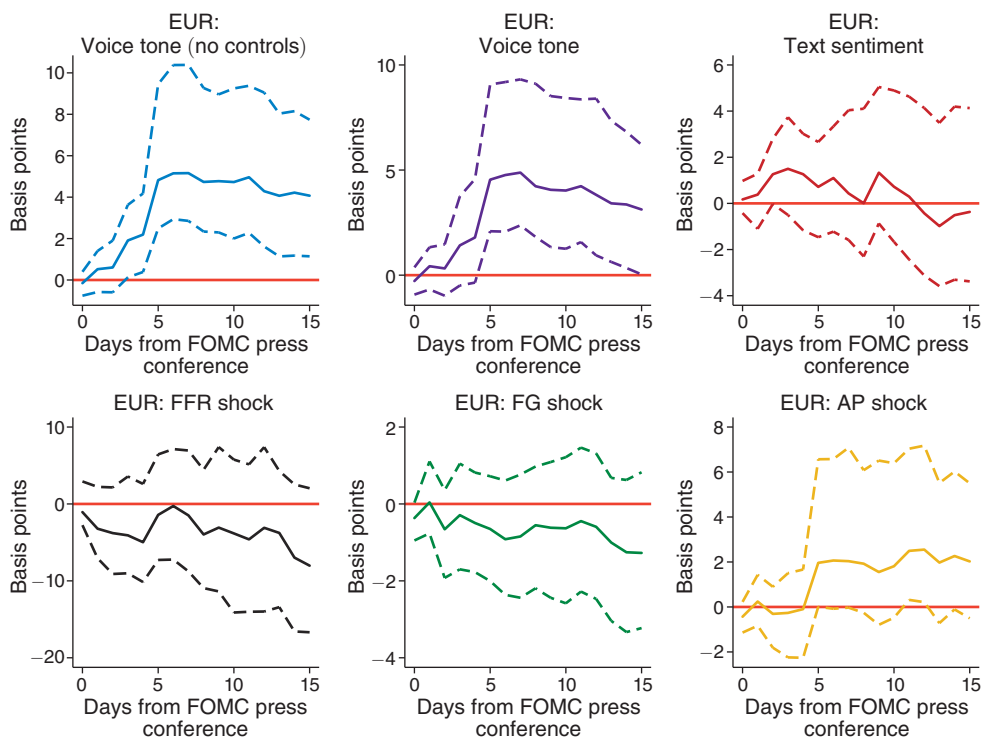


FIGURE 13. RESPONSE OF THE EURO TO ONE US DOLLAR (DOLLAR/EURO) EXCHANGE RATE TO POLICY ACTIONS AND MESSAGES

Notes: This figure reports the estimated slope coefficients b (specification (4)) for policy communication/actions. Dashed lines show 90 percent bias-corrected and accelerated bootstrap confidence intervals.

native tools to quantify the sentiment of policy texts can affect our estimates for responses to voice tone. First, we employ the Robustly Optimized BERT Pretraining Approach (RoBERTa), a modified version of BERT that was developed by Facebook AI (Liu et al. 2019). Compared to the original BERT, RoBERTa is trained on more data, larger batches, longer sequences, and dynamically changes the masking pattern applied to the training data while removing the next sentence prediction objective. At the cost of being significantly more computationally expensive, RoBERTa can yield a modest improvement in accuracy. Second, as mentioned earlier, the pretrained BERT sentiment classification is based on a corpus of training texts that is not tailored for monetary policy. As a result, we have to fine-tune the model to convert BERT's word embeddings into the hawkish/dovish spectrum. As an alternative to our approach, we use FinBERT (Araci 2019), a pretrained BERT model which was fine-tuned for the sentiment analysis of financial texts (rather than general texts).

Third, we use the search-and-count approach which has been widely used in the literature (e.g., Apel and Grimaldi 2014; Neuhierl and Weber 2019). In short, we build lists (dictionaries) of nouns, adjectives, and verbs that can indicate the stance of monetary policy (hawkish or dovish) and the strength of economic outlook. We

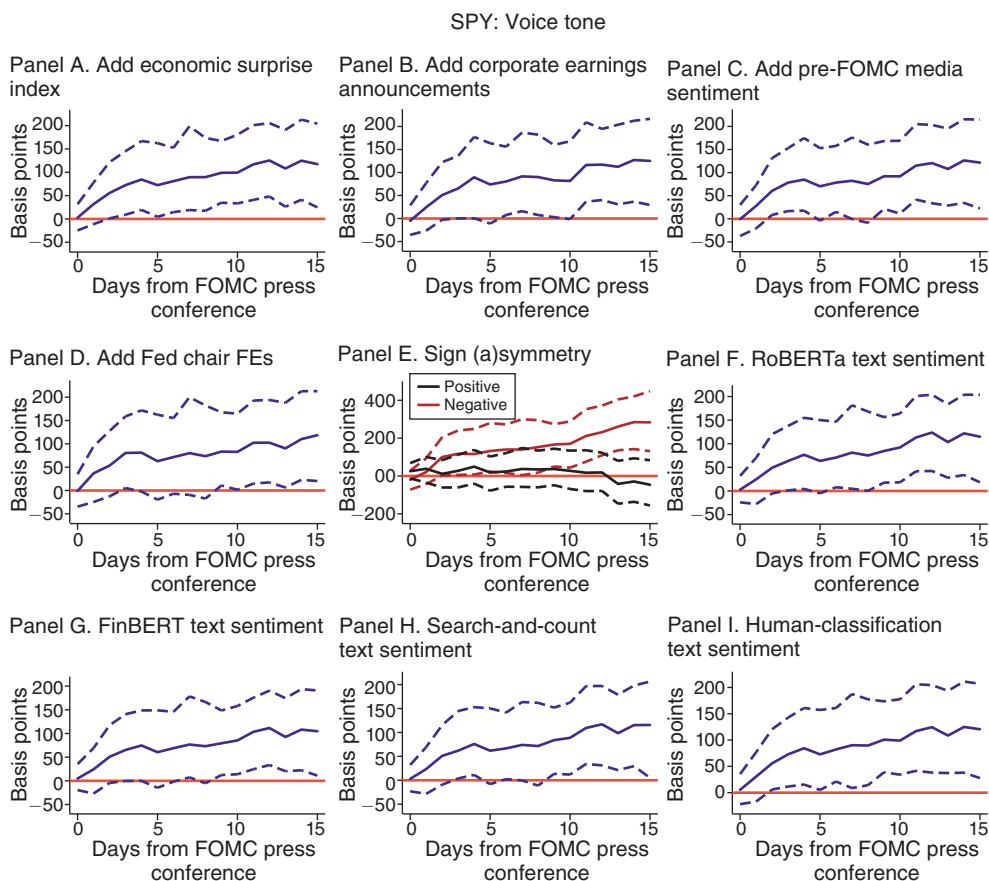


FIGURE 14. ROBUSTNESS CHECKS

Notes: This figure reports the estimated slope coefficients b (specification (4)) for voice tone. The outcome variable is SPY, the ETF that tracks the S&P 500 index. Dashed lines show 90 percent bias-corrected and accelerated bootstrap confidence intervals.

then compute the frequency of words in these dictionaries (online Appendix D1 provides more details). While this approach is less sophisticated than BERT, it is particularly transparent and easy to implement. Fourth, we ask a group of research assistants to score texts in policy statements, remarks, and Q&A sessions for the 2011–2019 sample. The scores vary from -10 (very hawkish) to $+10$ (very dovish). For each FOMC meeting, we compute the average score across texts and research assistants. While this approach may be more accurate in detecting the nuances of human communication, it is more subjective (and hence potentially less reproducible) than BERT.

Panels F–I of Figure 14 show that using alternative measures of text sentiment does not affect our conclusion that voice tone can move the financial markets. In online Appendix D2, we report results for additional robustness checks. In these checks, we apply an approach similar to Kozlowski, Taddy, and Evans (2019) and Jha, Liu, and Manela (2021) to measure the *intensity* of dovishness/hawkishness of

the policy texts and we experiment with allowing nonlinear terms in text sentiment. In short, we find similar results.

One could also argue that our measure of voice tone simply captures some features of the policy texts which were not fully accounted for in our text sentiment measure. While we cannot rule out this alternative explanation completely, we note that the results on the tone of voice are robust to different measures of text sentiment, ranging from human classification to state-of-the-art methods in machine learning. This robustness suggests that the voice tone measure could capture additional information which goes significantly further than the message captured by the text sentiment. As a result, voice tone and other forms of nonverbal communication can expand the toolkit for managing the expectations of the public.

High-Frequency Analysis.—Our results suggest that the effect of voice tone builds up gradually with only relatively small responses observed on impact at the daily frequency. We use a number of regressors to control for possible confounders (e.g., a piece of macroeconomic news is released on the day of a press conference). In this section, we use intraday data to zoom in on the high-frequency movements of asset prices. While this approach can sharpen our identification, it also likely amplifies noise as voice tone is measured with error, and financial markets can take time to process nonverbal cues from the Fed (recall that voice tone is a flow rather than a stock).

With this caveat in mind, we generate the precise timing (down to a second) for each answer during press conferences. We then match SPY prices to the timings. Finally, we estimate the following specification:

$$(6') \quad Outcome_{t \rightarrow t+h, m, s} = b_0^{(h)} + b_1^{(h)} VoiceTone_{t, m, s} + b_2^{(h)} TextSentiment_{t, m, s} \\ + \lambda_m + \gamma_s + error_t^{(h)},$$

where t is the start time of the answer to the s th question in meeting m , $t + h$ is h minutes after the end of the answer, λ_m is the fixed effect for meeting m , and γ_s is the fixed effect for the order of questions (i.e., question number s in a press conference). The variable *VoiceTone* takes values -1 (negative), 0 (neutral), and 1 (positive); *TextSentiment* takes values -1 (hawkish), 0 (neutral), and 1 (dovish); and $Outcome_{t \rightarrow t+h, m, s}$ is the price change for SPY between t and $t + h$.

We also estimate a modified version of specification (6'):

$$(6'') \quad Outcome_{t_0 \rightarrow t+h, m, s} = b_0^{(h)} + b_1^{(h)} VoiceTone_{t_0 \rightarrow t, m, s} \\ + b_2^{(h)} TextSentiment_{t_0 \rightarrow t, m, s} + \lambda_m + \gamma_s + error_t^{(h)},$$

where t_0 is the start of the answer for the first question in a press conference, $VoiceTone_{t_0 \rightarrow t, m, s}$ measures the cumulative tone of answers (calculated as in equation (2)) between t_0 and t , and $TextSentiment_{t_0 \rightarrow t, m, s}$ measures the cumulative text sentiment of answers (calculated as in equation (3)) between t_0 and t . As there is

clear dependency in the variables by construction, we use meetings as clusters for bootstrap.

Specification (6'), which we call the "flow" specification, examines the reaction of a financial variable answer by answer. In contrast, the "cumulative" specification (6'') focuses on the response of a financial variable to information conveyed since the start of a press conference. The main advantage of the "flow" specification is that the unit of analysis is an answer. However, each answer may be a noisy measure of policy stance and, hence, one may also want to use the "cumulative" specification which likely attenuates the noise. Note that both specifications include meeting fixed effects meaning that we control for a broad range of factors including the macroeconomic environment and the Fed chair's personality.

As shown in Figure 15, we observe a small and statistically significant response of stock prices to voice tone on impact: a positive voice tone raises the stock market by approximately one basis point. For the minute that follows the answer, we cannot reject the null of no response and we cannot reject the null of the stable response (i.e., the impact response is equal to subsequent responses). The "cumulative" specification suggests that, as the voice tone becomes clearer during the course of a press conference, the point estimates gradually increase with the horizon, but the estimates continue to be imprecise. The magnitude of the response is smaller than the magnitude observed in the analysis with daily data. This is consistent with our conjecture that it takes time for the market to interpret signals from the tone of voice.

F. Discussion and Additional Analysis

In general, our findings shed new light on the effectiveness of press conferences as a central bank communication tool. We show that, just as the actions of the Fed move financial markets, so too does the vocal aspect of FOMC press conferences. The vocal dimension of the central bank communication appears to convey information beyond that found in the content of the text, and market participants form their expectations and make their decisions based on that information.

What Is Communicated?—The estimated responses suggest that a more positive voice tone of a Fed chair leads to rate risk reduction, lower expected volatility, depressed inflation expectations, and increased stock prices. The exact information that is communicated is open to further inquiry, but tentative interpretations are possible.

One interpretation is that the tone of voice effectively works as a form of FG. For example, a positive voice tone could signal that the Fed is unlikely to change the policy stance in the near future. If rate risk is attenuated due to a no-change-in-policy signal, then volatility due to policy shocks is diminished, which is reflected by VIX and futures on VIX. Given that our sample is dominated by the zero lower bound, a lower rate risk then means lower inflation expectations (i.e., the Fed does not see a need to raise interest rates to fight inflation). If interest rates are unlikely to increase and the perceived future volatility is lower, the discount factor for future earnings could be lower, thus, pushing up the value of stocks. Furthermore, a decrease in policy uncertainty can have a direct positive effect on the economy (e.g., Husted, Rogers, and Sun 2020). In line with this interpretation, the estimated responses

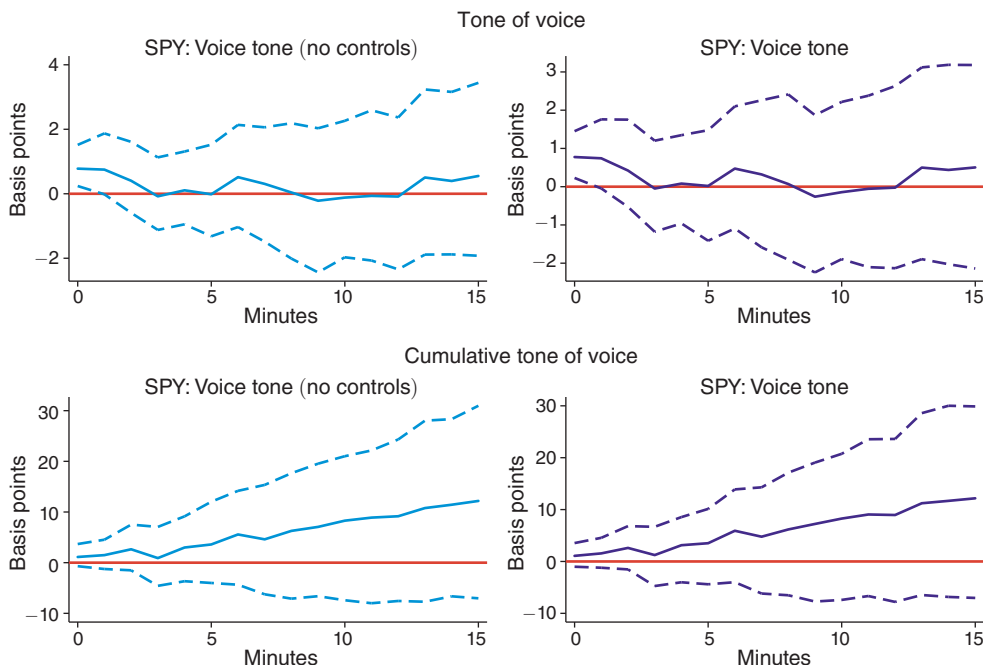


FIGURE 15. INTRADAY RESPONSES OF STOCK PRICES (SPY) TO VOICE TONE

Notes: The top panels show the impulse responses of SPY prices to voice tone variation answer by answer. These impulse responses are estimated using specification (6'). The bottom panels show the impulse responses of changes in SPY prices to cumulative voice tone variation from the start of the Q&A session in a press conference. These impulse responses are estimated using specification (6''). Dashed lines show 90 percent bias-corrected and accelerated bootstrap confidence intervals.

to forward guidance shocks and to voice tone are qualitatively similar for many variables.

Alternatively, these responses reflect some forms of the Fed information effect (Romer and Romer 2000), i.e., the notion that the Fed has superior information about the current or future state of the economy. For instance, the Fed chair may be satisfied with the inflation dynamics and the pace of the economic recovery after the Great Recession. Through a positive voice tone, they can communicate that monetary tightening aimed at fighting inflation is a lower probability event that reduces rate risk and uncertainty. A positive voice tone can also signal a brighter economic outlook which can reduce uncertainty (uncertainty is countercyclical) and raise expectations about future cash flows. These forces could generate a boom in the stock market. So which effect is at play? If the tone of voice was simply a form of FG, one would expect similar effects of the tone of voice and the forward guidance shock across asset classes. The differences in results for the bond market appear to be inconsistent with this interpretation. However, it has been documented that the bond market reactions to Fed communication could be weaker than, or even different from, the stock market reactions (e.g., Lucca and Moench 2015; Cieslak and Pang 2020; Ehrmann and Talmi 2020). Taken together, while there is suggestive

evidence in favor of both the information effect and FG effect, we are not able to provide conclusive evidence on which effect dominates with our data.

Hence, answering the question of which of the accounts above—and there could be other explanations—is a more accurate rationalization of financial market reactions to the nonverbal communication remains a challenging (separating FG and information effects is complex, see e.g., Bauer and Swanson 2020) but fruitful avenue for future research. Notwithstanding this issue, it is clear that the tone of voice can move multiple financial variables, hence, suggesting that the estimated responses are unlikely to be statistical flukes and are likely to capture some systematic forces in the data.

The Shape of the Response.— The difference in the shape of responses to various forms of policy communication and actions is another area for further research. Specifically, in contrast to step-like contemporaneous responses of financial variables to the Fed’s actions, the estimated responses to voice tone (and text sentiment) tend to build gradually over time with weak contemporaneous reactions. We can offer several conjectures to rationalize this pattern of the responses.

First, the tone may be a leading indicator for subsequent policy communication by the Fed chair and other officials. As more information is revealed progressively by the Fed via formal and informal channels, financial variables could take time to respond.³⁰ Indeed, it is not unusual for Fed officials to organize speeches and press conferences aimed at clarifying the position of the Fed after FOMC meetings. Perhaps the most striking example of such follow-up policy communication happened after the “taper tantrum” episode. At the press conference on June 19, 2013, Bernanke hinted at a reduction of the quantitative easing program, which led to significant movements in the financial markets. To contain these gyrations, a number of Fed officials rushed to clear up any potential confusion about the central bank’s intentions.

To assess the plausibility of this channel, we focus on Twitter, a social media platform that central banks have increasingly used to communicate with nonexperts (Ehrmann and Wabitsch 2022). Specifically, we obtain tweets published on the Fed’s Twitter accounts (i.e., the Board of Governors’ account and the regional Fed accounts) and apply the trained BERT model tailored for central bank communication (as used in the main analysis) to these tweets. The aggregate sentiment of the Fed’s tweets on day $h = [1; 15]$ after the press conference is measured as $TweetSentiment_h = \sum_{i=0}^h Sentiment_i$. We then use specification (4) to estimate the response of this tweet sentiment measure to the tone of voice measure. The results reported in panel A of Figure 16 suggest that, after a press conference with a more positive tone of voice, the sentiment of the Fed’s tweets is more dovish. This pattern is consistent with the gradual amplification of policy messages communicated initially via voice tone that are gradually incorporated into asset prices.

Second, the delayed response may capture the dynamics of trading where more-informed/attentive investors (e.g., those who are better at reading—perhaps subconsciously—nonverbal communication at press conferences) move first

³⁰We thank William English for suggesting this channel.

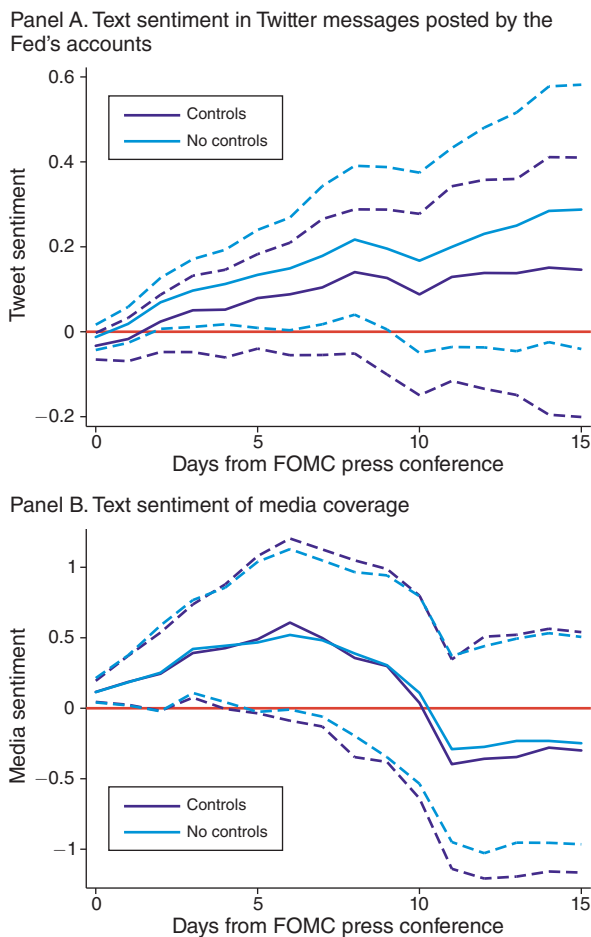


FIGURE 16. POST-FOMC POLICY COMMUNICATION AND MEDIA COVERAGE

Notes: Panel A plots impulse responses for cumulative text sentiment in tweets posted by the Fed's Twitter accounts (to a unit increase in voice tone). The specification is given by equation (4). Panel B plots the impulse responses for cumulative media sentiment to a unit increase in voice tone. The specification is given by equation (4). Dashed lines show 90 percent bias-corrected and accelerated bootstrap confidence intervals.

and the initial momentum is then amplified by less-informed investors, a mechanism explored in other contexts (e.g., Bikhchandani, Hirshleifer, and Welch 1992; Zhou and Lai 2009). Duffie (2010) and others argue that this can give rise to “slow capital” which can generate predictable movements in stock returns. As discussed extensively in Lucca and Moench (2015), the main challenge of explanations based purely on information or financial frictions is how one can rationalize a delayed or otherwise predictable response to one element of policy events (e.g., the predictable pre-FOMC announcement drift in stock returns) and a sharp, largely unforecastable response to another element of policy events (e.g., investors are not systematically surprised by changes in the Fed funds rate). We conjecture that information frictions are exacerbated for policy messages communicated via the tone of voice and other nonverbal channels. Indeed, market participants cannot rely on readily available objective measures of nonverbal communication that should trigger trading

activities. It can take time to filter out that part of policy communication (perhaps by rewatching a press conference)³¹ and to form a narrative in the market that can justify an adjustment in asset prices. To the extent that signals in nonverbal communication are decoupled from more standard measures of policy stance (recall that these objects are not perfectly correlated), one can observe rapid responses to changes in policy that are well understood (e.g., changes in the Fed funds rate) and delayed responses to less understood changes in policy (e.g., voice tone).

Third, media coverage may be an important force in financial markets (see Tetlock 2014 for a survey of this literature). Furthermore, financial markets may fail to incorporate information efficiently so that asset prices can respond to dated news. For example, Huberman and Regev (2001); Carvalho, Klagge, and Moench (2011); Tetlock (2011); and others document that financial markets react even to “stale news” (i.e., information that has been previously disclosed) covered by the media. To the extent that media coverage develops gradually while a consensus view about the message at a press conference is being reached, one then may observe a gradual response of the financial markets to variations in voice tone. Building on our exercise in Section III E, we can examine how the sentiment of media coverage for monetary policy evolves after FOMC meetings. Specifically, we use specification (4) with *MediaSentiment_t* as the dependent variable to construct an impulse response of media sentiment. We find a clear hump-shaped dynamic (panel B, Figure 16): after a positive voice tone, the media sentiment becomes more dovish on impact, it builds up for the next few days, peaks after around a week, and then converges back to zero. This pattern, coupled with the evidence of investor reactions to the media sentiment documented in the literature, suggests that FOMC-related news coverage and sentiment could be a channel through which the tone of voice can move financial markets and continued media coverage can contribute to the delayed responses.

IV. Concluding Remarks

Press conferences are an important communication tool for delivering and explaining monetary policy decisions to the public. Unlike press releases, transcripts, or minutes, a press conference contains both verbal and nonverbal channels. The latter offers an opportunity to communicate “soft” information. Machine learning applied to text analysis allowed researchers to measure messages in written policy documents more accurately to quantify the importance of soft information. Other parts of communication (emotions, moods, tones, body language) could be equally (if not more) important, thus, potentially enriching the policy toolkit. However, these forms of communication have proven to be particularly difficult to quantify. Building on recent advances in voice recognition and deep learning, we attempt to shed new light on the effects of nonverbal policy communication.

Our analysis of variation in the Fed chairs’ voice tone during Q&A sessions after FOMC meetings shows that nonverbal communication can have a statistically and

³¹ There is some anecdotal evidence that information revealed at FOMC press conferences disseminates gradually. For example, by March 21, 2021, the FOMC press conference held on March 17, 2021 had been watched 55,172 times on Yahoo Finance’s YouTube channel and 143,093 times on CNBC’s YouTube channel. By the time of writing this (December 2021), the number of views had risen to 59,988 and 174,148, respectively.

economically discernible effect on a variety of financial indicators. For example, our results suggest that the voice tone used in policy communication may have a significant effect on the stock market to a much greater extent than to that which is contained in the Fed's actions or actual words (texts). This reaction is consistent with the Fed communicating a more positive outlook for the economy and a lower probability of monetary tightening in the future. Inflation expectations and exchange rates also respond to voice tone. In contrast, the bond market appears to have more mixed reactions to vocal cues from the Fed chairs.

Although future research should dig deeper into understanding the nuances of using voice to communicate policy, our results clearly have important policy implications. *How* messages are spoken appears to be potentially as important as the content of the messages. That is, nonverbal communication is potentially a new instrument to deliver information to the public. The Fed watchers routinely sieve through policy texts to identify and interpret minute variations in words (e.g., a change from "modest" to "moderate"). With advances in voice/face recognition, one may expect another arms race in policy communication and, hence, one needs to be cautious when using nonverbal cues as a policy communication tool in order to avoid any unintended effects. This does not make the job of central bankers easier and potentially adds another qualification (voice control) to the job requirement. This also may become a prerequisite for any other roles which use a public arena for policy communication. Indeed, to paraphrase Ronald Reagan, how can a Fed chair not be an actor?

REFERENCES

- Apel, Mikael, and Marianna Blix Grimaldi. 2014. "How Informative Are Central Bank Minutes?" *Review of Economics* 65 (1): 53–76.
- Araci, Dogu. 2019. "FinBERT: Financial Sentiment Analysis with Pre-trained Language Models." Unpublished.
- Bauer, Michael D., and Eric T. Swanson. 2020. "The Fed's Response to Economic News Explains the 'Fed Information Effect'." NBER Working Paper 27013.
- Berger, Helge, Michael Ehrmann, and Marcel Fratzscher. 2011. "Monetary Policy in the Media." *Journal of Money, Credit and Banking* 43 (4): 689–709.
- Bhavan, Anjali, Pankaj Chauhan, and Rajiv Ratn Shah. 2019. "Bagged Support Vector Machines for Emotion Recognition from Speech." *Knowledge-Based Systems* 184: 104886.
- Bikhchandani, Sushil, David Hirshleifer, and Ivo Welch. 1992. "A Theory of Fads, Fashion, Custom, and Cultural Change as Informational Cascades." *Journal of Political Economy* 100 (5): 992–1026.
- Blinder, Alan S., Michael Ehrmann, Marcel Fratzscher, Jakob De Haan, and David-Jan Jansen. 2008. "Central Bank Communication and Monetary Policy: A Survey of Theory and Evidence." *Journal of Economic Literature* 46 (4): 910–45.
- Caffi, Claudia, and Richard W. Janney. 1994. "Toward a Pragmatics of Emotive Communication." *Journal of Pragmatics* 22 (3–4): 325–73.
- Campbell, Jeffrey R., Charles L. Evans, Jonas D. M. Fisher, and Alejandro Justiniano. 2012. "Macroeconomic Effects of Federal Reserve Forward Guidance." *Brookings Papers on Economic Activity* 42 (1): 1–54.
- Carvalho, Carlos, Nicholas Klagge, and Emanuel Moench. 2011. "The Persistent Effects of a False News Shock." *Journal of Empirical Finance* 18 (4): 597–615.
- Chava, Sudheer, Wendi Du, and Baridhi Malakar. 2021. "Do Managers Walk the Talk on Environmental and Social Issues?" SSRN Working Paper 3900814.
- Chava, Sudheer, Wendi Du, and Nikhil Paradkar. 2020. "More than Buzzwords? Firms' Discussions of Emerging Technologies in Earnings Conference Calls." SSRN 3862645.
- Chodorow-Reich, Gabriel. 2014. "Effects of Unconventional Monetary Policy on Financial Institutions." *Brookings Papers on Economic Activity* 44 (1): 155–227.
- Choudhury, Akash Roy, Anik Ghosh, Rahul Pandey, and Subhas Barman. 2018. "Emotion Recognition from Speech Signals using Excitation Source and Spectral Features." In *2018 IEEE Applied*

- Signal Processing Conference*, edited by Sovan Dalai, Debangshu Dey, Biswendu Chatterjee, and Susanta Ray, 257–61. Piscataway, NJ: IEEE.
- Cieslak, Anna, Adair Morse, and Annette Vissing-Jorgensen.** 2019. “Stock Returns over the FOMC Cycle.” *Journal of Finance* 74 (5): 2201–48.
- Cieslak, Anna, and Hao Pang.** 2020. “Common Shocks in Stocks and Bonds.” NBER Working Paper 28184.
- Cieslak, Anna, and Annette Vissing-Jorgensen.** 2021. “The Economics of the Fed Put.” *Review of Financial Studies* 34 (9): 4045–89.
- Coenen, Günter, Michael Ehrmann, Gaetano Gaballo, Peter Hoffmann, Anton Nakov, Stefano Nardelli, Eric Persson, and Georg Strasser.** 2017. “Communication of Monetary Policy in Unconventional Times.” European Central Bank Working Paper 2080.
- Coibion, Olivier, Yuriy Gorodnichenko, Saten Kumar, and Mathieu Pedemonte.** 2020. “Inflation Expectations as a Policy Tool?” *Journal of International Economics* 124 (C): 103297.
- Coibion, Olivier, Yuriy Gorodnichenko, and Michael Weber.** 2019. “Monetary Policy Communications and Their Effects on Household Inflation Expectations.” NBER Working Paper 25482.
- Curti, Filippo, and Sophia Kazinnik.** 2021. “Let’s Face It: Quantifying the Impact of Nonverbal Communication in FOMC Press Conferences.” SSRN 3782239.
- Devlin, Jacob, Ming-Wei Chang, Kenton Lee, and Kristina Toutanova.** 2018. “Bert: Pre-training of Deep Bidirectional Transformers for Language Understanding.” Unpublished.
- Dricu, Mihai, Leonardo Ceravolo, Didier Grandjean, and Sascha Frühholz.** 2017. “Biased and Unbiased Perceptual Decision-making on Vocal Emotions.” *Scientific Reports* 7 (1): 1–16.
- Duffie, Darrell.** 2010. “Presidential Address: Asset Price Dynamics with Slow-Moving Capital.” *Journal of Finance* 65 (4): 1237–67.
- Ehrmann, Michael, and Marcel Fratzscher.** 2007. “The Timing of Central Bank Communication.” *European Journal of Political Economy* 23 (1): 124–45.
- Ehrmann, Michael, Gaetano Gaballo, Peter Hoffmann, and Georg Strasser.** 2019. “Can More Public Information Raise Uncertainty? The International Evidence on Forward Guidance.” *Journal of Monetary Economics* 108 (C): 93–112.
- Ehrmann, Michael, and Jonathan Talmi.** 2020. “Starting from a Blank Page? Semantic Similarity in Central Bank Communication and Market Volatility.” *Journal of Monetary Economics* 111: 48–62.
- Ehrmann, Michael, and Alena Wabitsch.** 2022. “Central Bank Communication with Non-experts – A Road to Nowhere?” *Journal of Monetary Economics* 127: 69–85.
- Enders, Zeno, Franziska Hünnekes, and Gernot J. Müller.** 2019. “Monetary Policy Announcements and Expectations: Evidence from German Firms.” *Journal of Monetary Economics* 108: 45–63.
- Fatum, Rasmus, and Barry Scholnick.** 2008. “Monetary Policy News and Exchange Rate Responses: Do Only Surprises Matter?” *Journal of Banking and Finance* 32 (6): 1076–86.
- Federal Open Market Committee.** 2011–2019. “FOMC Press Conferences.” https://www.youtube.com/watch?v=HdV2VUNh4E&list=PL159CD41EB36CFE86&ab_channel=FederalReserve (accessed November 28, 2022).
- Federal Open Market Committee.** 2011–2019. https://www.federalreserve.gov/monetarypolicy/fomc_historical.htm (accessed November 28, 2022).
- Fleckenstein, Matthias, Francis A. Longstaff, and Hanno Lustig.** 2014. “The TIPS-Treasury Bond Puzzle.” *Journal of Finance* 69 (5): 2151–97.
- Gao, Yuanbo, Baobin Li, Ning Wang, and Tingshao Zhu.** 2017. “Speech Emotion Recognition Using Local and Global Features.” In *International Conference on Brain Informatics*, edited by Yi Zeng, Yong He, Jeanette Hellgren Koteleski, Maryann Martone, Bo Xu, Hanchuan Peng, and Qingming Luo, 3–13. Berlin: Springer Nature.
- Gorodnichenko, Yuriy, Tho Pham, and Oleksandr Talavera.** 2023. “Replication Data for: The Voice of Monetary Policy.” American Economic Association [publisher], Inter-university Consortium for Political and Social Research [distributor]. <https://doi.org/10.3886/E178302V1>.
- Gorodnichenko, Yuriy, and Walker Ray.** 2017. “The Effects of Quantitative Easing: Taking a Cue from Treasury Auctions.” NBER Working Paper 24122.
- Gorodnichenko, Yuriy, and Michael Weber.** 2016. “Are Sticky Prices Costly? Evidence from the Stock Market.” *American Economic Review* 106 (1): 165–99.
- Gürkaynak, Refet S., Brian Sack, and Eric T. Swanson.** 2005a. “Do Actions Speak Louder than Words? The Response of Asset Prices to Monetary Policy Actions and Statements.” *International Journal of Central Banking* 1 (1): 55–93.
- Gürkaynak, Refet S., Brian Sack, and Eric T. Swanson.** 2005b. “The Sensitivity of Long-Term Interest Rates to Economic News: Evidence and Implications for Macroeconomic Models.” *American Economic Review* 95 (1): 425–36.

- Hansen, Stephen, and Michael McMahon.** 2016. "Shocking Language: Understanding the Macroeconomic Effects of Central Bank Communication." *Journal of International Economics* 99: S114–33.
- Hansen, Stephen, Michael McMahon, and Andrea Prat.** 2018. "Transparency and Deliberation within the FOMC: A Computational Linguistics Approach." *Quarterly Journal of Economics* 133 (2): 801–70.
- Hansen, Stephen, Michael McMahon, and Matthew Tong.** 2019. "The Long-Run Information Effect of Central Bank Communication." *Journal of Monetary Economics* 108 (C): 185–202.
- Hattori, Masazumi, Andreas Schrimpf, and Vladyslav Sushko.** 2016. "The Response of Tail Risk Perceptions to Unconventional Monetary Policy." *American Economic Journal: Macroeconomics* 8 (2): 111–36.
- Haver Analytics.** 2022. <https://www.haver.com/> (accessed November 28, 2022).
- Hobson, Jessen L., William J. Mayew, and Mohan Venkatachalam.** 2012. "Analyzing Speech to Detect Financial Misreporting." *Journal of Accounting Research* 50 (2): 349–92.
- Hu, Allen, and Song Ma.** 2020. "Human Interactions and Financial Investment: A Video-Based Approach." Unpublished.
- Huberman, Gur, and Tomer Regev.** 2001. "Contagious Speculation and a Cure for Cancer: A Nonevent that Made Stock Prices Soar." *Journal of Finance* 56 (1): 387–96.
- Husted, Lucas, John Rogers, and Bo Sun.** 2020. "Monetary Policy Uncertainty." *Journal of Monetary Economics* 115 (C): 20–36.
- Jha, Manish, Hongyi Liu, and Asaf Manela.** 2021. "Natural Disaster Effects on Popular Sentiment toward Finance." *Journal of Financial and Quantitative Analysis* 56 (7): 2584–2604.
- Jordà, Òscar.** 2005. "Estimation and Inference of Impulse Responses by Local Projections." *American Economic Review* 95 (1): 161–82.
- Kozłowski, Austin C., Matt Taddy, and James A. Evans.** 2019. "The Geometry of Culture: Analyzing the Meanings of Class through Word Embeddings." *American Sociological Review* 84 (5): 905–49.
- Krishnamurthy, Arvind, and Annette Vissing-Jørgensen.** 2011. "The Effects of Quantitative Easing on Interest Rates: Channels and Implications for Policy." *Brookings Papers on Economic Activity* 42 (2): 215–65.
- Kuttner, Kenneth N.** 2001. "Monetary Policy Surprises and Interest Rates: Evidence from the Fed Funds Futures Market." *Journal of Monetary Economics* 47 (3): 523–44.
- Lamla, Michael J., and Dmitri V. Vinogradov.** 2019. "Central Bank Announcements: Big News for Little People?" *Journal of Monetary Economics* 108: 21–38.
- Likitha, M. S., Sri Raksha R. Gupta, K. Hasitha, and A. Upendra Raju.** 2017. "Speech Based Human Emotion Recognition using MFCC." In *2017 International Conference on Wireless Communications, Signal Processing and Networking*, 2257–60. Piscataway, NJ: IEEE.
- Liu, Yinhan, Myle Ott, Naman Goyal, Jingfei Du, Mandar Joshi, Danqi Chen, Omer Levy, Mike Lewis, Luke Zettlemoyer, and Veselin Stoyanov.** 2019. "RoBERTa: A Robustly Optimized BERT Pretraining Approach." Unpublished.
- Livingstone, S.R., and F.A. Russo.** 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* 13 (5): e0196391.
- Livingstone, S.R., and F.A. Russo.** 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: Dataset." *PLoS ONE*. <https://doi.org/10.5281/zenodo.1188975> (accessed November 28, 2022).
- Lucca, David O., and Emanuel Moench.** 2015. "The Pre-FOMC Announcement Drift." *Journal of Finance* 70 (1): 329–71.
- Mayew, William J., and Mohan Venkatachalam.** 2012. "The Power of Voice: Managerial Affective States and Future Firm Performance." *Journal of Finance* 67 (1): 1–43.
- Mehrabian, Albert.** 1971. *Silent Messages: Implicit Communication of Emotions and Attitudes*. Belmont, CA: Wadsworth Publishing.
- Neuhierl, Andreas, and Michael Weber.** 2019. "Monetary Policy Communication, Policy Slope, and the Stock Market." *Journal of Monetary Economics* 108 (C): 140–55.
- Nexis Uni.** 2011–2019. <https://www.nexisuni.com> (access 28 November 2022).
- Pan, Yixiong, Peipei Shen, and Liping Shen.** 2012. "Speech Emotion Recognition Using Support Vector Machine." *International Journal of Smart Home* 6 (2): 101–108.
- Pichora-Fuller, M.K., and K. Dupuis.** 2010. "Toronto Emotional Speech Set." University of Toronto. <https://hdl.handle.net/1807/24487> (accessed November 28, 2022).
- Romer, Christina D., and David H. Romer.** 2000. "Federal Reserve Information and the Behavior of Interest Rates." *American Economic Review* 90 (3): 429–57.

- Rosa, Carlo. 2011a. "The High-Frequency Response of Exchange Rates to Monetary Policy Actions and Statements." *Journal of Banking and Finance* 35 (2): 478–89.
- Rosa, Carlo. 2011b. "Words That Shake Traders: The Stock Market's Reaction to Central Bank Communication in Real Time." *Journal of Empirical Finance* 18 (5): 915–34.
- Rozkrut, Marek, Krzysztof Rybiński, Lucyna Sztaba, and Radosław Szwaja. 2007. "Quest for Central Bank Communication: Does It Pay to Be 'Talkative'?" *European Journal of Political Economy* 23 (1): 176–206.
- Swanson, Eric T. 2021. "Measuring the Effects of Federal Reserve Forward Guidance and Asset Purchases on Financial Markets." *Journal of Monetary Economics* 118: 32–53.
- Swanson, Eric T. 2021. "Measuring the Effects of Federal Reserve Forward Guidance and Asset Purchases on Financial Markets: Dataset." *Journal of Monetary Economics*. <http://www.socsci.uci.edu/~swanson2/researchpublished.html> (accessed November 28, 2022).
- Tetlock, Paul C. 2011. "All the News That's Fit to Reprint: Do Investors React to Stale Information?" *Review of Financial Studies* 24 (5): 1481–1512.
- Tetlock, Paul C. 2014. "Information Transmission in Finance." *Annual Review of Financial Economics* 6 (1): 365–84.
- Tiingo. 2011–2019. <https://api.tiingo.com/documentation/end-of-day> (accessed November 28, 2022).
- Twitter Full-archive Search API. 2022. <https://developer.twitter.com/en/docs/twitter-api/tweets/search/introduction> (accessed November 28, 2022).
- Verma, Devika, and Debajyoti Mukhopadhyay. 2016. "Age Driven Automatic Speech Emotion Recognition System." In *2016 International Conference on Computing, Communication and Automation*, edited by Parma Nand Astya, Abhishek Swaroop, Vishnu Sharma, and Manjeet Singh, 1005–10. Piscataway, NJ: IEEE.
- Visschedijk, Gillian C., Ard W. Lazonder, Anja van der Hulst, Nathalie Vink, and Henny Leemkuil. 2013. "Modelling Human Emotions for Tactical Decision-Making Games." *British Journal of Educational Technology* 44 (2): 197–207.
- WRDS. 2022. "TAQ". <https://wrds-www.wharton.upenn.edu/> (accessed November 28, 2022).
- Wu, Jing Cynthia, and Fan Dora Xia. 2016. "Measuring the Macroeconomic Impact of Monetary Policy at the Zero Lower Bound." *Journal of Money, Credit and Banking* 48 (2–3): 253–91.
- Wu, J.C., and F.D. Xia. 2016. "Measuring the Macroeconomic Impact of Monetary Policy at the Zero Lower Bound: Dataset." *Journal of Money, Credit, and Banking*. <https://sites.google.com/view/jingcynthiawu/shadow-rates> (accessed November 28, 2022).
- Yahoo Finance. 2011–2019. "Company Earnings Calendar." <https://finance.yahoo.com/calendar/earnings/> (accessed November 28, 2022).
- Yahoo Finance. 2011–2019. "iShares Interest Rate Hedged Corporate Bond ETF (LQDH)." <https://finance.yahoo.com/quote/LQDH/history?p=LQDH> (accessed November 28, 2022).
- Zhou, Rhea Tingyu, and Rose Neng Lai., 2009. "Herding and Information Based Trading." *Journal of Empirical Finance* 16 (3): 388–93.
- Zisad, Sharif Noor, Mohammad Shahadat Hossain, and Karl Andersson. 2020. "Speech Emotion Recognition in Neurological Disorders Using Convolutional Neural Network." In *Brain Informatics: 13th International Conference Proceedings*, edited by Mufti Mahmud, Stefano Vassanelli, M. Shamim Kaiser, and Ning Zhong, 287–96. Berlin, Germany: Springer Nature.