IIMB-WP N0. 622/2022

WORKING PAPER NO: 622

Does the choice of words in the Fed's Board of Governors' speeches matter?

Abhinav Anand
Assistant Professor
Finance & Accounting
Indian Institute of Management Bangalore
Bannerghatta Road, Bangalore – 5600 76
abhinav.anand@iimb.ac.in

Sankarshan Basu
Professor
Finance & Accounting
Indian Institute of Management Bangalore
Bannerghatta Road, Bangalore – 5600 76
sankarshan.basu@iimb.ac.in

Jalaj Pathak
Assistant Professor
Finance & Accounting
Great Lakes Institute of Management
Chennai
jalaj.p@greatlakes.edu.in

Ashok Thampy
Professor
Finance & Accounting
Indian Institute of Management Bangalore
Bannerghatta Road, Bangalore – 5600 76
ashok.thampy@iimb.ac.in

Year of Publication – July 2022

Does the choice of words in the Fed's Board of Governors' speeches matter?*

Abhinav Anand† Sankarshan Basu‡ Jalaj Pathak§
Ashok Thampy¶

Abstract

Yes it does. The US stock market shows significant movement in response to the tone of Fed speeches, the day the speech is delivered. Negative speeches depress returns and amplify volatility, and the stock market moves more strongly in response to forward-looking speeches. In contrast, current text quantification techniques fail to exhibit any market impact of speeches. Our study introduces two improvements: i) a sentence-based n-gram analysis that quantifies the connotation of multi-clausal phrases (e.g., 'a slowdown in business profitability'); and ii) usage of an augmented financial dictionary which incorporates 'valence shifters': adjectives, adverbs and conjunctions (e.g., 'large', 'slightly', 'although') which alter the connotation of text but have been ignored in literature. We show that valence shifter usage in Fed speeches is the highest during the Great Recession and the Eurozone crisis, and that valence shifters are used to inject more positivity in speeches.

Keywords: Central Bank Communication, Tone Analysis, Financial Text Analysis, Federal Reserve speeches

JEL Classification: G14, G18, G28, G41

*We thank Sanchit Jain for excellent research assistance and Bidisha Chakrabarty for constructive comments and criticisms.

†Indian Institute of Management Bangalore, Finance & Accounting

‡Indian Institute of Management Bangalore, Finance & Accounting

§Indian Institute of Management Lucknow, Finance & Accounting

¶Indian Institute of Management Bangalore, Finance & Accounting

"Monetary policy is 98% talk and only 2% action."
Former Fed chairman Ben Bernanke on the show '60 Minutes'.[1]

1 Introduction

Due to its prime importance in conducting monetary policy and maintaining financial stability, all aspects of Federal Reserve communication are watched very closely by market participants. Although a large collection of papers have been published analyzing the impact of press releases and FOMC statements [Lucca and Trebbi, 2009, Hansen and McMahon, 2016, Gonzalez and Tadle, 2021], we examine a very important yet understudied tool in the Federal Reserve communication toolkit: the role of speeches delivered by members of the Board of Governors of the Federal Reserve.[2]

In general, central bank communication has been found to be significantly associated with a wide variety of economic variables such as interest rates [Kohn and Sack, 2003, Demiralp and Jorda, 2004, Lucca and Trebbi, 2009, Smales and Apergis, 2017]; money supply [Gerlach, 2007]; currency markets [Dossani, 2018] as well as stock return and volatility [Ehrmann and Fratzscher, 2004, Apergis and Pragidis, 2019, Brusa et al., 2020, Bodilsen et al., 2021]. According to Schmeling and Wagner [2019], central bank communication impacts market expectations, which in turn, could influence asset prices owing to its informational content [Savor and Wilson, 2013].

In particular, since the Fed Board of Governors' main duty is the administering of monetary policy and the stewardship of the Federal Open Market Committee, their speeches can contain critical information regarding their members' subjective estimates and priors for key policy variables, and can illuminate their past and future voting decisions. The Federal Reserve (and indeed all central banks in general) intends to communicate to markets by means of forward guidance, its

[1]See link to the news story here: https://www.economist.com/books-and-arts/2015/10/17/more-talk-more-action

[2]Neuhierl and Weber [2019] in a related paper, investigate speeches of FOMC chairs and vice-chairs, and show that a more hawkish tone predicts faster monetary policy tightening.

preferences regarding future trajectories of relevant policy variables such as short term interest rates, inflation, inflation expectations etc. For the Board of Governors, it is essential that such speech-based communication is conveyed accurately, and is interpreted in the manner envisaged by the Fed. Any miscommunication or misinterpretation in this regard can prove quite costly to the financial markets in particular and to the economy in general. Hence, a study of how the Fed's Board of Governors' speeches, and their choice of words move markets is vital for the policymaker who wishes to transmit accurate information clearly and unambiguously to market participants. Moreover, insofar as central bank communication can itself be used for policy implementation as suggested in Guthrie and Wright [2000], any evidence which connects the impact of Fed speeches to movements in the markets helps the central banker in gauging whether it is successfully conveying its message.[3]

From the perspective of investors as well, it is vital that their evaluation of the Federal Reserve speeches—both in content and intent—is accurate and in line with the objectives of the Fed. Stock market securities are priced at a premium to risk-free assets such as T-bills whose yields are directly affected by both benchmark interest rates as well as future inflation expectations—both of which are influenced by Federal Reserve forward guidance. Hence clearly, all stock market participants scrutinize speeches delivered by the Fed Board of Governors extremely closely. Hence, it is quite likely that speeches by members of the Board of Governors do end up influencing movements in the US stock market. Presumably, speeches with positive content and tone should improve stock market sentiment, and those with dire warnings about the current (or future) state of affairs should depress market expectations.

Despite such a clear theoretical implication of Fed speeches impacting markets, we show that current methods of financial texts' tone quantification show no impact of Fed speeches on market outcomes. We re-evaluate the impact of Fed speeches on the US market, and pay particular attention to the Board of Governors' proclivity to employ nuanced language which makes use of multi-clausal

[3]This is reflected in the quotation from Ben Bernanke used at the beginning of our paper.

phrases and connotation-altering modifiers (e.g., adjectives, adverbs etc.) that makes their content harder to interpret.

In order to analyze the tone of speeches delivered by officials of the Federal Reserve, we subject the speeches to financial text analysis. The Loughran and McDonald dictionary [Loughran and McDonald, 2011] along with the "bag-of-words" and n-gram approach [Tetlock, 2007, Li, 2008, Tetlock et al., 2008] are key tools, and have been used to examine the impact of varied financial communication—ranging from central bank communication to conference call tone [Brockman et al., 2017] and CEO letters [Boudt and Thewissen, 2019].[4] We evaluate the impact of Board of Governors' speeches on US market outcomes using the standard LM dictionary with bag-of-words approach (LM hereafter); show that such an exercise indicates no impact of Fed speeches on US markets; and contribute to this literature by introducing improvements in financial texts' tone quantification process, which captures the sizeable impact of Fed speeches on US market outcomes. We do this via two innovations:

1. using the sentence as the unit of analysis, which is a plausible solution of the as yet unsolved problem of how many words to include at a time in the tone quantification process [Andreevskaia and Bergler, 2008], and

2. by using "valence shifters"—adjectives and adverbs such as "but", "large", "barely" etc.—which modify the meaning of the sentence [Kennedy and Inkpen, 2006, Polanyi and Zaenen, 2006, Schulder et al., 2018] but have not been granted any weight in the LM dictionary.[5]

   As an illustration, consider the following simple sentences:

   (a) We expect a decrease in losses next quarter.

   (b) We expect a *slight* decrease in losses next quarter.

   (c) We expect a *major* decrease in losses next quarter.

   (d) We expect *not much* decrease in losses next quarter.

[4] "bag-of-words" uses one word at a time while n-gram uses a cluster of $n \geq 2$ words at a time.
[5] The full list of valence shifters used in our speech sample in table A1.

(e) We expect a *large* decrease in losses next quarter *although* demand *has fallen.*

Clearly all sentences enumerated above have different connotations. The 'base' sentence is the first, and it expects a decreases in losses next quarter. Each succeeding sentence modifies its tone by describing it, e.g., 'slight', 'major', 'not much', 'large', 'although'—all of which have been ignored thus far in standard approaches such as LM. In fact, according to the LM approach, all the above sentences are assigned a tone of 0. Not only does it ignore the valence shifters, it also fails to include the connotation of the word 'decrease'.[6] However, our approach correctly classifies each sentence's tone by recognizing the connotation-altering role of valence shifters and by including 'demand has fallen' as the relevant 3-gram for the final sentence.

Further, in order to extract the tone from Fed speeches, in addition to the usage of the LM dictionary, we also use the dictionary specified in Apel and Grimaldi [2014] and Apergis and Pragidis [2019] which characterize text with respect to central bank communication. We extract polar phrases to accurately capture the connotation of verb-noun combinations such as "increasing stability" or "decrease in confidence". Our technique which combines multiple lexicons, unigrams, n-grams, polar phrases along with valence shifters and the usage of sentence as the unit of analysis ensures that manual intervention in assigning values to text is minimal which circumvents problems arising out of subjectivity, discrete classification etc. and enhances replicability and comparability for further research [Picault and Renault, 2017]. Further, the automated route to financial texts' tone extraction is superior in terms of both speed and scale as compared to manual or semi-automated tone assignment.

Our technique decomposes a speech into its constituent sentences and for each sentence, applies a suitable n-gram and assigns polarity to it by consulting an augmented LM dictionary with valence shifters and central bank-relevant terms. Somewhat relatedly, Pennebaker et al. [2003] argue that the entire corpus of text

[6]'Decrease' in itself is ambiguous: 'decreasing profits' is negative but 'decreasing losses' is positive.

and individual sentences within it, must be considered while assessing the meaning of the text. DuBay [2007] outlines how cognitive theorists and linguists in the 1970s elaborated that the meaning of a text is not in the independent words but is rather constructed by making inferences and interpretations on the whole. Our insistence on sentence-level n-gram categorization, and on valence shifters, ensures that this dictum is obeyed since it is able to assign proper weights to multi-clausal phrases, as well as to adjectives and adverbs—both of which can completely alter the connotation of the text.

We show that valence shifter usage in Fed speeches is the highest during episodes of market distress such as the Great Recession and the Eurozone debt crisis—both coinciding during the leadership of Ben Bernanke. We also find that Fed speeches use valence shifters to inject nuance into their text, and employ it to make otherwise positive speeches more positive, and negative speeches less negative. The added prolixity of speech text, as proxied by higher-than-usual valence shifter usage is conspicuous by its absence during times of relative calm in external market conditions. In interpreting such results however, we advocate caution, since aggregate valence shifter usage and more nuanced speech text cannot directly be attributed to the leadership of the Federal Reserve, since it also reflects the effect of policy preferences, market environment, and the prevailing uncertainties in the Fed's internal estimates.

We emphasize our paper's main contribution as follows. The current bag-of-words with LM dictionary approach fails to show any impact of Fed's Board of Governors' speeches on the movement of US stock index returns. This suggests—counter-intuitively—that markets do not react in any way to speeches delivered by Fed Board of Governors. However, once the polarity of multi-clausal phrases is accounted for, and valence-shifters and augmented financial lexicons are employed, Fed speeches do end up displaying a significant impact on US stock returns.

In particular, we show that Fed speeches—especially those that are forward-looking—impact the returns of US stock indices positively on the same day as the speech is delivered, implying that (all else equal) positive speeches raise returns and negative speeches depress returns. We also show that Fed speeches impact the volatility of the US stock markets negatively on the same day as the speech is delivered, implying that (all else equal) positive speeches decrease market volatility and negative speeches amplify market volatility. Further, we demonstrate that speeches on topics relevant to risk premia in the financial markets have a much greater impact on market returns than those on other topics; and that (all else equal) positive Fed speeches reduce the US term premium for both short and long maturity bonds.

The paper is organized as follows, section 2 is the literature review for Federal Reserve communication in particular, and central bank communication in general. Section 3 specifies the methodology for tone calculation followed by section 4 which describes the data sources. Section 5 investigates usage of valence shifters across time, topics and Fed Chairs. Section 6 outlines the impact of Fed speeches on US index returns and volatility. Section 7 tests for robustness and finally, section 8 offers concluding remarks.

2 Literature review

Due to the perceived economic and financial importance of central banks, a diverse number of studies have investigated their impact. For example, among the earliest such studies, Guthrie and Wright [2000] investigate how central bank statement—rather than open market operations—can be used to implement monetary policy in New Zealand. Romer and Romer [2004] analyze central bank communication using subjective assessment of the content and examine its impact on monetary policy. Bennani [2020] creates a measure of Fed chair's "confidence" and find that this measure is positively and significantly related to investor sentiment. Dybowski and Kempa [2020] use a topic modelling approach to examine the stance of the European Central Bank communication and find that the ECB has shifted its focus from monetary policy towards the stability of European financial system. Gonzalez and Tadle [2021] find that the press releases of most central banks converge during periods of international crises.

In particular, central bank communication has been found to impact several aspects of the financial markets. We outline some of these in the subsections below.

2.1 Impact of Fed communication

2.1.1 Impact on stock returns and volatility

Rosa [2011] examines the intraday impact of FOMC statements on the US intraday stock and volatility indices and reports significant results. Savor and Wilson [2013] find that the average market return and Sharpe ratio are significantly higher on important announcement days. Lucca and Moench [2015] report large average excess pre-FOMC returns on the US equities but no impact on treasury bills. Gu et al. [2018] report that the U.S stock prices tend to increase after the FOMC announcements which are accompanied by SEP (Summary of Economic Projections) releases. Cieslak et al. [2019] inspect the association between the equity premium and FOMC meetings days and report major impact in weeks 0, 2, 4 and 6 of the FOMC cycle. Bodilsen et al. [2021] also report that FOMC meetings followed by press conferences are significantly associated with stock return.

2.1.2 Impact on interest rates, treasury yields, currency markets etc.

Kohn and Sack [2003] analyze central bank communication using dummy categorization of the content and find that it significantly impacts the interest rates. Pasquariello [2007] present evidence on how central bank interventions significantly impact the price formation in the foreign exchange markets. Lucca and Trebbi [2009] use Google search and Factiva based news articles in an n-gram approach to analyze FOMC announcements and find that they are significantly associated with treasury yields. Hansen and McMahon [2016] use a topic analysis approach on FOMC communication to analyze its impact on the market using a FAVAR framework and report significant association with treasury yields but not on real economic variables. On similar lines, Smales and Apergis [2017] extract the readability of monetary policy statements using the Flesch-Kincaid index and present its impact on 10 year T-bills. Dossani [2018] examines how the central bank press conferences impact risk premia in the currency markets and finds significant results.

2.2 Impact of ECB and other central banks' communication

Jansen and De Haan [2006] examine comments by European central bankers on the interest rate, inflation, and economic growth in Eurozone and find that such comments are often contradictory to each other. Gerlach [2007] discusses interest rate related statements made by the ECB and their respective impact using subjective dummy classification of the statement by the authors. Picault and Renault [2017] use n-gram and term weighing approach to quantify ECB communication and find that "markets are more (less) volatile on the day following a conference with a negative (positive) tone about the Euro area economic outlook". Schmeling and Wagner [2019] and Apergis and Pragidis [2019] also quantify central bank tone and find that it is significantly associated with both return and volatility. Bennani et al. [2020] examine the ad-hoc communication by the ECB and find that the text measure derived from such communication provides additional information about future monetary policy decisions and is also significantly associated with the future ECB rate changes. Most recently, Leombroni et al. [2021] report that ECB's monetary policy communication on regular announcement days led to significant yield spread during sovereign debt crises.

To summarize, while there has been a large collection of papers on different aspects of central bank communication's impact on market outcomes in general, the Fed Board of Governors' speeches have not been analyzed systematically thus far. Further, anecdotal evidence—especially from hedge fund operators, and active fund managers—suggests heavy interest from market participants and institutional investors in the speeches delivered by the Board of Governors. Our paper contributes by investigating this relatively understudied aspect of Fed communication in detail.

3 Methodology

3.1 Tone Quantification

We decompose a Fed speech into the collection of its constituent sentences. The tone of the speech is the average tone of the sentences it is composed of. In instances where there are multiple speeches on the same day, the content for all speeches is analyzed as that belonging to one composite speech.

To identify sentences in the text, we adopt the following procedure. We parse the content and convert it to all lower cases. We remove references (if any) from the content, and identify all possible punctuation marks in the text. Following this, the text between two full stops, a full stop and a question mark, and between two question marks is classified as a sentence. We classify words in each sentence into two categories: valence shifters (adjectives, adverbs and adversative conjunctions), and polar (positive/negative) words/phrases.

The collection of polar words are taken from the LM dictionary [Loughran and McDonald, 2011] and the phrases are extracted according to Apel and Blix Grimaldi [2012] and Apergis and Pragidis [2019]. The phrases comprise two parts: a verb and a noun. The nouns are taken from the Economist's "Economics Dictionary"[7] and the verb list includes all the verb forms not classified in the LM dictionary, such as "increase", "decrease", "reduce", "fall", "raise" etc. These verb-noun combinations are identified as n-gram units ($2 \leq n \leq 5$) and are categorized as either positive or negative. For example, phrases with a noun and a verb-form such as "increasing growth", or "rising employment" are treated as positive and others such as "increase in unemployment", "fall in output" and "decrease in growth" are categorized as negative. We find that for the full sample of speeches, approximately 51% of sentences contain one or more polar words from the LM dictionary and an additional 14% sentences contain one or more of the macro-related nouns and verbs not classified in the LM dictionary. Thus, by augmenting the LM dictionary with verb-noun combinations, a more extensive portion of speeches can be quantified as compared to using just n-grams or bag-of-words with LM dictio-

[7] https://www.economist.com/economics-a-to-z/

nary based approach. In effect, our insistence on the sentence as the unit of tone quantification, and the usage of the augmented LM dictionary with multi-clausal verb-noun combinations, leads to an n-gram analysis (n words at a time) where n changes from sentence to sentence.

In addition to the sentence based n-gram analysis, we employ adjectives, adverbs and (adversative) conjunctions—which modify the meaning of sentences—to impart polarity to words and phrases which have been ignored in the LM dictionary. The valence shifters and their respective weights are taken from Kennedy and Inkpen [2006], Polanyi and Zaenen [2006] and Schulder et al. [2018]. These valence shifters are further classified into four categories: amplifiers ("absolutely", "acutely", "very"), de-amplifiers ("barely", "faintly", "few"), negators ("not", "cannot") and adversative conjunction ("despite", "but"). The amplifiers, de-amplifiers, and adversative conjunction are given a weight of 0.8: positive for an amplifier, negative for a de-amplifier, negative for the words before adversative conjunction; and positive for the words after adversative conjunction. This is because adversative conjunction such as "but" will amplify the tone after it and weight down the tone before it.[8] The negators are given a value of -1. The default weight of 0.8 is as per the existing literature but we verify our results by varying the weight of valence shifters from 0.5 to 0.9 and confirm that the findings continue to hold. For example, in the section on robustess, we re-examine our benchmark results—which remain essentially unchanged—when valence shifters are assigned a weight of 0.5.

We note that the term "valence shifters" has been used in Máté et al. [2021] for characterizing verbs such as "increase" in verb-noun combinations in order to quantify the Hungarian central bank tone. However, our usage of the term 'valence shifters' is quite distinct since it is adapted for use from the communications and computational linguistics literature and is used for characterizing adverbs and adjectives that modify the meaning of sentences. We do use the verb-noun combinations employed by Máté et al. [2021] but go a step further to incorporate the effect of adverbs and adjectives such as 'massive', 'only', 'but', 'faintly' etc. on the

[8] For example, consider the sentence, "The economy is doing well *but* the rising prices are a concern."

meaning, and hence the tone of sentences.

To summarize, for each sentence, first, the polar words/phrases are identified and given the weight of +1/-1, following which valence shifters are identified around each polar word/phrase from the beginning till the end of the sentence. Thus, each polar word/phrase along with its set of valence shifters are classified as a word cluster for each sentence.

In comparison to the sentence-level n-gram approach and augmented LM dictionary, the existing bag-of-words (unigram) with LM dictionary approach can lead to incorrect quantification of tone. As an illustration, consider the following hypothetical sentences below:

1. We expect to witness an increase in business activity.
2. We expect to witness a *slight* increase in business activity.
3. We expect to witness a *major* increase in business activity.
4. We expect to witness a *not much* increase in business activity.
5. We expect to witness a *large* increase in business activity *although* demand has *fallen*.

Clearly, while superficially similar, all sentences enumerated above are quite different in their connotation. For all hypothetical example sentences presented above, the unigram LM dictionary methodology assigns a score of 0. This is because valence shifters ('slight', 'major', 'not much', 'large') are ignored, and words like 'increase' are assigned zero weight since 'profit increase' has positive connotation, while 'unemployment increase' has a negative connotation; and hence a unigram approach is incapable of assigning polarity to it. More illustrations of sentence-level comparisons between our method and the LM approach are presented in table 1.

As another, more realistic illustration, we include from our sample, a speech delivered by the then Vice Chair of the Board of Governors (Donald Kohn) on March 16, 2006:

Table 1: Illustration of tones of sentences as per new (NM) and old (LM) methodologies

Valence Shifter Type	Valence Shifter Word	Sentence	Multi-Clausal Phrase	Tone LM	Tone New Methodology	Comment
None	NA	"We expect to witness an increase in business activity."	increase in business activity	0	+0.16	"increase" is not quantified in bag-of-words approach and hence tone is 0.
De-Amplifier	"slight"	"We expect to witness a *slight* increase in business activity."	increase in business activity	0	+0.02	"slight" discounts the positive impact of "increase in business activity" thus dampening the positive connotation of the sentence.
Amplifier	"major"	"We expect to witness a *major* increase in business activity."	increase in business activity	0	+0.25	"major" amplifies the impact of the multi-clausal phrase, thus intensifying the positive tone of the sentence.
Negator	"not"	"We expect to witness a *not much* increase in business activity."	increase in business activity	0	+0.02	"not" changes the sign of the amplifier "much" and thus decreases the positive tone of the multi-clausal phrase.
Amplifier Adversative Conjunction	"large" "although"	"We expect to witness a *large* increase in business activity *although* demand *has fallen*."	increase in business activity, demand has fallen	0	+0.25	"large" increases the impact of positive multi-clausal phrase but "although" moderates the negative effect of the multi-clausal phrase "demand has fallen".

> *"In general, we have a very poor understanding of the forces driving speculative bubbles and the role played by monetary policy"*

Using the bag-of-words with LM dictionary, the tone of the above sentence is calculated as:

$$\frac{(-1)[\text{=poor}]}{11} = -0.09$$

Now, using the methodology specified in this paper, the tone is calculated as follows.

First, polar words/phrases are identified from the sentence followed by valence shifters around these polar words/phrases. For example, when the first polar word (PW_1) is identified in the sentence, our method looks for valence shifters prior to it (PW_1) i.e., from the beginning of the sentence. Similarly, for the next polar word (PW_2), the search for valence shifters occurs between PW_1 and PW_2 and so on. This procedure is conducted for all valence shifters. Thus each sentence is divided into clusters with respect to polar words/phrases. In terms of the speech fragment analyzed before, our procedure can be broken down into the following steps:

1. *In general, we have a **very poor** understanding of the forces driving speculative bubbles and*

2. *the role played by monetary policy*

Thus, the sentence is divided into two clusters with ***very*** being a valence shifter (amplifier) to the polar word "**poor**" in the first cluster.

The tone is calculated is as follows:

$$(-0.8)[\text{=very}] + (-1)[\text{=poor}] = -1.8$$

$$\frac{(-1.8)[\text{=first cluster}] + (0)[\text{=second cluster}]}{12} = -0.15$$

The number of non stop-words in the denominator is one unit higher in case of the new methodology due to the enumeration of one valence shifter ('very') which was ignored in the existing methodology.

Comparing the tones of the sentence due to the existing methodology ($=-0.09$) and the new methodology ($=-0.15$) reveals a stark difference (66%) between the degree of negativeness embedded in just one sentence. Hence the comparable effect on the whole speech corpus can be substantial. While the existing methodology classifies the sentence as slightly negative, the new methodology categorizes it as quite negative—primarily on account of correctly identifying "very" as a negative tone intensifier. This aspect is completely ignored in the existing methodology.

3.2 Empirical design

We test the hypothesis that daily returns of the US stock market indices are significantly associated with the Federal Reserve speech tone.

The following regression specifications are tested for the returns and for the volatility:

$$R_t = a_0 + b_0 Tone_t + \sum_{i=1}^{3} c_i R_{t-i} + d_1 * Time\_Controls_t + \\ d_2 * Speech\_Controls_t + d_3 * Macro\_Controls_t + d_4 * Position\_FE + u_t \quad (1)$$

$$Vol_t = a_0 + b_0 Tone_t + \sum_{i=1}^{3} c_i Vol_{t-i} + d_1 * Time\_Controls_t + \\ d_2 * Speech\_Controls_t + d_3 * Macro\_Controls_t + d_4 * Position\_FE + u_t \quad (2)$$

In both equations above, time controls include the day of the week and month dummy, and speech controls include average words per sentence ('AWPS') and percentage of complex words ('Per_CW'). These two variables of speech controls are the main constituents of the three widely used readability measures: the Fog Index, the Flesch-Kincaid Index and the SMOG Index. Thus we use these speech controls to account for the readability and complexity of speeches [Gunning, 1952, Li, 2008, Biddle et al., 2009, Miller, 2010]. The lags of return as control are kept in accordance with previous studies which examine the impact of central bank communication on index returns [Ehrmann and Fratzscher, 2007, Born et al., 2014, Gertler and Horvath, 2018].

In addition, we include macroeconomic variables as control factors. These include the real exchange rate, the term premium and the Bloomberg Economic Surprise Index (ESI). The term premium is calculated as the difference between the yield of 1 and 20 year bond. The Bloomberg ESI calculates the surprise element as the percentage point difference between analysts' forecasts of a wide variety of economic variables—such as jobless claims, pending home sales, consumer confidence, index of industrial production etc.—and the published value of economic data. In order to account for time-invariant, fixed aspects of the Fed personnel who delivers the speech, we include 'Position' fixed effects as well.

4 Data

Data for this study come from several sources: Federal Reserve's speeches are downloaded from the Federal Reserve website (https://www.federalreserve.gov/newsevents/speeches.htm); and data for stock indices, VIX, controls and macro variables are taken from Bloomberg. The Fed Funds rate data are downloaded from the St. Louis Fed website (FRED).[9]

All speeches are downloaded automatically using web parsing from the official Federal Reserve website. The speech data are available for the Fed from January 2006 to February 2020. Our sample excludes FOMC announcements, since we are interested only on the effects of speeches by Fed Board of Governors. The Fed database contains speeches for all members of the Federal Reserve Board and these officials can be either Chairpersons, Vice-Chairs or Governors.[10] The total sample

[9] https://fred.stlouisfed.org/series/FEDFUNDS.

[10] The collection of speeches contains one speech by 'Other officials' but we exclude it from the

includes speeches by Federal Reserve Chairpersons, Vice-Chairpersons, Governors, Vice Chair for Supervision, and Director of the Division of Monetary Affairs. In our sample, out of the total 797 speeches, 241 are by Chairpersons.

5 The Board of Governors' Choice of Words

5.1 Prevalence of valence shifters in Fed speeches

Overall, there are 797 speeches in our sample, with an average of 4.1 speeches delivered per month. The average speech contains 3482 words; the longest speech has 10923 words;[11] while the shortest contains 237 words. A large majority of speeches (547) have an overall negative tone, with the mean speech tone being -0.06. The highest value of speech tone in our sample is 0.29, while the lowest is -0.34. The standard deviation is 0.09.

Overall, about 38% of sentences contain at least one valence shifter, with the highest proportion being that of amplifiers (53%), followed by negators (19%), adversative conjunctions (17%), and de-amplifiers (11%).

Table 2 presents examples of the presence and usage of various types of valence shifters in the speeches of the Federal Reserve along with the difference in tone quantification using the LM method and the new methodology (NM) introduced in this study. It is clear from the entries in the table that usage of words such as 'very', 'few' and 'but' can substantially alter the tone of the sentence.

Figure 1 presents the time series barplots of the overall percentage of valence shifters in Fed speeches as well as its four components over the years. As is clear from visually inspecting the figure, the usage of valence shifters in Fed's Board of Governors' speeches is the highest during the market distress episodes of the Great Recession and the Eurozone debt crisis. Aggregate valence shifter usage shoots up to its highest recorded value of around 61% during the Great Recession (May 2008); and during the peak of the Eurozone debt crisis (2010–2011), valence

current sample.

[11] This corresponds to the composite speech which is constructed after having converted all speeches delivered during a day into one.

Table 2: Examples of usage of valence shifters in Federal Reserve speeches

Valence Shifter Type	Valence Shifter Word	Sentence	Date and Speaker	Tone LM	Tone New Methodology	% change	Comment
Amplifier	"very"	"in general, we have a ***very*** poor understanding of the forces driving speculative bubbles and the role played by monetary policy."	Donald Kohn 16-03-2006	-0.09	-0.15	-67%	"very" accentuates the impact of "poor" thus intensifying the negative tone of the sentence.
De-Amplifier	"few"	"the reports on first-quarter earnings have been quite positive, and available measures of credit quality, such as credit ratings and loan defaults, show ***few*** signs of stress."	Mark Olson 25-05-2006	-0.060	-0.009	+83%	"few" discounts the negative impact of "stress" thus ameliorating the negative connotation of the sentence.
Adversative Conjuction	"but"	"meanwhile the exchange rate remains strong –***but*** being relatively stable is attracting little attention."	Stanley Fisher 19-05-2016	0.20	0.10	-100%	"but" discounts the impact of "strong" thus decreasing the impact of the phrase "the exchange rate remains strong".

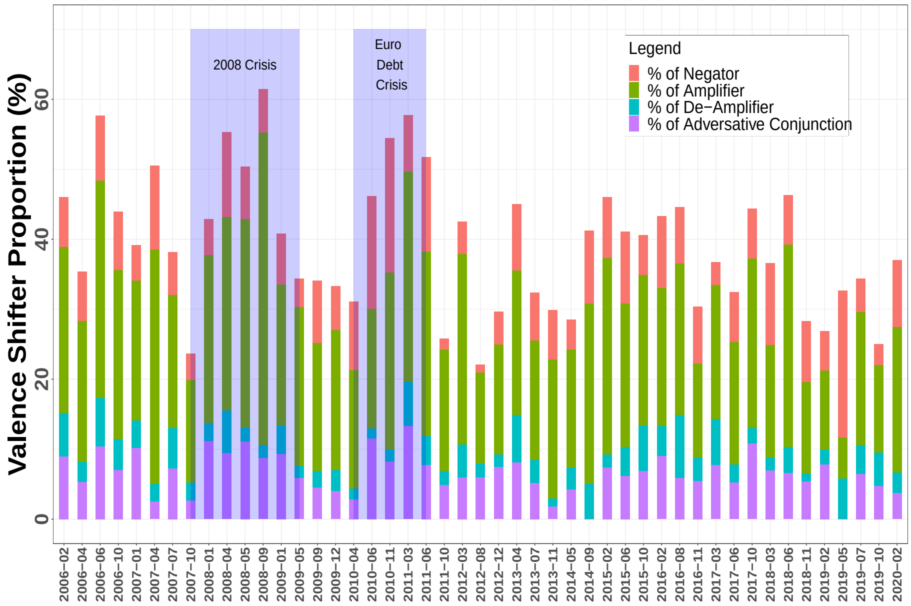


Figure 1: The figure presents the composition of overall percentage of valence shifters in Fed's Board of Governors' speeches, and its respective four components (negator, amplifier, de-amplifier and adversative conjunction) across time. The ticks on the x-axis correspond to delivery of speeches over the years 2006–2020.

shifter usage continues to be quite high—varying between 45%–55%. Interestingly, among valence shifters, the highest use is that of amplifiers (e.g., 'very'). Visual inspection also shows that during the two crisis episodes the share of adversative conjunctions ('however', 'although' etc.), and that of negators ('never', 'cannot' etc.) is at its highest.

High usage of valence shifters introduces semantic complexity into speeches, thereby making text harder to interpret. Hence increased proportion of valence shifters in Fed speeches during periods of market distress suggests an injection of additional nuance which remains conspicuous by its absence during tranquil market conditions. In particular, negators are harder to interpret [Carpenter and Just, 1975, Fischler et al., 1983, Christensen, 2009], and usage of such language has been shown by firms in the MD&A section of 10-K reports to exaggerate positive developments and understate the negative [Anand et al., 2022]. On the other hand, since these times were particularly turbulent, high usage of valence shifters could also be attributed to high prevailing uncertainty—not just externally in the markets, but also in Fed's Board of Governors' internal estimates.

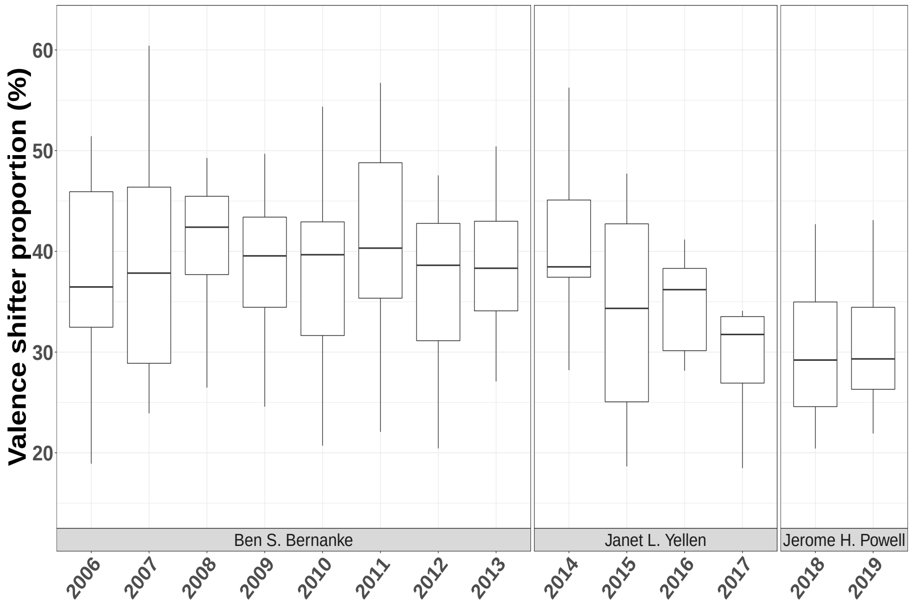
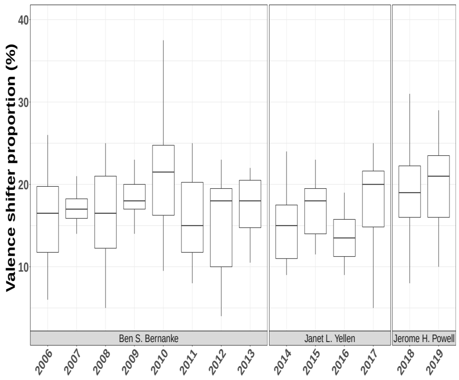


Figure 2: The figure presents the composition of overall percentage of valence shifters in Fed's Board of Governors' speeches, and its respective four components (negator, amplifier, de-amplifier and adversative conjunction) across time. The ticks on the x-axis correspond to delivery of speeches over the years 2006–2020.

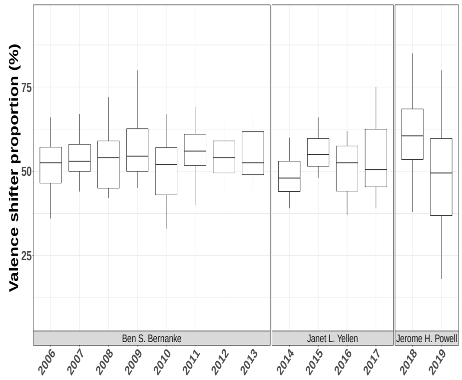
Visually inspecting figure 2 suggests that on aggregate, valence shifter usage is the highest during the tenure of Ben Bernanke, coinciding during the period of the Great Recession and the Eurozone debt crisis. Valence shifter usage appears to have dropped after 2014 during the tenure of Janet Yellen, and has lowered even further during Jerome Powell's chairmanship.

Further, figure 3 subdivides the valence shifer usage across Fed Chairs into its constituents—negators, amplifiers, de-amplifiers and adversative conjunctions. The following aggregate picture emerges: Ben Bernanke's tenure displays the highest use of negators (in 2010); Janet Yellen's tenure is marked by a high usage of de-amplifiers ('barely', 'faintly'), while Jerome Powell's tenure shows high usage of amplifiers.

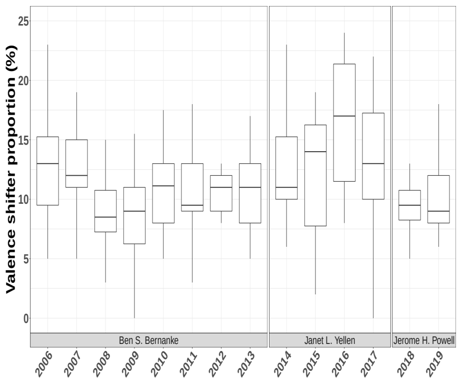
We note that such findings on the Fed's Board of Governors' choice of words cannot be directly attributed to the personalities leading the Fed. Each period was beset with its own challenges and external market conditions. Yet, aggregate difference in usage styles does shed an interesting light on the mindset, and policy priorities of Fed governors during different leaders' chairmanship.



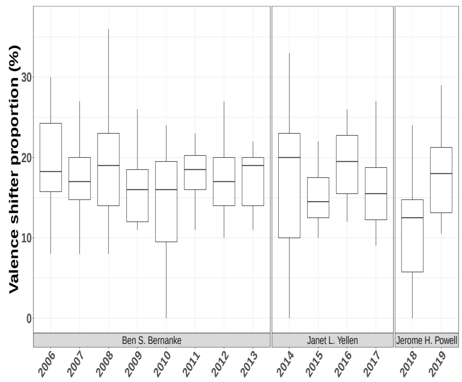
(a) Negator



(b) Amplifier



(c) De-Amplifier



(d) Adversative Conjunction

Figure 3: The plots present different proportions of valence shifter usage by its type: negators, amplifiers, de-amplifiers and adversative conjunction. The three sections present the tenure of Ben Bernanke, Janet Yellen and Jerome Powell respectively.

5.2 New methodology (NM) versus the old (LM)

To compare the difference between the level of speech tone, as quantified by the new (NM) and the old (LM) methodologies, we present figure 4 which calculates the relative difference between the two time series as $\frac{NM-LM}{LM}$. The median change is approximately 2 times and the maximum positive and negative changes in NM as compared to LM are 63.75 and -106.45 *times* respectively.

We also find that there are often large (absolute) differences in speech tones calculated according to NM and LM. For example, the tone of the speech by Governor Randall Kroszner on 30th November 2007, titled "Innovation, Information, and Regulation in Financial Markets" had a very mild negative tone of -0.0005 as per the LM methodology, whereas its NM tone was about 100 times more negative,

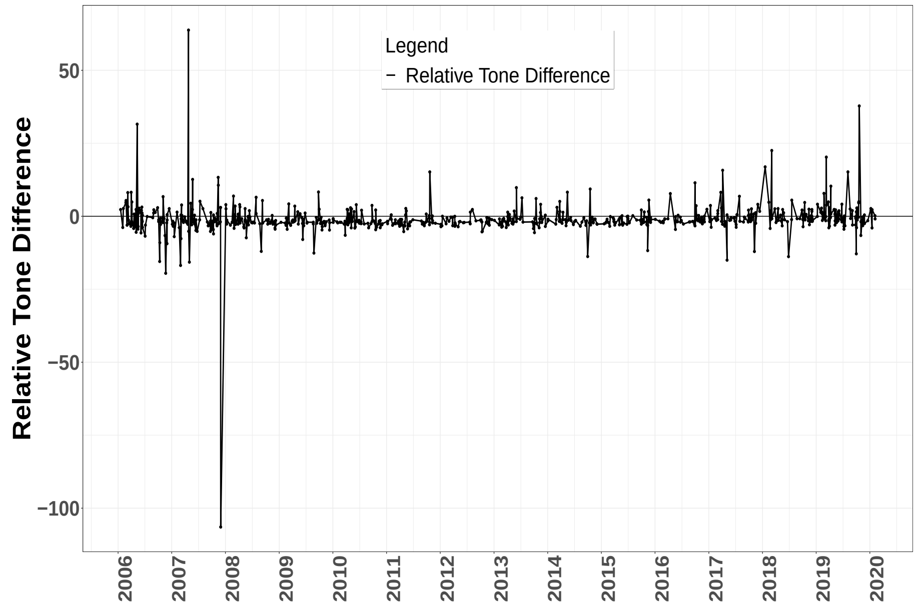


Figure 4: Relative tone difference between the old (LM) and new methodology (NM) calculated as $\frac{NM-LM}{LM}$.

at -0.0554. Similarly the tone of the speech delivered by Chairman Ben Bernanke in his welcoming remarks on 25th April 2007, had a slight positive LM tone of 0.0026 but its NM tone was much more positive at 0.1712.

Table 3: Relative Difference between LM and NM

Panel A: Both NM and LM Positive: 111/797 (14.14%)		
Case	% of sample	p-value
NM > LM	106 (95.50%)	0
NM < LM	5 (4.50%)	0.99
Panel B: Both NM and LM Negative: 616/797 (77.28%)		
Case	% of sample	p-value
NM > LM	42 (6.82%)	0.99
NM < LM	574 (93.18%)	0

Note: This table presents NM and LM tone difference and its relation to whether Fed Board of Governors use valence shifters to amplify or understate speech tone. NM denotes connotation according to the 'multi-clausal phrases and valence shifter' based new methodology, and 'LM' denotes the methodology taken from Loughran and McDonald [2011]. The p-value is that for the T test for equality of means.

We further examine valence shifter usage in table 3 where we analyze the difference in tone calculated using the new methodology (NM) and the existing methodology (LM). Panel A presents cases for which both LM and NM speech tones are positive (111 out of 797 speeches) and shows that out of those 111 speeches, in ∼96% cases the NM tone has a higher magnitude than that of the LM tone, with corresponding p-value 0. Its implication is that for positive speeches, the tone's positivity is amplified by the usage of valence shifters.[12] On the other hand, for speeches with negative tone—according to both NM and LM, comprising ∼77% of the total speeches—the tone according to NM is is less negative than that computed as per LM for ∼93% of the cases. Again, this implies that valence shifters are employed to decrease the negative-ness of the speech tone. These differences are quite significant, with corresponding p-values being 0.[13]

Table 4: Speech Statistics: New vs. Existing Methodology

Statistic	New	Existing
Min	-4.5365	-2.5655
Max	3.0173	1.4241
Mean	-0.1577	-0.1436
Median	-0.1507	-0.1925
SD	0.5582	0.3715
IQR	0.7492	0.5241

Note: This table presents the summary statistics for the speech tone of the sentences containing valence shifters calculated using the new methodology and the LM dictionary based bag-of-words approach.

Further, table 4 presents the difference in speech tone statistics for sentences with valence shifters, using the existing LM methodology and the new methodology introduced in this study. It presents clear evidence that the range of the speech tones is higher for the new methodology (NM) i.e., the minimum is lower and the maximum is higher in NM; the median calculated under NM is higher than that under the existing methodology; and the standard deviation and inter-quartile range are higher under the new method. The mean, being more susceptible to the presence of outliers, displays a lower value in the new method as compared to the old method.

Taken together, this suggests that the full variability of speech tones is systematically underestimated when valence shifters are ignored, as is done in the

[12]For 5 out of the 117 speeches the positive LM tone exceeds the NM tone but the corresponding p-value is 0.99, implying insignificance.

[13]For 42 out of the 797 speeches the negative magnitude of the LM tone exceeds the NM but the corresponding p-value is 0.99, implying insignificance.

bag-of-words with LM dictionary methodology. To further establish that the difference in speech tones outlined between the two techniques is significantly distant from each other, we examine the distance between the two speech tone distributions using the Kolmogorov Smirnov (KS) test, where the D-statistic (distance) has a value of 0.17, and the p-value of 0.

6 Impact on the US stock market

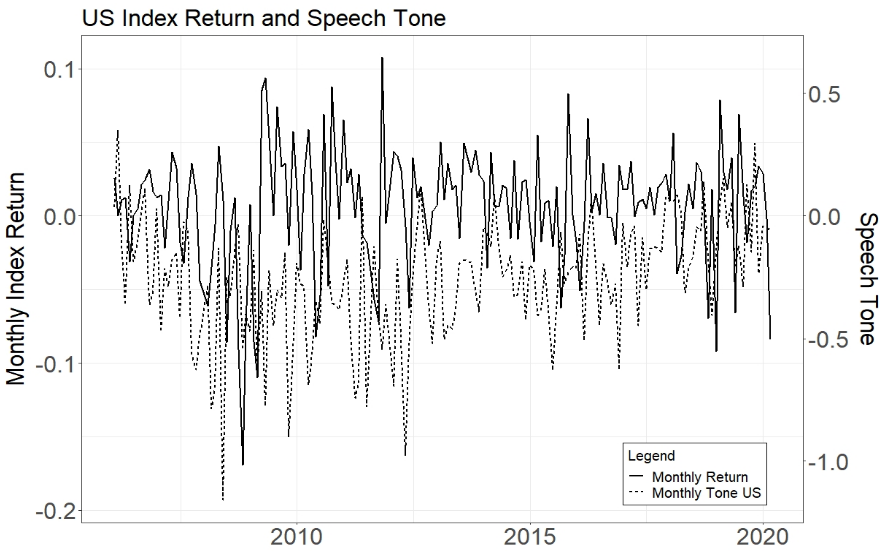


Figure 5: The time series of monthly S&P 500 index returns on the primary y axis; and the monthly speech tone on the secondary y-axis.

Figure 5 presents the time series of monthly S&P 500 index returns on the primary y axis, and the monthly speech tone on the secondary y-axis.[14] Broadly speaking, the two time series tend to co-move with each other which leads us to hypothesize a significant statistical relationship between the Federal Reserve speech tone and the US benchmark stock index return.

Similarly figure 6 presents the time series of monthly speech tone and the Fed Funds rate on the primary and secondary y-axes respectively. As the fig-

[14]The reason for choosing to display monthly movements in the two time series is due to their amenability for easy visual inspection.

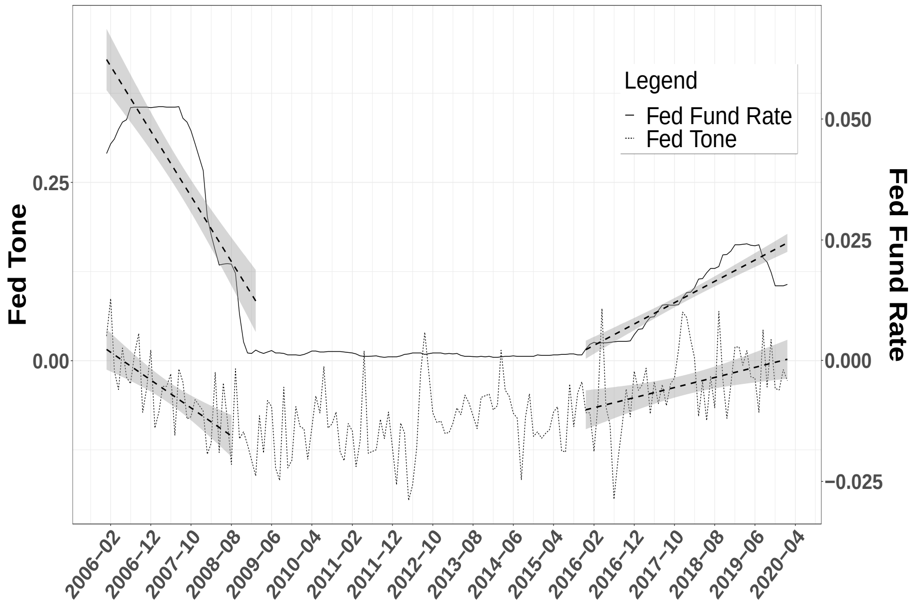


Figure 6: The time series of monthly Fed Tone on the primary y axis; and the monthly Fed Rate on the secondary y-axis.

ure demonstrates, there is significant comovement between the two time series. Broadly speaking there are three regimes: i) 2006–2009 in which the Fed Funds rate and the Fed speech tone show negative trends and falling values; ii) 2009–2016 where there is hardly any movement in the Fed Funds rate, and the Fed speech tone also displays no discernible trend; and iii) 2016–2020 where both the Fed Funds rate and the Fed speech tone show positive trends and increasing values.

Together, figures 5 and 6 provide strong preliminary visual evidence that there is a plausible statistical relationship between the new Fed speech tone introduced in this study and the US stock index return, as well as the US Fed Funds rate.

In the following subsections, we present extensive evidence that Fed speeches move US stock markets by impacting both returns and volatility of the market index. Our main result is that for both returns and volatility, Fed speeches impact stock markets on the same day as the speech is delivered. We also show that Federal Reserve's Board of Governors' speeches also impact the US term premium.

6.1 Impact on the S&P 500 daily returns

We examine the impact of Fed Board of Governors' speech tone on daily returns of the benchmark S&P 500 index. Table 5 presents the results in line with the regression specification in equation (1). The regression methodology is that of ordinary least squares with heteroskedasticity and autocorrelation consistent (HAC) errors.

In addition to the full set of speeches, we also include separately, the subsample of 'forward-looking' speeches (denoted as 'FL Speeches' in the table) and examine their impact on the US market indices for returns and volatility. The emphasis on the subset of forward-looking speeches is consistent with arguments in Ehrmann and Fratzscher [2007], who suggest that central banks mostly use communications as an expectation-management tool. Moreover, an added advantage is that forward-looking and future expectation based communication is less likely to be endogenous [Gertler and Horvath, 2018]. To classify speeches as forward-looking, we consider the pool of speeches that feature a higher-than-average proportion of terms associated with forward-looking communication, which are generally used to convey premeditated plans and actions. These include "believe", "estimate", "anticipate", "plan", "predict", "hope", "seek", "expect", "likely", "intend", "potential", "is likely to", "with the intent" etc. We calculate the frequency of such words and phrases for each speech in our sample and only consider the subsample of speeches for which the frequency is above the mean. In this way, 376 speeches are identified from our initial sample of 797 as forward-looking.

Table 5 presents the benchmark results with respect to the regression specification presented in equation (1). The main finding is that the Federal Reserve speech tone, quantified using the methodology specified in this study, significantly impacts the daily US stock index returns on the same day as the speech is delivered. Further, the coefficient estimate is positive (0.010) which implies that (all else equal) speeches with negative tone lead to a fall in the daily index return; and those with positive tone lead to a rise in the daily index return. Similarly, for the subset of speeches that are forward-looking, the impact of the speech is felt on the market index the same day as the delivery of the speech, and it is positive and

Table 5: Impact of Federal Reserve speech tone on the S&P 500 daily returns

$R_t = a_0 + b_0 Tone_t + d * Controls + u_t$				
Variables	**All Speeches**		**FL Speeches**	
NM	0.010*			
	(0.006)			
LM		0.023		
		(0.019)		
NM			0.030**	
			(0.013)	
LM				0.036
				(0.043)
Speech Controls	Yes	Yes	Yes	Yes
Macro Controls	Yes	Yes	Yes	Yes
Time Controls	Yes	Yes	Yes	Yes
Return Controls	Yes	Yes	Yes	Yes
Position FE	Yes	Yes	Yes	Yes
R^2	0.132	0.130	0.202	0.187
N	503	503	256	256

Note: This table presents the results from regressing daily index returns on speech tone (and controls). The results are reported in line with equation (1). The standard errors are reported in the parentheses and are all heteroskedasticity and autocorrelation (HAC) robust. 'FL Speeches' denotes the subsample of forward-looking speeches. The controls include three lags of return, day of the week, month dummy as well as speech level controls (average words per sentence (AWPS) percentage of complex words (Per_CW)); along with macroeconomic controls: the real exchange rate (Ex_Rate), Term Premium and the Bloomberg Economic Surprise Index (ESI), as well as the speech-giver's Position fixed effects. ***, ** and * indicate that the coefficient estimate are significantly different from zero at the 1 percent, 5 percent and 10 percent levels respectively.

statistically significant, with a coefficient that is three times the value as that for a generic speech. A unit standard deviation move in forward-looking Fed speech tone leads to 0.16 standard deviation move in the market index—which is about 2.5 times more than that for the whole set of speeches. In addition, the benchmark result is robust to the inclusion of several types of controls—whether they be speech-level controls, macro controls, lagged returns (3), day-of-the-week/month controls, or designation-level 'position' fixed effects for the Fed personnel delivering the speech. Moreover, the corresponding set of results for the LM based bag-of-words methodology fails to exhibit a coefficient significantly different from 0—for both the full sample and for the forward-looking speech subsample.

6.2 Impact on daily volatility

We also test whether the impact extends to the volatility of the US stock market index. To test this specification, we analyze speech tone effect on daily changes in the Chicago Board Options Exchange's (CBOE) Volatility Index (VIX) in line with the regression specification in equation (2).

Table 6: Impact of Federal Reserve speech tone on the VIX

$VIX_t = a_0 + b_0 Tone_t + d * Controls + u_t$				
Methodology	**All Speeches**		**FL Speeches**	
NM	-0.001			
	(0.001)			
LM		-0.003		
		(0.006)		
NM			-0.229***	
			(0.087)	
LM				-0.356
				(0.290)
Speech Controls	Yes	Yes	Yes	Yes
Macro Controls	Yes	Yes	Yes	Yes
Time Controls	Yes	Yes	Yes	Yes
Return Controls	Yes	Yes	Yes	Yes
Position FE	Yes	Yes	Yes	Yes
R^2	0.077	0.076	0.150	0.125
N	423	423	223	223

Note: This table presents the results from regressing daily VIX changes on speech tone (and controls). The results are reported in line with equation (2). The standard errors are reported in the parentheses and are all heteroskedasticity and autocorrelation (HAC) robust. The controls include three lags of return, day of the week, month dummy as well as speech level controls (average words per sentence (AWPS) percentage of complex words (Per_CW)); along with macroeconomic controls: the real exchange rate (Ex_Rate), Term Premium and the Bloomberg Economic Surprise Index (ESI), as well as the speech-giver's Position fixed effects. ***, ** and * indicate that the coefficient estimate are significantly different from zero at the 1 percent, 5 percent and 10 percent levels respectively.

Table 6 presents our results on the effect of speech tone on the changes in VIX. Our main result is that the speech tone of the forward-looking Federal Reserve speeches significantly impacts the daily US stock index realized volatility on the same day as the speech is delivered. Further, the coefficient estimate is negative (-0.229 for the forward-looking speeches) which implies that (all else equal) speeches with negative tone lead to amplified daily volatility; and those with positive tone lead to falls in daily volatility. A one standard deviation change in the forward-looking speech tone is associated with 0.22 standard deviation fall in VIX.

The current bag-of-words based LM dictionary approach fails to exhibit coefficient estimates' values significantly different from 0.

6.3 Impact of speeches on topics related to risk premia in financial markets

In order to investigate whether the impact of Federal Reserve speeches varies by the subject matter and content of the speeches, we conduct topic analysis using Latent Dirichlet Allocation (LDA) [Blei et al., 2003, Hansen et al., 2018].

Prior studies have found that there is a significant relationship between central bank communication and risk premia observed in the financial markets [Cieslak et al., 2019, Leombroni et al., 2021]. Further, Cieslak and Schrimpf [2019] report that the non-monetary component accounts for more than half the central bank communication and is significantly associated with financial markets outcomes.

In line with these observations, we segregate speeches which prominently feature words and terms strongly associated with risk premia in the financial markets.[15] We find that about 37% of the speeches in our sample incorporate such terms related to risk premia in the financial markets to a significant degree. Interestingly, these speeches feature a much higher proportion of valence shifters (42%) than the rest of the speeches (35%).

Our main findings are presented in table 7 and are quite similar to the benchmark results in table 5. The tone, as calculated according to the new methodology shows significant association with the S&P 500 returns on the same day as the speech is delivered. The coefficient estimate (0.030) is positive, indicating that (financial-market-risk-premia themed) positive speeches raise index returns and negative speeches depress returns. The same set of results follow even more strongly for the forward-looking speech sample also—both in terms of coefficient magnitude and its economic and statistical significance. Further, the impact of Fed speeches on topics related to risk premia in financial markets exceeds the

[15] The full list of words used in this categorization is as follows: "banks", "financial markets", "risk", "capital", "banking", "credit", "firms", "reserves", "liquidity", "interest rate", "crisis", "regulatory", "assets", "stress", "regulation", "basel", "lending", "insurance", "treasury", "leverage".

Table 7: Impact of Federal Reserve speech tone on topics related to risk premia in the financial markets on the S&P 500 daily returns

Methodology	All Speeches		FL Speeches	
Panel A: S&P500 Return				
$R_t = a_0 + b_0 Tone_t + d * Controls + u_t$				
NM	0.030**			
	(0.012)			
LM		0.062		
		(0.044)		
NM			0.050***	
			(0.023)	
LM				0.126
				(0.099)
R^2	0.162	0.148	0.368	0.348
N	183	183	102	102
Panel B: VIX				
$Vol_t = a_0 + b_0 Tone_t + d * Controls + u_t$				
NM	-0.218***			
	(0.079)			
LM		-0.415		
		(0.310)		
NM			-0.298*	
			(0.173)	
LM				-0.243
				(0.693)
R^2	0.177	0.079	0.089	0.203
N	162	162	104	104
Speech Controls	Yes	Yes	Yes	Yes
Macro Controls	Yes	Yes	Yes	Yes
Time Controls	Yes	Yes	Yes	Yes
Return Controls	Yes	Yes	Yes	Yes
Position FE	Yes	Yes	Yes	Yes

Note: This table presents the results from regressing daily index returns on speech tone (and controls). The results are reported in line with equation (1). The standard errors are reported in the parentheses and are all heteroskedasticity and autocorrelation (HAC) robust. The controls include three lags of return, day of the week, month dummy as well as speech level controls (average words per sentence (AWPS) percentage of complex words (Per_CW)); along with macroeconomic controls: the real exchange rate (Ex_Rate), Term Premium and the Bloomberg Economic Surprise Index (ESI) and the speech-giver's Position fixed effects. ***, ** and * indicate that the coefficient estimate are significantly different from zero at the 1 percent, 5 percent and 10 percent levels respectively.

benchmark results—both in statistical and economic significance—for the full set of speeches, as well as the pool of forward-looking speeches. In contrast, like the previously tabulated results, the bag-of-words based LM dictionary approach fails to display significance of any kind.

6.4 Impact on the US term premium

Gilchrist et al. [2019] examine the impact of the US monetary policy on dollar denominated sovereign bonds and find that US monetary easing leads to a significant narrowing of credit spreads on these bonds. Similarly, Tillmann [2020] examines the impact of monetary policy surprises on term structure of interest rates and reports that policy tightening leads to a significantly smaller increase in long-term bond yields. On similar lines, we also examine the impact of Fed speech tone on the US term premium and the results are presented in table 8. The term premium is calculated using the methodology specified in Adrian et al. [2013]. The data for the calculated term premium are available from the New York Fed website.[16]

Table 8: Impact of Fed speeches on the US bonds' term premia

	2 year	5 year	7 year	10 year
Panel A: All speeches				
NM Coefficient	-0.387**	-0.954**	-1.181**	-1.374**
	(0.175)	(0.408)	(0.498)	(0.576)
R^2	0.609	0.556	0.578	0.610
N	514	514	514	514
Panel B: Forward-looking speeches				
NM Coefficient	-0.692**	-1.449**	-1.835***	-2.223**
	(0.312)	(0.718)	(0.526)	(1.010)
R^2	0.605	0.550	0.509	0.615
N	261	261	261	261
Speech Controls	Yes	Yes	Yes	Yes
Macro Controls	Yes	Yes	Yes	Yes
Time Controls	Yes	Yes	Yes	Yes
Return Controls	Yes	Yes	Yes	Yes
Position FE	Yes	Yes	Yes	Yes

Note: This table presents the results from regressing term premia on speech tone (and controls). The results are reported in line with equation (1). The standard errors are reported in the parentheses and are all heteroskedasticity and autocorrelation (HAC) robust. The controls include three lags of return, day of the week, month dummy as well as speech level controls (average words per sentence (AWPS) percentage of complex words (Per_CW)); along with macroeconomic controls: the real exchange rate (Ex_Rate) and the Bloomberg Economic Surprise Index (ESI) and speech-giver's Position fixed effects. ***, ** and * indicate that the coefficient estimate are significantly different from zero at the 1 percent, 5 percent and 10 percent levels respectively.

Our findings can be summarized thus: for all maturity periods—2, 5, 7, 10 years—the Fed Board of Governors' speeches impact term premia significantly negatively. Further, the magnitude of the impact steadily rises as the maturity

[16] https://www.newyorkfed.org/research/data_indicators/term_premia.html

lengthens. Moreover, the impact is even more pronounced—both statistically and economically—for forward-looking set of speeches. To summarize (all else equal) an increase in Fed speech positivity reduces term premia for US bonds for all durations, and the degree of reduction in premia is higher for speeches which are forward-looking. Our results are consistent with Bundick et al. [2017], who argue that a positive Fed outlook leads to a fall in economic uncertainty and thus a fall in term premium component of yields.

7 Robustness

We subject the results to a number of robustness tests: i) changing valence shifter weights, ii) accounting for FOMC announcements, and iii) investigating potential reverse causality.

7.1 Changing valence shifter weight

Table 9: Impact of Federal Reserve speech tone on the S&P 500 daily returns with valence shifter Weight = 0.5

$R_t = a_0 + b_0 Tone_t + d * Controls + u_t$		
Variables	**All Speeches**	**FL Speeches**
NM	0.010*	
	(0.005)	
NM		0.025*
		(0.013)
Speech Controls	Yes	Yes
Macro Controls	Yes	Yes
Time Controls	Yes	Yes
Return Controls	Yes	Yes
Position FE	Yes	Yes
R Square	0.131	0.195
N	503	256

Note: This table presents the results from regressing daily index returns on speech tone (and controls) using tone as per the valence shifter weight 0.5. The results are reported in line with equation (1). The standard errors are reported in the parentheses and are all heteroskedasticity and autocorrelation (HAC) robust. The controls include three lags of return, day of the week, month dummy as well as speech level controls (average words per sentence (AWPS) percentage of complex words (Per_CW)); along with macroeconomic controls: the real exchange rate (Ex_Rate), Term Premium and the Bloomberg Economic Surprise Index (ESI) and Position fixed effects. ***, ** and * indicate that the coefficient estimate are significantly different from zero at the 1 percent, 5 percent and 10 percent levels respectively.

The valence shifter weight is chosen as 0.8 to conform with the weight suggested in literature. However, we verify that the results survive the changing of weights to 0.5 in table 9 and find that the new speech tone still significantly and positively impacts the index returns for all speeches as well as forward looking speeches.

7.2 Accounting for FOMC meetings

Table 10: Impact of Federal Reserve speech tone on the S&P500 daily returns (FOMC)

$R_t = a_0 + b_0 Tone_t + d * Controls + u_t$				
Variables	**All Speeches**		**FL Speeches**	
NM	0.007 (0.007)			
LM		0.011 (0.021)		
NM			0.028** (0.014)	
LM				0.010 (0.059)
Speech Controls	Yes	Yes	Yes	Yes
Macro Controls	Yes	Yes	Yes	Yes
Time Controls	Yes	Yes	Yes	Yes
Return Controls	Yes	Yes	Yes	Yes
Position FE	Yes	Yes	Yes	Yes
R Square	0.182	0.182	0.182	0.182
N	445	445	227	227

Note: This table presents the results from regressing daily index returns on speech tone (and controls). The pool of speeches is smaller—after excluding those which are delivered one-week prior to, or one-week post the conclusion of an FOMC meeting. The results are reported in line with equation (1). The standard errors are reported in the parentheses and are all heteroskedasticity and autocorrelation (HAC) robust. The controls include three lags of return, day of the week, month dummy as well as speech level controls (average words per sentence (AWPS) percentage of complex words (Per_CW)); along with macroeconomic controls: the real exchange rate (Ex_Rate), Term Premium and the Bloomberg Economic Surprise Index (ESI) and the speech-giver's Position fixed effects. ***, ** and * indicate that the coefficient estimate are significantly different from zero at the 1 percent, 5 percent and 10 percent levels respectively.

To separate the effect of FOMC meetings on Fed speeches and their putative impact on stock markets, we verify that the impact of Fed speeches persists even after excluding those which are delivered one-week prior to, or one-week post the conclusion of an FOMC meeting. Table 10 displays the results of this exercise which closely mirror the benchmark results. In particular, the forward-looking speeches show significantly positive impact on the S&P 500 returns on the day the speeches are delivered, while no such significance is observed for the LM dictionary based bag-of-words approach.

7.3 Investigating reverse causality: Do returns influence speeches?

Table 11: Impact of the S&P 500 daily returns on Federal Reserve speech tone

	$Tone_t = a_0 + a_n R_{t-n} + d * Controls + u_t$					
	$n = 0$	$n = 1$	$n = 2$	$n = 3$	$n = 4$	$n = 5$
NM	0.390	-0.089	-0.129	-0.349	0.176	-0.003
	(0.258)	(0.296)	(0.250)	(0.254)	(0.325)	(0.260)
Speech Controls	Yes	Yes	Yes	Yes	Yes	Yes
Time Controls	Yes	Yes	Yes	Yes	Yes	Yes
Macro Controls	Yes	Yes	Yes	Yes	Yes	Yes
Return Controls	Yes	Yes	Yes	Yes	Yes	Yes
Position FE	Yes	Yes	Yes	Yes	Yes	Yes

Note: This table presents the results from regressing speech tone on daily index returns (and controls). The results are reported in line with equation (1). The standard errors are reported in the parentheses and are all heteroskedasticity and autocorrelation (HAC) robust. The controls include three lags of return, day of the week, month dummy as well as speech level controls (average words per sentence (AWPS) percentage of complex words (Per_CW)); along with macroeconomic controls: the real exchange rate (Ex_Rate), Term Premium and the Bloomberg Economic Surprise Index (ESI) and Position fixed effects. ***, ** and * indicate that the coefficient estimate are significantly different from zero at the 1 percent, 5 percent and 10 percent levels respectively.

Since most of the Fed speeches are written in advance, and hence are not meant to be reacting to special developments in the markets, the scope of reverse causality is minimal. However, as a precautionary measure, we formally test for reverse causality by calculating the impact of the S&P 500 index returns on Fed speech tone for the next five days. The controls employed are the same as in regression specification (1), except for lags of returns. The results are presented in table 11 and we find that, as expected, the daily index return does not have any significant impact on the Fed speech tone for any lag.

8 Concluding remarks

We show that the choice of words in the Fed's Board of Governors' speeches moves the US stock market and volatility indices on the day the speeches are delivered.

Positive speeches raise returns, reduce volatility and suppress term premia. Our technique of Fed's speeches' tone quantification improves upon the current techniques of financial text analysis by offering two innovations: i) usage of the sentence as the unit of the n-gram analysis, which solves the problem of how many words to include at a time in the tone quantification procedure; and ii) usage of valence shifters, which are adjectives and adverbs which modify the meaning and tone of sentence but have been ignored so far in financial text analysis. A comparative analysis shows that the currently popular technique with the LM dictionary and bag-of-words approach fails to show any impact significantly different from 0.

We show that valence shifter usage in Fed speeches is the highest during episodes of market distress such as the Great Recession and the Eurozone debt crisis—both coinciding during the leadership of Ben Bernanke. We also find that Fed speeches use valence shifters to inject nuance into their text, and employ it to make otherwise positive speeches more positive, and negative speeches less negative. In interpreting such results however, we advocate caution, since aggregate valence shifter usage and more nuanced speech text cannot directly be attributed to the leadership of the Federal Reserve, since it also reflects the effect of prevailing uncertainties, policy preferences and market environment.

Appendices

A List of Valence Shifters

The table A1 below specifies the valence shifters encountered in the speeches analyzed in this study.

Table A1: List of Valence Shifters

Word	Classification	Weight	Word	Classification	Weight
almost	de-amplifier	0.8	not	negator	-1
although	adversative-conjuction	0.8	only	de-amplifier	0.8
barely	de-amplifier	0.8	particular	amplifier	0.8
but	adversative-conjuction	0.8	particularly	amplifier	0.8
cannot	negator	-1	partly	de-amplifier	0.8
certain	amplifier	0.8	purpose	amplifier	0.8
certainly	amplifier	0.8	quite	amplifier	0.8
colossal	amplifier	0.8	rarely	de-amplifier	0.8
considerably	amplifier	0.8	real	amplifier	0.8
deep	amplifier	0.8	really	amplifier	0.8
deeply	amplifier	0.8	seldom	de-amplifier	0.8
definitely	amplifier	0.8	serious	amplifier	0.8
dont	negator	-1	seriously	amplifier	0.8
enormous	amplifier	0.8	severe	amplifier	0.8
enormously	amplifier	0.8	severely	amplifier	0.8
especially	amplifier	0.8	significant	amplifier	0.8
extreme	amplifier	0.8	significantly	amplifier	0.8
extremely	amplifier	0.8	slightly	de-amplifier	0.8
few	de-amplifier	0.8	somewhat	de-amplifier	0.8
greatly	amplifier	0.8	sure	amplifier	0.8
hardly	de-amplifier	0.8	surely	amplifier	0.8
heavily	amplifier	0.8	totally	amplifier	0.8
heavy	amplifier	0.8	true	amplifier	0.8
high	amplifier	0.8	truly	amplifier	0.8
highly	amplifier	0.8	vast	amplifier	0.8
however	adversative-conjuction	0.8	very	amplifier	0.8
huge	amplifier	0.8	whereas	adversative-conjuction	0.8
hugely	amplifier	0.8	decidedly	amplifier	0.8
least	de-amplifier	0.8	definite	amplifier	0.8
little	de-amplifier	0.8	immense	amplifier	0.8
massive	amplifier	0.8	immensely	amplifier	0.8
massively	amplifier	0.8	incalculable	amplifier	0.8
more	amplifier	0.8	incredibly	de-amplifier	0.8
most	amplifier	0.8	sparsely	de-amplifier	0.8

Continued on next page

Table A1 – continued from previous page

Word	Classification	Weight	Word	Classification	Weight
much	amplifier	0.8	vastly	amplifier	0.8
neither	negator	-1	uber	amplifier	0.8
never	negator	-1	cant	negator	-1
majorly	amplifier	0.8	faintly	de-amplifier	0.8
none	negator	-1	wont	negator	-1

Note: This table presents the list of valence shifters along with their classification and weight.

Table A2: Impact of Federal Reserve speech tone on the DJIA daily returns

$R_t = a_0 + b_0 Tone_t + d * Controls + u_t$				
Variables	**All Speeches**		**FL Speeches**	
NM	0.009*			
	(0.005)			
LM		0.024		
		(0.017)		
NM			0.031***	
			(0.011)	
LM				0.051
				(0.040)
Speech Controls	Yes	Yes	Yes	Yes
Macro Controls	Yes	Yes	Yes	Yes
Time Controls	Yes	Yes	Yes	Yes
Return Controls	Yes	Yes	Yes	Yes
Position FE	Yes	Yes	Yes	Yes
R Square	0.117	0.115	0.193	0.175
N	503	503	256	256

Note: This table presents the results from regressing daily index returns on speech tone (and controls). The results are reported in line with equation (1). The standard errors are reported in the parentheses and are all heteroskedasticity and autocorrelation (HAC) robust. The controls include three lags of return, day of the week, month dummy as well as speech level controls (average words per sentence (AWPS) percentage of complex words (Per_CW)); along with macroeconomic controls: the real exchange rate (Ex_Rate), Term Premium and the Bloomberg Economic Surprise Index (ESI) and Position fixed effects. ***, ** and * indicate that the coefficient estimate are significantly different from zero at the 1 percent, 5 percent and 10 percent levels respectively.

References

Tobias Adrian, Richard K Crump, and Emanuel Moench. Pricing the term structure with linear regressions. *Journal of Financial Economics*, 110(1):110–138, 2013.

Abhinav Anand, Sankarshan Basu, and Jalaj Pathak. A new proxy of financial texts' readability, 2022. IIM Bangalore Research Paper No 648.

Alina Andreevskaia and Sabine Bergler. When specialists and generalists work together: Overcoming domain dependence in sentiment tagging. In *Proceedings of ACL-08: HLT*, pages 290–298, 2008.

Mikael Apel and Marianna Blix Grimaldi. The information content of central bank minutes, 2012. Working Paper Series 261, Sveriges Riksbank (Central Bank of Sweden).

Mikael Apel and Marianna Blix Grimaldi. How informative are central bank minutes? *Review of Economics*, 65(1):53–76, 2014.

Nicholas Apergis and Ioannis Pragidis. Stock price reactions to wire news from the European Central Bank: Evidence from changes in the sentiment tone and international market indexes. *International Advances in Economic Research*, 25 (1):91–112, 2019.

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

Hamza Bennani, Nicolas Fanta, Pavel Gertler, and Roman Horvath. Does central bank communication signal future monetary policy in a (post)-crisis era? the case of the ecb. *Journal of International Money and Finance*, 104:102167, 2020.

Gary C Biddle, Gilles Hilary, and Rodrigo S Verdi. How does financial reporting quality relate to investment efficiency? *Journal of Accounting and Economics*, 48(2-3):112–131, 2009.

David M Blei, Andrew Y Ng, and Michael I Jordan. Latent Dirichlet allocation. *Journal of Machine Learning Research*, 3:993–1022, 2003.

Simon Bodilsen, Jonas N Eriksen, and Niels S Grønborg. Asset pricing and FOMC press conferences. *Journal of Banking & Finance*, 128:106–163, 2021.

Benjamin Born, Michael Ehrmann, and Marcel Fratzscher. Central bank communication on financial stability. *The Economic Journal*, 124(577):701–734, 2014.

Kris Boudt and James Thewissen. Jockeying for position in ceo letters: Impression management and sentiment analytics. *Financial Management*, 48(1):77–115, 2019.

Paul Brockman, James E Cicon, Xu Li, and S McKay Price. Words versus deeds: Evidence from post-call manager trades. *Financial Management*, 46(4):965–994, 2017.

Francesca Brusa, Pavel Savor, and Mungo Wilson. One central bank to rule them all. *Review of Finance*, 24(2):263–304, 2020.

Brent Bundick, Trenton Herriford, and Andrew Smith. Forward guidance, monetary policy uncertainty, and the term premium, 2017. Research Working Paper RWP 17-7, Federal Reserve Bank of Kansas City.

Patricia A Carpenter and Marcel A Just. Sentence comprehension: a psycholinguistic processing model of verification. *Psychological review*, 82(1):45, 1975.

Ken Ramshøj Christensen. Negative and affirmative sentences increase activation in different areas in the brain. *Journal of Neurolinguistics*, 22:1–17, 2009.

Anna Cieslak and Andreas Schrimpf. Non-monetary news in central bank communication. *Journal of International Economics*, 118:293–315, 2019.

Anna Cieslak, Adair Morse, and Annette Vissing-Jorgensen. Stock returns over the FOMC cycle. *The Journal of Finance*, 74(5):2201–2248, 2019.

Selva Demiralp and Oscar Jorda. The response of term rates to Fed announcements. *Journal of Money, Credit and Banking*, pages 387–405, 2004.

Asad Dossani. Central bank tone and currency risk premia, 2018. Working Paper.

William DuBay. *Unlocking Language: The classic studies in readability*. BookSurge Publishing, 2007.

T. Philipp Dybowski and Bernd Kempa. The European Central Bank's monetary pillar after the financial crisis. *Journal of Banking & Finance*, 121(C), 2020.

Michael Ehrmann and Marcel Fratzscher. Taking stock: Monetary policy transmission to equity markets. *Journal of Money, Credit and Banking*, pages 719–737, 2004.

Michael Ehrmann and Marcel Fratzscher. The timing of central bank communication. *European Journal of Political Economy*, 23(1):124–145, 2007.

Ira Fischler, Paul A Bloom, Donald G Childers, Salim E Roucos, and Nathan W Perry Jr. Brain potentials related to stages of sentence verification. *Psychophysiology*, 20(4):400–409, 1983.

Stefan Gerlach. Interest rate setting by the ECB, 1999-2006: Words and deeds. *International Journal of Central Banking*, 3(3):1–46, 2007.

Pavel Gertler and Roman Horvath. Central bank communication and financial markets: New high-frequency evidence. *Journal of Financial Stability*, 36:336–345, 2018.

Simon Gilchrist, Vivian Yue, and Egon Zakrajšek. US monetary policy and international bond markets. *Journal of Money, Credit and Banking*, 51:127–161, 2019.

Mario Gonzalez and Raul Cruz Tadle. Monetary policy press releases: An international comparison, 2021. Working Papers Central Bank of Chile 912.

Chen Gu, Alexander Kurov, and Marketa Halova Wolfe. Relief rallies after fomc announcements as a resolution of uncertainty. *Journal of Empirical Finance*, 49:1–18, 2018.

Robert Gunning. *Technique of clear writing.* McGraw-Hill, 1952.

Graeme Guthrie and Julian Wright. Open mouth operations. *Journal of Monetary Economics*, 46(2):489–516, 2000.

Stephen Hansen and Michael McMahon. Shocking language: Understanding the macroeconomic effects of central bank communication. *Journal of International Economics*, 99:S114–S133, 2016.

Stephen Hansen, Michael McMahon, and Andrea Prat. Transparency and deliberation within the FOMC: a computational linguistics approach. *The Quarterly Journal of Economics*, 133(2):801–870, 2018.

David-Jan Jansen and Jakob De Haan. Look who's talking: ECB communication during the first years of EMU. *International Journal of Finance & Economics*, 11(3):219–228, 2006.

Alistair Kennedy and Diana Inkpen. Sentiment classification of movie reviews using contextual valence shifters. *Computational Intelligence*, 22(2):110–125, 2006.

Donald Lewis Kohn and Brian Sack. Central bank talk: does it matter and why?, 2003. Finance and Economics Discussion Series, Divisions of Research & Statistics and Monetary Affairs, Federal Reserve Board.

Matteo Leombroni, Andrea Vedolin, Gyuri Venter, and Paul Whelan. Central bank communication and the yield curve. *Journal of Financial Economics*, 141 (3):860–880, 2021.

Feng Li. Annual report readability, current earnings, and earnings persistence. *Journal of Accounting and Economics*, 45(2-3):221–247, 2008.

Tim Loughran and Bill McDonald. When is a liability not a liability? Textual analysis, dictionaries, and 10-Ks. *The Journal of Finance*, 66(1):35–65, 2011.

David O Lucca and Emanuel Moench. The pre-FOMC announcement drift. *The Journal of Finance*, 70(1):329–371, 2015.

David O Lucca and Francesco Trebbi. Measuring central bank communication: an automated approach with application to FOMC statements, 2009. National Bureau of Economic Research.

Ákos Máté, Miklós Sebők, and Tamás Barczikay. The effect of central bank communication on sovereign bond yields: The case of Hungary. *PLoS ONE*, 16(2): e0245515, 2021.

Brian P Miller. The effects of reporting complexity on small and large investor trading. *The Accounting Review*, 85(6):2107–2143, 2010.

Andreas Neuhierl and Michael Weber. Monetary policy communication, policy slope, and the stock market. *Journal of Monetary Economics*, 108:140–155, 2019.

Paolo Pasquariello. Informative trading or just costly noise? an analysis of central bank interventions. *Journal of Financial Markets*, 10(2):107–143, 2007.

James W Pennebaker, Matthias R Mehl, and Kate G Niederhoffer. Psychological aspects of natural language use: Our words, our selves. *Annual Review of Psychology*, 54(1):547–577, 2003.

Matthieu Picault and Thomas Renault. Words are not all created equal: A new measure of ECB communication. *Journal of International Money and Finance*, 79:136–156, 2017.

Livia Polanyi and Annie Zaenen. Contextual valence shifters. In *Computing Attitude and Affect in Text: Theory and Applications*, pages 1–10. Springer, 2006.

Christina D Romer and David H Romer. A new measure of monetary shocks: Derivation and implications. *American Economic Review*, 94(4):1055–1084, 2004.

Carlo Rosa. Words that shake traders: The stock market's reaction to central bank communication in real time. *Journal of Empirical Finance*, 18(5):915–934, 2011.

Pavel Savor and Mungo Wilson. How much do investors care about macroeconomic risk? Evidence from scheduled economic announcements. *Journal of Financial and Quantitative Analysis*, 48(2):343–375, 2013.

Maik Schmeling and Christian Wagner. Does central bank tone move asset prices?, 2019. Working Paper.

Marc Schulder, Michael Wiegand, Josef Ruppenhofer, and Stephanie Köser. Introducing a lexicon of verbal polarity shifters for English. In *Proceedings of the Eleventh International Conference on Language Resources and Evaluation (LREC 2018)*, Miyazaki, Japan, May 2018. European Language Resources Association (ELRA).

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

Paul C Tetlock. Giving content to investor sentiment: The role of media in the stock market. *The Journal of Finance*, 62(3):1139–1168, 2007.

Paul C Tetlock, Maytal Saar-Tsechansky, and Sofus Macskassy. More than words: Quantifying language to measure firms' fundamentals. *The Journal of Finance*, 63(3):1437–1467, 2008.

Peter Tillmann. Monetary policy uncertainty and the response of the yield curve to policy shocks. *Journal of Money, Credit and Banking*, 52(4):803–833, 2020.