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 – January 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 index shows significant statistical association with the tone of Fed speeches on the day the speech is delivered. We find that negative speeches depress returns and amplify volatility. The stock market moves more strongly in response to forward-looking speeches, and to speeches whose tone is negative. We capture financial texts' tone better than current techniques due to two innovations introduced in this study: i) a sentence-based ngram 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 (such as 'large', 'slightly', 'although' etc.) which alter the connotation of text but have been ignored in financial text analysis. We show that valence shifter usage in Fed speeches is the highest during the Great Recession and the Eurozone debt crisis.

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

JEL Classification: G14, G18, G28, G41

[*]Indian Institute of Management Bangalore, Finance & Accounting
[†]Indian Institute of Management Bangalore, Finance & Accounting
[‡]Great Lakes Institute of Management, 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 & Trebbi 2009, Hansen & McMahon 2016, Gonzalez & 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 (Neuhierl & Weber 2019).

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

The Federal Reserve (and indeed all central banks in general) intends to communicate to markets by means of forward guidance, its 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 communication in the form of speeches to its audiences 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

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

particular and to the economy in general. Hence, a study of how Fed's Board of Governors' speeches and their choice of words impact 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 & 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.[2]

From the perspective of investors as well, it is vital that their evaluation of the Federal Reserve speeches—both in content and in its intent—is accurate and is in line with the objectives of the Fed. The Fed is entrusted with conducting monetary policy, and several studies cited above show how its communication significantly impacts interest rates and inflation expectations. 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 watch 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 markets. 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. We formally evaluate the validity of such hypotheses by examining whether the tone of speeches by the Fed Board of Governors impacts movements in the US stock market indices, both in terms of returns as well as volatility. We also investigate the Board of Governors' proclivity to employ nuanced language which makes use of multi-clausal phrases and connotation-altering modifiers (e.g., adjectives, adverbs etc.) that makes the coverage of themes in their speeches (such as inflation-related, labor-market-related etc.) harder to interpret (Anand et al. 2022).

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

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 & McDonald 2011) along with the "bag-of-words" and ngram approach have been among the key tools in this field 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 & Thewissen 2019).[3] We offer improvements over the current technique of financial texts' tone quantification by proposing extensions to the "bag-of-words" and ngram approach (Tetlock 2007, Li 2008, Tetlock et al. 2008) and introduce 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 & 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 & Inkpen 2006, Polanyi & Zaenen 2006, Schulder et al. 2018) but have not been granted any weight in the LM dictionary.[4]

   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.

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

[3] In the process of tone quantification "bag-of-words" refers to using one word at a time whereas ngram refers to using a cluster of $n \geq 2$ words at a time.

[4] The full list of valence shifters used in the speeches is presented in the appendices in table A.1.

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 financial dictionaries such as the LM dictionary (Loughran & McDonald 2011). In fact, according to the bag-of-words with LM dictionary approach, all the above sentences are assigned a tone of 0. Not only does it ignore the valence shifters, it fails to include the connotation of the word 'decrease'.[5] However, our approach correctly classifies each sentence's tone by recognizing the connotation-altering role of valence shifters and including 'decrease in losses' as the relevant 3-gram at the sentence level.

Further, in order to extract the tone from Fed speeches, in addition to the usage of the LM dictionary for accurately quantifying the tone of financial text, we also use the dictionary specified in Apel & Grimaldi (2014), Apergis & Pragidis (2019) which characterizes text with respect to central bank communication.[6] 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, ngrams, 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 & 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 ngram and assigns polarity to it by consulting an augmented LM dictionary with valence shifters. Somewhat relatedly, Pennebaker et al. (2003) argue that the entire corpus of text and individual sentences within

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

[6]According to Picault & Renault (2017) there is a significant correlation between the tones calculated from the two dictionaries.

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 of sentence-level ngram 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 find that the highest overall proportion of valence shifters in speeches occurs during the tenure of Ben Bernanke, especially during the Great Recession (2007-08). This is intriguing since, in particular, high frequency of negators have been shown to be related to more difficulty in processing (Christensen 2009), and high valence shifter usage has been shown to exaggerate positive and hide negative news from investors by US firms (Anand et al. 2022). Insofar as usage of valence shifters introduces lexical/semantic complexity into speeches, thereby making it harder to interpret, heightened usage of such language by the Fed Board of Governors during recessions suggests an injection of obfuscation.

We subject the Fed's Board of Governors' speeches to our modified tone extraction process and evaluate whether the tone of speeches impacts movements in the US stock market indices, both in terms of returns as well as in terms of volatility. We show that Fed speeches—especially those that are forward-looking—impact the returns of US stock indices positively—both for the S&P 500 and the DJIA—on the same day as the speech is delivered, implying that (all else equal) positive speeches are associated with increased returns and negative speeches with decreased returns. We also show that Fed speeches impact the volatility of the US stock markets negatively—both for the S&P 500 volatility and the VIX—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. Another finding is that the US stock markets react much more strongly to negative speeches—which form a vast majority of total speeches—than to positive speeches. 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.

We also present detailed stock market intraday evidence—based on 30 minute interval returns, as well as on 30 minute intraday changes in VIX—on how the effect of Fed speeches percolates down to the changes in benchmark index levels and volatilities. We show that the markets react to Fed speeches during the later intervals of the day on which the speech is delivered, and during the earlier intervals of the next day. This is quite reasonable since if the speech is delivered on say, 2 PM, the markets will respond to its contents only after 2 PM on the same day and during the early trading hours over the next day.

Finally, we also demonstrate that our new method provides associative significance with both index returns as well as volatilities even after controlling for the presence of speech tone from the LM dictionary based bag-of-words approach. This indicates that our measure has significance over and above that supplied by speech tone quantification from the current method. We rule out reverse association and show that past stock returns have no significant influence on the tone U.S. Federal Reserve speeches. We also allay concerns regarding endogeneity by isolating forward-looking speeches and by showing that their impact on the stock markets is even higher as compared to the full sample.

The paper is organized as follows, section 2 is the literature review for Federal Reserve communication in particular, and central bank communication in general. 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. 6 outlines the impact of Fed speeches on US index returns and volatility—on a daily, as well as intraday basis. 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 set of studies have investigated their impact. For example, among the earliest such studies, Guthrie & Wright (2000) investigate how central bank statement—rather than open market operations—can be used to implement monetary policy in New Zealand. Romer & 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 & 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 & 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 U.S. intraday stock and volatility indices and reports significant results. Savor & Wilson (2013) check whether investors care about macroeconomic announcements and find that the average market return and Sharpe ratio are significantly higher on important announcement days. Lucca & 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 post 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 & Sack (2003) analyze central bank communication using dummy categorization of the content and find that it significantly impacts the interest rates. Lucca & Trebbi (2009) use Google search and Factiva based news articles in an ngram approach to analyze FOMC announcements and find that they are significantly associated with treasury yields. Hansen & 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 & 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 & 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 & Renault (2017) use ngram 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 & Wagner (2019) and Apergis & 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.

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. After downloading the speeches, 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 & McDonald 2011) and the phrases are extracted according to Apel & Blix Grimaldi (2012) and Apergis & 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 ngram 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 approxi-

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

mately 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 ngrams or bag-of-words with LM dictionary 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 ngram analysis (n words at a time) where n changes from sentence to sentence.

In addition to the sentence based ngram 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 & Inkpen (2006), Polanyi & 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, table A.2 in the appendices presents 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 'va-

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

lence 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 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 ngram 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.

Table 2 displays the sentence connotation according to our sentence-level ngram with valence shifters and augmented LM dictionary approach.

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:

> *"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:

Table 1: Examples of usage of valence shifters in Federal Reserve speeches

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".

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

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

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 as well as intraday movements in 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:

$$R_t = a_0 + b_n Tone_{t-n} + \sum_{i=1}^{3} c_i R_{t-i} + d_1 * Controls_t + \\ d_2 * SpeechControls_t + d_3 * MacroControls_t + \gamma_t \quad (1)$$

$$Vol_t = a_0 + b_n Tone_{t-n} + \sum_{i=1}^{3} c_i Vol_{t-i} + d_1 * Controls_t +$$
$$d_2 * SpeechControls_t + d_3 * MacroControls_t + \gamma_t \quad (2)$$

n ranges from 0 to 5. Controls include the day of the week and month dummy; as well as average words per sentence (awps) and percentage of complex words (per_CW) as speech level controls. The two variables of speech controls (awps and per_CW) 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 & Fratzscher 2007, Born et al. 2014, Gertler & Horvath 2018).

In addition, we include macroeconomic variables as control factors. These include the real exchange rate and the Bloomberg Economic Surprise Index (ESI). 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.

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); intraday data on index returns are taken from TAQ; 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]

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

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 text-related 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 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, about 38% of sentences contain at least one valence shifter. Out of the collection of valence shifters used in Fed speeches, the highest proportion is 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.

Insert table 2 about here.

To further examine the difference between tone calculated using the methodology introduced in this study and the LM dictionary based bag-of-words approach, we plot the tone of sentences containing valence shifters calculated by both methods. We present boxplots of speech tones under the two methods and compare their salient features in figure 1. Clearly, the tone of speeches under the new method-

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

ology shows a wider range than that in the old methodology, and the speech tone median under the new technique assumes a higher value than its counterpart.

Insert figure 1 about here.

Table 3 presents the difference in speech tone statistics for sentences with valence shifters, using the existing/old methodology (bag-of-words and LM dictionary) and the new methodology introduced in this study. From figure 1, as well as from table 3, it is clear 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.

Insert table 3 about here.

Taken together, this suggests that the full variability of speech tones is systematically underestimated when valence shifters are ignored, as is done in the 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 1.68×10^{-9}. Thus the distance between the two speech tone distributions (D statistic) is indeed significantly different for the two methodologies, with corresponding p-value being indistinguishable from 0 upto 8 decimal places.

5.2 Fed's valence shifter usage across time

Insert figure 2 about here.

Figure 2 presents the time series barplots of the overall percentage of valence shifters as well as its four components over the years. The highest use, overall, is that of amplifiers. Interestingly, the highest overall proportion of valence shifters (∼60%) occurs during the Great Recession and the Eurozone debt crisis. Insofar as usage of valence shifters introduces lexical/semantic complexity into speeches, thereby making text harder to interpret (Anand et al. 2022) heightened usage of such language by the Fed Board of Governors during recessions suggests an injection of obfuscation. In particular, negators are harder to interpret (Carpenter & Just 1975, Fischler et al. 1983, Christensen 2009), and usage of such language has been shown by firms to exaggerate positive developments and understate the negative in the MD&A section (Anand et al. 2022).

Insert figure 3 about here.

Similarly, figure 3 presents the distribution of valence shifters in 'hawkish', 'dovish', positive and negative speeches. The classification of hawkish and dovish speeches is done in accordance with Apel & Grimaldi (2014) and Malmendier et al. (2021). To identify a speech as dovish/hawkish we match the list of "goals" (Apel & Grimaldi 2014) with dovish and hawkish words.[11] Each matched occurrence is marked as −1 for dovish, and +1 for hawkish. We identify the counterparts within a sentence instead of a five gram, as advocated in Malmendier et al. (2021) since sentence-based ngrams are more accurate. The net dove-hawk index is defined as the difference in dovish and hawkish score for each speech. If the overall score is negative (positive) it is classified as "dovish" (hawkish).

$$\text{Net Dove-Hawk Index} := \frac{Hawkish}{(Hawkish + Dovish)} - \frac{Dovish}{(Dovish + Hawkish)}$$

Figure 3 presents the boxplots of valence shifter proportion based on the categorizations: hawkish, dovish, positive and negative respectively. As can be observed, for hawkish and negative speeches, the proportion of valence shifters used is much higher as compared to dovish and positive speeches.

[11]The following terms are categorized as 'goals': "inflation", "cyclical position", "growth","price", "wages", "oil prices", "development" and "unemployment". Dovish terms consist of the following: "decrease", "slow", "weak", "low". Hawkish terms use the following: "increase", "fast", "strong" and "high".

5.3 Valence shifter usage across Fed Chairpersons

Our sample duration comprises three regimes based on the tenure of the Federal Reserve's Board of Governors' Chairpersons: Ben Bernanke, Janet Yellen and Jerome Powell. We enquire if different chairpersons had different predilection for the usage of valence shifters in their speeches. Figure 4 presents the boxplots of valence shifter usage in speeches of the three chairpersons.

Insert figure 4 about here.

Visually inspecting figure 4 suggests that overall, valence shifter usage has fallen over time. Median valence shifter usage proportion is the highest for Bernanke (2006–2013), followed by that for Janet Yellen, and Jerome Powell. For the year 2007, we observe the maximum usage of valence shifters in a speech; and for the year 2008, we note the maximum median usage of valence shifters for speeches given in a year—both of which fall under Bernanke's regime. Both years also fall under the shadow of the Great Recession. Yellen's speeches show an equally high proportion of valence shifter usage in the year 2014, but her speeches, especially in the years 2015–2017, show much lower usage, while for Jerome Powell, the frequency is even rarer.

Insert table 4 about here.

Table 4 presents results for the difference in valence shifter usage across hawkish and dovish speeches, as well as that across different Fed Chairpersons. We use the T test for differences in means, and the Kolmogorov-Smirnov test for differences in distribution, and find that valence shifter usage—across both speech type (hawkish/dovish), as well as across Chairpersons—shows statistically significant differences in all cases.

We investigate further, the choice of three major themes—labour market, inflation and output—across the three Fed Chairs: Bernanke, Yellen, and Powell in figure 5. To identify the words and phrases related to the three themes, we first identify terms related to eight sub-themes: 'employment', 'unemployment', 'growth', 'demand', 'output', 'development', 'inflation' and 'price'. Further, for each sub-theme we calculate the proportion of words as a percentage of the total. Then, we average the percentages for relevant sub themes (such as employment and unemployment for "Labour Market") to get the percentage for each theme.

Insert figure 5 about here.

As visual inspection of figure 5 shows, Ben Bernanke's speeches featured a stable usage of the three major themes, with the least frequently mentioned theme being labor-market related. In contrast, Janet Yellen's speeches show a very wide variation across the themes: labor market, inflation and output. Her 2014 speeches show hardly any mention of the labor markets, but it commands a health proportion of her speeches in 2015, after reaching a peak ($\sim 15\%$) in 2016 only to plummet sharply in 2017—her last year. Inflation appears to be her favorite theme and dominates all others, peaking at 25% in 2015. Output-related themes are covered at the expense of labor-related themes across her speeches. For Jerome Powell, only two years' set of speeches are not enough to establish any patterns yet, except the order of importance: inflation, followed by output, and finally by labor market. To see if valence shifter usage varies across the three major themes, we present figure 6. The main observation is that from 2010–2017—across both Bernanke's and Yellen's tenure—the usage of valence shifters was the most frequent, by a large margin, in particular, for themes related to inflation. The proportion of valence shifters is the maximum in 2012 and 2015 for inflation dominated speeches.

Insert figure 6 about here.

5.4 Speech Statistics

We first provide basic descriptive statistics for the frequency of speeches, the number of words contained therein; the daily index return statistics; and for the speech tone calculated according to the methodology specified in this study. Table 5 presents the summary statistics.

Overall, there are 797 speeches in our sample, with an average of 4.1 speeches delivered per month. A large majority of speeches (547) have an overall negative tone. The average speech contains 3482 words; the longest speech has 10923 words;[12] while the shortest contains 237 words.

In keeping with the preponderance of speeches with a negative tone, the mean speech tone of our sample is -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.

Further, as table 5 shows, the daily return of the two benchmark stock indices of the US: the S&P 500 and the DJIA display a mean daily return of 0.023% and have 251 trading days per year. Their values for the maximum, minimum and the standard deviation of daily returns are also extremely close suggesting a very high correlation between the two indices.

Insert table 5 about here.

Figure 8 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. The reason for choosing to display monthly movements in the two time series is due to their amenability for easy visual inspection. 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.

Insert figure 8 about here.

Similarly figure 9 presents the time series of monthly speech tone and the Fed Funds rate on the primary and secondary y-axes respectively. As the figure demonstrates, there is significant comovement between the two time series. Broadly speaking there are three regimes from 2006–2020: 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.

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

Insert figure 9 about here.

Together, figures 8 and 9 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 examine this putative relationship more comprehensively.

6 Impact on the US stock market

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. In order to add weight to our investigation, we conduct this analysis on both the daily as well as the intraday levels.

6.1 Impact on the S&P 500 daily returns

We examine the impact of speech tone on both daily returns of the benchmark S&P 500 index as well as the 2 day open and close return. Table 6 presents the results of both specifications—daily as well as 2 day open and close return in line with regression specification (1). The methodology is that of ordinary least squares with heteroskedasticity and autocorrelation consistent (HAC) errors.[13]

Insert table 6 about here.

The main finding is that the Federal Reserve speech tone significantly impacts the daily US stock index returns contemporaneously i.e., on the same day as the speech is delivered. Further, the coefficient estimate is positive (0.011) which implies that (all else equal) speeches with negative tones lead to a fall in the daily index return; and those with positive tones lead to a rise in the daily index return. Based on our result, a one standard deviation change in the Federal Reserve speech

[13]All standard errors reported in this study are HAC robust.

tone is associated with 0.07 standard deviation change in daily market return for the S&P 500 index.[14]

Since the timestamp of the speeches are not available, it could be possible that some speeches are delivered after trading hours, or during the fag end of trading hours. Presumably, in such cases, the impact could be observed on the day of the speech as well as the next day after the speech has been delivered. To account for such possibilities, we calculate 2 day returns for the S&P 500 index using the opening price of the day on which the speech is delivered and the closing price of the day after the speech-day. The speech tone in this case is also significantly associated with the 2 day return on the next day after the speech is delivered. In the table, we have defined interval 0 as the period for which return is calculated using the opening price on day 0 (speech day) and the closing price on day 1; interval 1 as the period with return calculated from the opening price on day 1 and closing price on day 2 and so on.

6.2 Impact of forward-looking speeches

As specified in Ehrmann & Fratzscher (2007), it is important to examine forward-looking statements with respect to central bank communication, since central banks mostly use it 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 & Horvath 2018). Hence we consider the subsample of US Fed speeches which feature a higher-than-average proportion of terms and words associated with forward-looking statements and examine their tone's impact on the index return.

To identify forward-looking communication, we look for specific words and phrases which are generally used to convey pre-meditated plans and actions. These include "believe", "estimate", "anticipate", "plan", "predict", "hope", "seek", "expect", "likely", "intend", "potential", "is likely to", "with the intent" etc. We

[14] All results in this study remain consistent and broadly similar in terms of impact even if only speeches by Governor/Chairperson and the Deputy Governor/Vice Chairperson are considered (approximately 90% of the sample).

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. Thus 376 speeches are identified from our initial sample of 797 as forward-looking.

Table 7 presents the impact of forward-looking speech tone on the S&P 500 Index returns—both daily, as well as the 2 day return. The results are quite similar to table 6 in that the estimated coefficients are significantly positive. However, the magnitude of the coefficient as well as its economic significance is much higher for the subsample of forward-looking speeches. A one standard deviation movement in speech tone is associated with 0.18 standard deviation movement in daily index return, which is 2.5 times the impact observed in table 6. Further, for the two-day return, the impact is observed on both days—the speech-day as well as on the day following.

Insert table 7 about here.

6.3 Impact on daily volatility

Apart from analyzing the impact of Fed speeches on daily benchmark returns, 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 i) daily realized volatility of the S&P 500 index, and ii) daily changes in the Chicago Board Options Exchange's (CBOE) Volatility Index (VIX) in line with the regression specification in equation (2). We calculate the daily realized volatility by demeaning the squared residual returns and then calculating the mean of the demeaned residual over five days in line with Tetlock (2007).

Insert table 8 about here.

Table 8 presents our results on the effect of speech tone on the daily realized volatility of the S&P 500 index as well as VIX. Our main result is that the Federal Reserve speech tone significantly impacts the daily US stock index realized volatility contemporaneously i.e., on the same day as the speech is delivered. Further, the coefficient estimate is negative (-0.0002) which implies that speeches with negative tones are associated with a rise in daily volatility; and those with positive tones are associated with a fall in daily volatility.

In line with our previous result on daily realized volatility we find that Fed speeches significantly impact changes in the daily VIX contemporaneously with a negative regression estimate (-0.083) implying that speeches with a positive tone reduce changes in VIX while those with negative tones amplify it. We also note that the size of the coefficient for VIX is much higher than that for the daily realized volatility for the S&P 500 suggesting a more powerful impact of the speeches on VIX.

6.4 Impact on the S&P 500 intraday returns

In order to examine in greater detail how the impact of Fed speeches percolates down to the changes in benchmark index levels, we resort to an intraday analysis where we investigate the effect of speech tones on 30 minute interval returns of the S&P 500 index as well as VIX. The results of our examination are presented in table 9. Since our dataset on Fed speeches does not carry a time-stamp, we examine its effect on the market on both the day the speech was delivered (Day 0), as well as on the next day (Day 1). This is because, if the speech was delivered, for example, at 2 PM, the markets will be able to react to its content only after 2 PM on the same day and during the early hours of trading on the subsequent day. This is exactly what we observe: the Fed speeches display an impact on day 0 intraday returns at intervals 4, 8 and 12 with a positive sign; and on the next day (day 1) at intervals 1 (positive sign) and 9 (negative sign). The preponderance of significant intraday impact with positive signs suggests a reason why the contemporaneous impact of speeches on daily returns also retains the same sign.

Similar to the analysis of speech impact on intraday returns, we follow the effect of speech tones on intraday changes in the VIX calculated at 30 minute intervals. Again, a priori we expect that the effect (if any) will be more pronounced on the later intervals of day 0 and on the earlier intervals on day 1.

As expected, the Fed speeches exhibit an effect on intraday VIX changes on day 0 at intervals 4 and 12 with a negative sign; and on day 1 at intervals 1 and 4 with a negative sign. Again, the preponderance of significant intraday impact with negative signs suggests why the contemporaneous impact of speeches on daily changes in VIX also retains the same sign.

Insert table 9 about here.

6.5 Impact of positive vs. negative speeches on S&P 500 intraday returns

As described in table 5, the tone of a large majority of speeches is negative. In order to examine if the intraday impact of speeches with a positive tone is significantly different from those with a negative tone, we introduce a dummy variable in equation (1) which assumes a value 1 in case the tone is positive and 0 otherwise. We also add an interaction term of the dummy with the speech tone to capture interaction effects. The results are presented in table 10. Consistent with prior results, the speech tone displays positive significance on day 0 and 1 for intervals 4, 8 and 12; and intervals 1 and 4 respectively. The speech tone dummy assumes significantly negative values on day 0 and 1 for intervals 8; and intervals 4 and 9 respectively.[15] Further, the interaction term displays negative significance on day 0 for interval 10. Hence the results imply that speeches with a positive tone impact US stock returns much more weakly than those with negative tones; or in other words, the stock markets react much more strongly to negative Fed speeches than to positive speeches.[16]

Insert table 10 about here.

6.6 Impact based on topic and content

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

[15] For interval 11 on day 0 the speech tone dummy reports a significantly positive value.

[16] Similar results are observed in case of daily return. However, they are not presented for brevity.

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 & 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.[17] 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 11. The results are quite similar to our benchmark results in table 6 with the speech tone being significantly associated with the S&P 500 returns on the same day as the speech is delivered. However, the economic significance of the results for such a subsample of speeches is much higher, with the coefficient on the day 0 speech tone assuming a value 2.8 times as compared to the one in table 6.

Insert table 11 about here.

6.7 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

[17] 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".

the US term premium and the results are presented in table 12. 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.[18] We find that the Fed speech tone impacts the US term premium significantly and the coefficient is negative. Thus, all else equal, positive speech tones are associated with a fall in US term premium. This result holds not just for the 2-year, but also for the 5-year, 7-year and 10-year term premium. This is expected as per Bundick et al. (2017), where they specify that a positive Fed outlook leads to a fall in economic uncertainty and thus a fall in term premium.

Insert table 12 about here.

7 Robustness

For robustness we examine the other most important US stock index—the DJIA—and subject its daily as well as intraday returns to the same analysis as that for the S&P 500.

In addition, we also test whether the speech tone from our new methodology remains statistically significant in the presence of the speech tone from the existing, LM based methodology.

7.1 Impact on DJIA

In table 13 we presents results for the regression in which speech tone of the Federal Reserve is the independent variable and the DJIA daily index return is the dependent variable. The controls include the day of the week and month dummy, three lags of daily return, along with speech level controls—average words/sentence and the percentage of complex words—as well as macroeconomic controls in line with the specification in equation (1).

Insert table 13 about here.

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

The table indicates that the results are almost identical to that with the S&P 500 index returns outlined in table 6. The Fed speech tone impacts daily DJIA return contemporaneously with the coefficient estimate (0.011) displaying a positive sign which signifies that speech tone and daily DJIA returns co-move in the same direction.

Table 14 presents results for intraday 30-minute interval returns for the DJIA index. Again, the results are almost identical to those obtained for the intraday 30-minute interval returns for the S&P 500 index in table 6.

Insert table 14 about here.

The Federal Reserve speech tone impacts the intraday DJIA returns on day 0 at intervals 4, 8 and 12—all positively—and on day 1 at intervals 1 (positive) and 9 (negative). Again the overall positive impact of speech tone on intraday DJIA returns suggest why contemporaneous daily DJIA return impact retains its positive sign. Overall, it is not surprising that both daily and intraday DJIA returns exhibit an almost identical effect to that of the S&P 500 since their summary statistics are so closely associated with each other, as shown in table 5.

We also conduct both daily and intraday analysis of the impact of Fed speeches on DJIA volatility and note that due to the very closely associated nature of returns in both indices, the results are quite similar to the previous set of findings for the S&P 500. We also examine the impact of forward-looking speech tone on DJIA return and find almost identical results. For brevity, however, we do not display the full tables on volatility and forward-looking speech tone analysis.

7.2 Investigating reverse causality: Do returns influence speeches?

Although subsample analysis with forward-looking speeches avoids the endogeneity problem, as further precaution, we formally test for reverse causality by calculating the impact of the S&P 500 index returns on Fed speech tones. The controls employed are the same as in regression specification (1), except that in place of lags of returns we use lags of speech tone as control. The results are presented in table 15 and we find that the daily index return does not have any significant impact on the Fed speech tone for any lag.

Insert table 15 about here.

7.3 Comparison with the LM dictionary and bag-of-words approach

How much does our new methodology contribute towards explaining index returns over-and-above the impact of the existing LM dictionary based bag-of-words approach? We address this major concern in this subsection. In order to facilitate such a comparison, we augment the benchmark regression specification in equation (1) by adding the speech tone from the LM dictionary based bag-of-words approach as an additional control. Further, to allay concerns of multicollinearity arising due to the introduction of two speech tones with potentially high correlation, we conduct ridge regression and present the results in table 16.

Our main finding is that even in the presence of the speech tone using the existing methodology (EM) the tone extracted from our technique (NM) retains its significance and displays a significantly positive coefficient one day after the delivery of the speech. For all other lags—from day 0 to day 5—there is no significance for either of the two speech tones. In particular, EM displays no further associative significance at any lag in the presence of our new tone extraction methodology. We obtain the same result with the DJIA index but do not report the full details for brevity.

Insert table 16 about here.

8 Conclusion

Our study improves upon the current techniques of financial text analysis by offering two innovations: i) usage of the sentence as the unit of the ngram 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. Our application of this new methodology to the quantification of the impact of Fed speeches on the US stock markets indicate that the speeches impact stock market returns and volatility on the same day as they are delivered; that negative speeches have more impact than positive speeches; and that the US stock market reacts more strongly to forward-looking speeches. We also show that during crises, the Fed Chair's speeches feature an unusually high usage of valence shifters which make speeches harder to interpret.

Tables and Figures

8.1 Tables

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	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	“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	-0.06	-0.009	“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	“but” discounts the impact of “strong” thus decreasing the impact of the phrase “the exchange rate remains strong”.

Table 3: 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 calculated using the new methodology and the LM dictionary based "bag-of-words" approach (existing methodology).

Table 4: Difference in mean valence shifter usage

	p-value T-test	p-value KS test	D-Statistic KS test
Hawkish vs Dovish	0.000001	0.00009	0.22
Positve vs Negative	0.0008	0.002	0.26
Bernanke vs Yellen	0.07	0.07	0.23
Yellen vs Powell	0.01	0.05	0.33
Bernanke vs Powell	2.5e-06	8.96e-05	0.47

Note: This table presents the p value for T test for difference between mean values and the D statistic and p value for the difference between distributions using the Kolmogorov-Smirnov test.

Table 5: Speech Statistics

Variable/Time Period	Max	Min	Mean	SD
Speech Words	10923	237	3482	1662.56
Speech Tone	0.2949	-0.3403	-0.0605	0.0864
S&P 500 Index	11.5800	−11.9840	0.0227	1.2574
DJIA Index	11.3650	−12.9265	0.0233	1.2089

Note: This table presents the summary statistics for the number of words in the speeches, tone of the speeches as well as the Index daily return (%).

Table 6: Impact of Federal Reserve speech tone on the S&P 500 daily returns

	$R_t = a_0 + a_n Tone_{t-n} + d * Controls + u_t$					
	$n = 0$	$n = 1$	$n = 2$	$n = 3$	$n = 4$	$n = 5$
S&P 500 return	0.011* (0.006)	0.003 (0.006)	0.001 (0.006)	−0.001 (0.006)	0.0002 (0.007)	0.002 (0.006)
	Interval 0	Interval 1	Interval 2	Interval 3	Interval 4	Interval 5
S&P 500 Index (2 day return)	0.006 (0.008)	0.012* (0.007)	−0.001 (0.007)	0.002 (0.007)	0.002 (0.007)	0.003 (0.006)
Speech Controls	Yes	Yes	Yes	Yes	Yes	Yes
Controls	Yes	Yes	Yes	Yes	Yes	Yes
Macro Controls	Yes	Yes	Yes	Yes	Yes	Yes

Note: This table presents the results from regressing daily and 2 day 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 and percentage of complex words); along with macroeconomic controls: the real exchange rate and the Bloomberg Economic Surprise Index. ***, ** and * indicate that the coefficient estimate are significantly different from zero at the 1 percent, 5 percent and 10 percent levels respectively.

Table 7: Impact of Federal Reserve's forward-looking speech tone on the daily return

	$R_t = a_0 + a_n Tone_{t-n} + d * Controls + u_t$					
	$n = 0$	$n = 1$	$n = 2$	$n = 3$	$n = 4$	$n = 5$
S&P 500 Index	0.028** (0.014)	0.007 (0.009)	0.014 (0.011)	−0.002 (0.010)	−0.005 (0.011)	−0.003 (0.011)
	Interval 0	Interval 1	Interval 2	Interval 3	Interval 4	Interval 5
S&P 500 Index (2 day return)	0.027* (0.016)	0.024* (0.014)	0.008 (0.013)	0.010 (0.013)	0.005 (0.012)	−0.004 (0.012)
Speech Controls	Yes	Yes	Yes	Yes	Yes	Yes
Controls	Yes	Yes	Yes	Yes	Yes	Yes
Macro Controls	Yes	Yes	Yes	Yes	Yes	Yes

Note: This table presents the results from regressing daily and 2 day 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 and percentage of complex words); along with macroeconomic controls: the real exchange rate and the Bloomberg Economic Surprise Index. ***, ** and * indicate that the coefficient estimate are significantly different from zero at the 1 percent, 5 percent and 10 percent levels respectively.

Table 8: Impact of Federal Reserve speech tone on the S&P 500 daily realized volatility

	$Vol_t = a_0 + a_n Tone_{t-n} + d * Controls + u_t$					
	$n = 0$	$n = 1$	$n = 2$	$n = 3$	$n = 4$	$n = 5$
S&P 500 Vol	-0.0002^{*}	-0.00006	-0.0001	0.00009	0.0002	-0.0003^{*}
	(0.0001)	(0.0002)	(0.0002)	(0.0001)	(0.0001)	(0.0001)
VIX	-0.083^{*}	-0.032	-0.037	-0.046	-0.018	-0.004
	(0.045)	(0.043)	(0.044)	(0.041)	(0.049)	(0.042)
Speech Controls	Yes	Yes	Yes	Yes	Yes	Yes
Controls	Yes	Yes	Yes	Yes	Yes	Yes
Macro Controls	Yes	Yes	Yes	Yes	Yes	Yes

Note: This table presents the results from regressing daily index realized volatility and VIX on speech tone (and controls). We calculate the daily realized volatility by demeaning the squared residual returns and then calculating the mean of the demeaned residual over five days (Tetlock 2007). 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 the realized volatility, day of the week, month dummy as well as speech level controls (average words per sentence and percentage of complex words) along with macroeconomic controls: real exchange rate and Bloomberg Economic Surprise Index. ***, ** and * indicate that the coefficient estimate are significantly different from zero at the 1 percent, 5 percent and 10 percent levels respectively.

Table 9: Impact on the intraday S&P 500 intraday returns and VIX

	Intraday Impact			
	S&P 500 Index Return		VIX	
Interval	Day 0	Day 1	Day 0	Day 1
Interval 1	0.006	0.012***	−0.026	−0.044**
	(0.003)	(0.004)	(0.026)	(0.023)
Interval 2	−0.001	−0.0007	−0.00007	0.002
	(0.001)	(0.002)	(0.009)	(0.012)
Interval 3	−0.0004	−0.001	0.003	0.006
	(0.001)	(0.001)	(0.008)	(0.008)
Interval 4	0.003***	0.002	−0.019**	−0.014**
	(0.001)	(0.001)	(0.007)	(0.007)
Interval 5	0.00003	−0.0008	−0.007	0.008
	(0.001)	(0.001)	(0.006)	(0.008)
Interval 6	−0.0005	0.0001	0.002	−0.006
	(0.001)	(0.001)	(0.007)	(0.005)
Interval 7	0.0005	−0.0007	−0.004	0.001
	(0.001)	(0.001)	(0.006)	(0.006)
Interval 8	0.002**	0.001	−0.003	−0.00008
	(0.001)	(0.001)	(0.006)	(0.006)
Interval 9	0.0001	−0.002*	−0.002	−0.0006
	(0.001)	(0.001)	(0.006)	(0.006)
Interval 10	−0.0005	0.0004	0.004	−0.001
	(0.001)	(0.001)	(0.007)	(0.006)
Interval 11	−0.0007	−0.001	0.007	−0.003
	(0.001)	(0.001)	(0.007)	(0.008)
Interval 12	0.002*	−0.002	−0.013*	0.010
	(0.001)	(0.001)	(0.007)	(0.009)
Interval 13	−0.0007	−0.0007	−0.004	0.0007
	(0.002)	(0.002)	(0.012)	(0.011)
Interval 14	0.0006	−0.0002	0.006	−0.008
	(0.0005)	(0.0002)	(0.009)	(0.010)
Controls	Yes	Yes	Yes	Yes
Macro Controls	Yes	Yes	Yes	Yes
Speech Controls	Yes	Yes	Yes	Yes

Note: This table presents the results from the regression on daily central bank speech tone. The dependent variable is the intraday 30 min index returns in line with equation (1). The one difference as compared to the daily regression is that three lags of intraday 30 minute return are not kept as additional controls. The standard errors are reported in the parentheses and are all Heteroskedasticity and Autocorrelation (HAC) robust.

Table 10: Impact of positive vs. negative speeches on intraday S&P 500 returns

Return Interval	**Day 0**			**Day 1**		
Interval	Speech Tone	Tone Dummy	Interaction Term	Speech Tone	Tone Dummy	Interaction Term
Interval 1	0.010 (0.005)	−0.0009 (0.001)	−0.006 (0.013)	0.016*** (0.006)	−0.0002 (0.001)	−0.020 (0.013)
Interval 2	0.0009 (0.002)	−0.0003 (0.0006)	−0.004 (0.007)	−0.004 (0.002)	0.0004 (0.0005)	0.008 (0.007)
Interval 3	−0.001 (0.002)	0.0005 (0.0004)	−0.003 (0.006)	−0.002 (0.002)	−0.00002 (0.0003)	0.006 (0.004)
Interval 4	0.003** (0.001)	0.00007 (0.0003)	0.0006 (0.004)	0.004** (0.001)	−0.0006* (0.0004)	−0.0008 (0.004)
Interval 5	0.001 (0.001)	−0.0001 (0.0003)	−0.004 (0.005)	−0.0009 (0.002)	0.0003 (0.0003)	−0.004 (0.004)
Interval 6	0.00002 (0.001)	−0.0003 (0.0002)	0.002 (0.003)	0.001 (0.002)	0.00004 (0.0003)	−0.005 (0.003)
Interval 7	−0.0003 (0.001)	−0.0001 (0.0003)	0.005 (0.003)	−0.001 (0.001)	−0.00001 (0.0003)	0.001 (0.003)
Interval 8	0.004*** (0.001)	−0.0006* (0.0003)	0.001 (0.004)	0.001 (0.001)	−0.0001 (0.0003)	−0.0008 (0.004)
Interval 9	0.002 (0.001)	−0.0002 (0.0003)	−0.004 (0.004)	−0.001 (0.001)	−0.0004* (0.0002)	0.004 (0.003)
Interval 10	0.0002 (0.001)	0.0004 (0.0004)	−0.009* (0.004)	0.001 (0.002)	0.00006 (0.0003)	−0.004 (0.003)
Interval 11	−0.001 (0.002)	0.0007* (0.0003)	−0.007 (0.005)	−0.001 (0.002)	−0.00007 (0.0003)	0.004 (0.004)
Interval 12	0.004* (0.002)	0.00005 (0.0004)	−0.007 (0.004)	−0.002 (0.002)	−0.0004 (0.0004)	0.005 (0.005)
Interval 13	0.001 (0.004)	−0.0002 (0.0005)	−0.006 (0.007)	−0.001 (0.003)	0.0005 (0.0004)	−0.003 (0.005)
Interval 14	0.00003 (0.0004)	0.0003 (0.0002)	−0.002 (0.001)	−0.0001 (0.0002)	−0.00004 (0.00008)	0.0003 (0.0006)
Controls	Yes	Yes	Yes	Yes	Yes	Yes
Macro Controls	Yes	Yes	Yes	Yes	Yes	Yes
Speech Controls	Yes	Yes	Yes	Yes	Yes	Yes

Note: This table presents the results from the regression of the S&P 500 index's intraday 30 min returns on the Federal Reserve speech tone in line with equation (1) with the addition of a dummy for positive speech tone and an interaction term of speech tone and the dummy variable as additional explanatory variables. The one difference as compared to the daily regression is that three lags of intraday 30 minute returns are not kept as additional controls. The standard errors are reported in the parentheses and are all Heteroskedasticity and Autocorrelation (HAC) robust. The controls include the day of the week, month dummy as well as speech level controls (average words per sentence and percentage of complex words) along with macroeconomic controls: real exchange rate and the Bloomberg Economic Surprise Index. ***, ** and * indicate that the coefficient estimate are significantly different from zero at the 1 percent, 5 percent and 10 percent levels respectively.

Table 11: Impact on S&P daily returns based on topic analysis

	$R_t = a_0 + a_n Tone_{t-n} + d * Controls + u_t$					
	$n = 0$	$n = 1$	$n = 2$	$n = 3$	$n = 4$	$n = 5$
S&P500 Index	0.028^{**} (0.013)	0.018 (0.012)	0.005 (0.011)	-0.008 (0.013)	0.0008 (0.012)	-0.0005 (0.012)
Controls	Yes	Yes	Yes	Yes	Yes	Yes
Macro Controls	Yes	Yes	Yes	Yes	Yes	Yes
Speech Controls	Yes	Yes	Yes	Yes	Yes	Yes

Note: This table presents the results from regressing S&P daily index returns on speech tone (and controls). The speech sample consists of speeches that feature terms prominently associated with risk premia in the financial markets. 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 and percentage of complex words) along with macroeconomic controls: real exchange rate and Bloomberg Economic Surprise Index. ***, ** and * indicate that the coefficient estimate are significantly different from zero at the 1 percent, 5 percent and 10 percent levels respectively.

Table 12: U.S. Term Premium

Term	2 year	5 year	7 year	10 year	Controls	Speech Controls	Fixed Effects	Macro Controls
TP	-0.838^{***} (0.154)	-1.760^{***} (0.302)	-2.201^{***} (0.374)	-2.645^{***} (0.452)	Yes	Yes	Yes	Yes

Note: This table presents the results from regression on yield component (Term Premium). The coefficients are reported for the impact of the U.S. Federal Reserve speech tone. The standard errors are reported in the parentheses and are all Heteroskedasticity and Autocorrelation (HAC) robust. The controls include day of the week, month dummy as well as speech level controls (average words per sentence and percentage of complex words) along with macro controls (real exchange rate and Bloomberg Surprise Index) and 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 13: Impact on DJIA daily returns

	$R_t = a_0 + a_n Tone_{t-n} + d * Controls + u_t$					
	$n = 0$	$n = 1$	$n = 2$	$n = 3$	$n = 4$	$n = 5$
DJIA	0.011^{**} (0.006)	0.004 (0.005)	0.001 (0.005)	-0.0003 (0.005)	-0.0008 (0.006)	0.001 (0.005)
Controls	Yes	Yes	Yes	Yes	Yes	Yes
Macro Controls	Yes	Yes	Yes	Yes	Yes	Yes
Speech Controls	Yes	Yes	Yes	Yes	Yes	Yes

Note: This table presents the results from regressing daily DJIA 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 and percentage of complex words) along with macroeconomic controls: real exchange rate and Bloomberg Economic Surprise Index. ***, ** and * indicate that the coefficient estimate are significantly different from zero at the 1 percent, 5 percent and 10 percent levels respectively.

Table 14: Impact on DJIA intraday returns

Interval	Day 0	Day 1
Interval 1	0.004	0.012***
	(0.003)	(0.004)
Interval 2	−0.0004	−0.001
	(0.001)	(0.001)
Interval 3	−0.0006	−0.001
	(0.001)	(0.001)
Interval 4	0.003***	0.001
	(0.001)	(0.001)
Interval 5	0.00002	−0.0009
	(0.001)	(0.001)
Interval 6	−0.0002	−0.00005
	(0.001)	(0.001)
Interval 7	0.00004	−0.0007
	(0.001)	(0.001)
Interval 8	0.002**	0.001
	(0.001)	(0.001)
Interval 9	0.0003	−0.002**
	(0.001)	(0.001)
Interval 10	−0.0003	0.0002
	(0.001)	(0.001)
Interval 11	−0.0004	−0.0008
	(0.001)	(0.001)
Interval 12	0.002*	−0.002
	(0.001)	(0.001)
Interval 13	−0.0002	−0.0009
	(0.002)	(0.002)
Interval 14	−0.0001	−0.0002
	(0.0002)	(0.0002)
Controls	Yes	Yes
Macro Controls	Yes	Yes
Speech Controls	Yes	Yes

Note: For this table, the dependent variable is the DJIA intraday 30 min index returns. The results are reported in line with equation (1). The one difference is that three lags of intraday 30 minute returns are not kept as additional controls. The standard errors are reported in the parentheses and are all Heteroskedasticity and Autocorrelation (HAC) robust. The controls include the day of the week, month dummy as well as speech level controls (average words per sentence and percentage of complex words) along with macroeconomic controls: real exchange rate and Bloomberg Economic Surprise Index. ***, ** and * indicate that the coefficient estimate are significantly different from zero at the 1 percent, 5 percent and 10 percent levels respectively.

Table 15: Impact of the S&P 500 daily returns on Federal Reserve speech tone

	$Tone_t = a_0 + a_n R_{t-n} + d * Controls + u_t$					
	$n = 0$	$n = 1$	$n = 2$	$n = 3$	$n = 4$	$n = 5$
S &P Return	0.319	−0.196	−1.501	0.219	0.198	−3.206
	(0.609)	(0.790)	(3.839)	(0.818)	(1.626)	(2.916)
Speech Controls	Yes	Yes	Yes	Yes	Yes	Yes
Controls	Yes	Yes	Yes	Yes	Yes	Yes
Macro Controls	Yes	Yes	Yes	Yes	Yes	Yes

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

Table 16: Impact on daily S&P 500 returns with current methodology tone (EM) as an additional control

	Day 0		Day 1		Day 2		Day 3		Day 4		Day 5	
Methodology	NM	EM	NM	EM	NM	EM	NM	EM	NM	EM	NM	EM
S&P500	0.002 (0.002)	0.001 (0.002)	0.015** (0.006)	−0.007 (0.006)	−0.00004 (0.003)	0.001 (0.003)	0.0009 (0.002)	0.001 (0.002)	−0.000003 (0.000004)	−0.000000 (0.000004)	−0.0007 (0.002)	0.001 (0.002)
Speech Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Macro Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Note: This table presents the results from performing ridge regression on daily S&P 500 index returns on speech tone calculated using both the methodology specified in this study (NM) and the existing "bag-of-words" LM dictionary approach (EM). The results are reported in line with equation (1) with EM as an additional control. 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 and percentage of complex words) along with macroeconomic controls: real exchange rate and Bloomberg Economic Surprise Index. ***, ** and * indicate that the coefficient estimate are significantly different from zero at the 1 percent, 5 percent and 10 percent levels respectively.

8.2 Figures

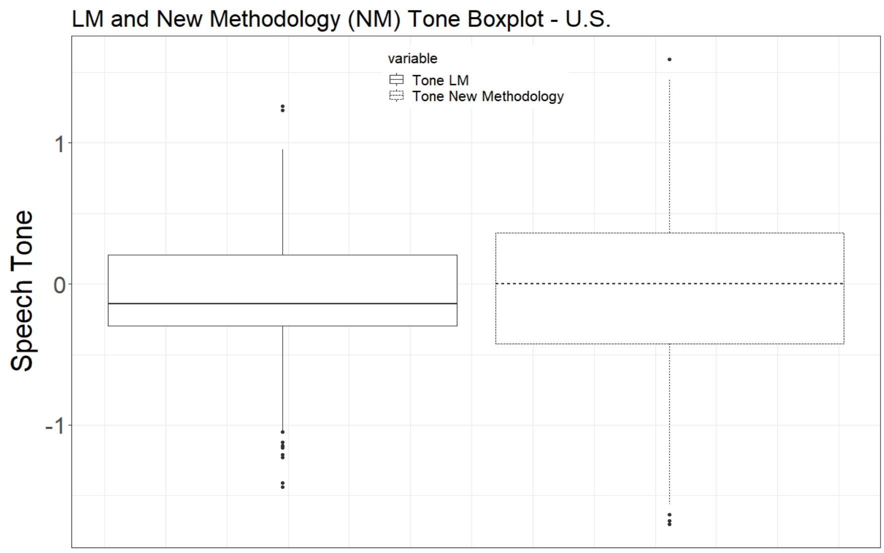


Figure 1: Boxplots of speech tones under the old and new methodologies. The LM tone (solid line) is speech tone calculated using the "bag-of-words" approach and the LM dictionary whereas the new methodology tone (dotted line) is the tone calculated by the methodology specified in this study

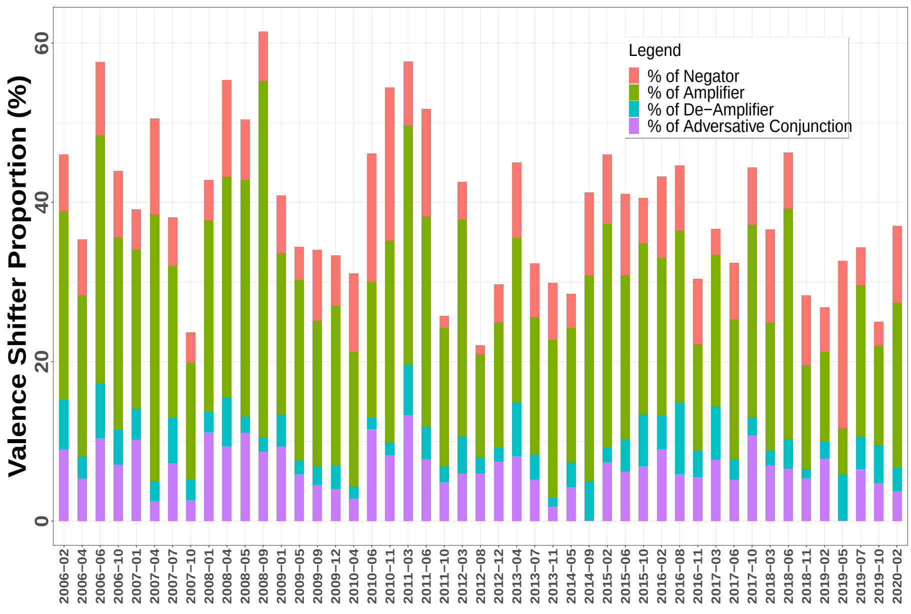


Figure 2: The figure presents the movement of overall percentage of valence shifters in speech and its respective four components (negator, amplifier, de-amplifier and adversative conjunction across time.

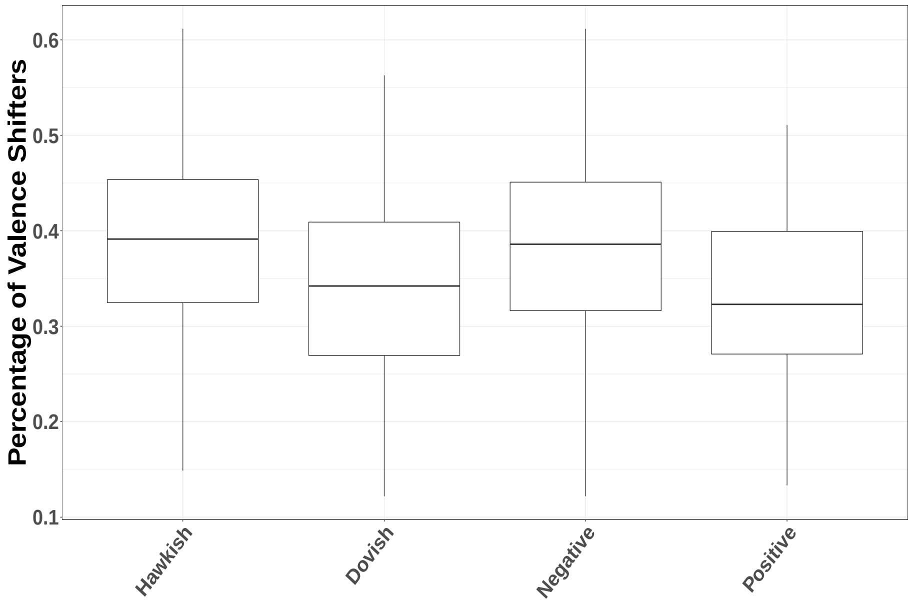


Figure 3: The figure presents the movement of overall percentage of valence shifters in speech for Hawkish, Dovish, Positive and Negative Speeches respectively.

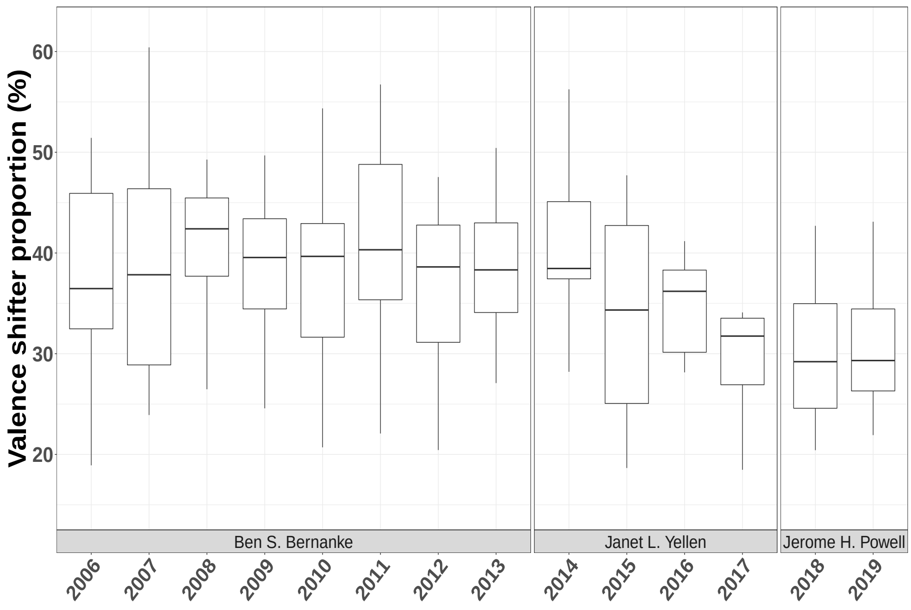


Figure 4: The boxplot above presents the time series of the proportion of valence shifter usage in Fed speeches of Chairpersons over 2006 to 2019. The three sections present the tenure of Ben Bernanke, Janet Yellen and Jerome Powell respectively.

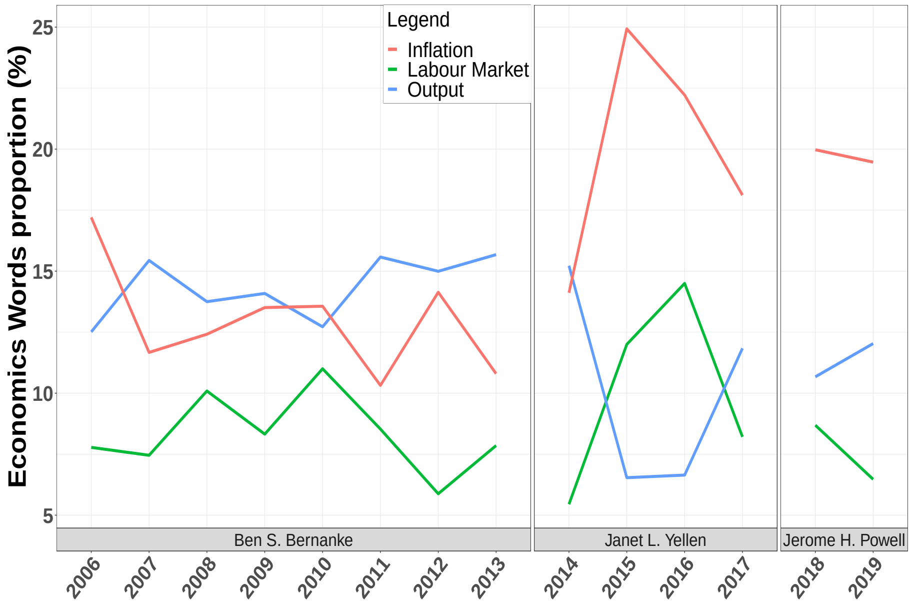


Figure 5: The figure above presents the time series of the proportion of words used in three major economic themes by Chairpersons over 2006 to 2019. The three sections present the tenure of Ben Bernanke, Janet Yellen and Jerome Powell respectively.

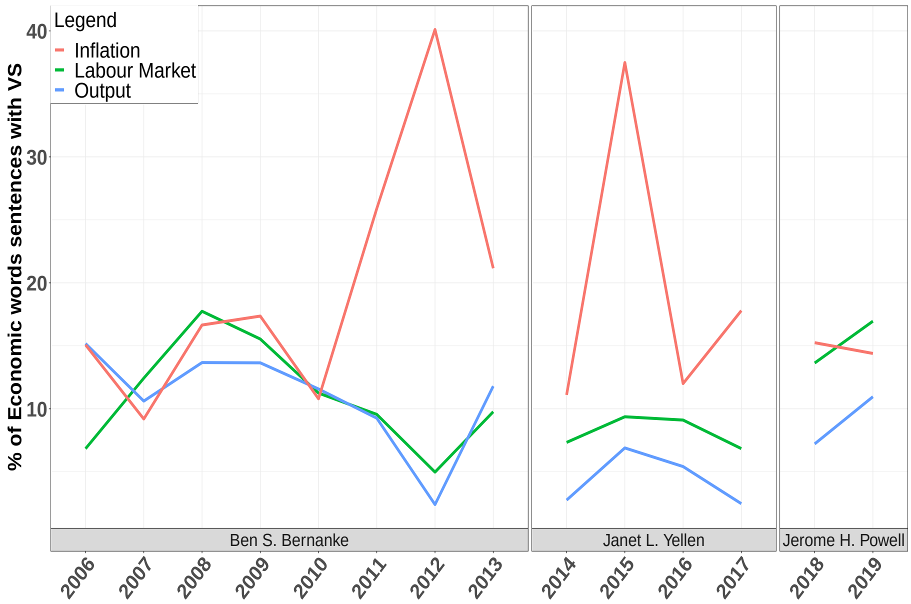
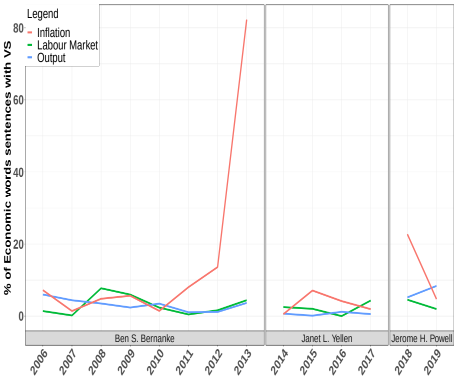
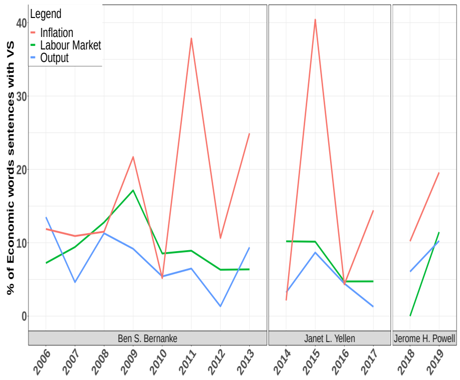


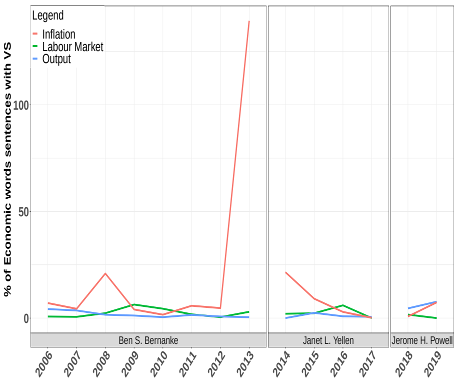
Figure 6: The figure above presents the time series of the proportion of words used in three major economic themes by Chairpersons over 2006 to 2019. The three sections present the tenure of Ben Bernanke, Janet Yellen and Jerome Powell respectively.



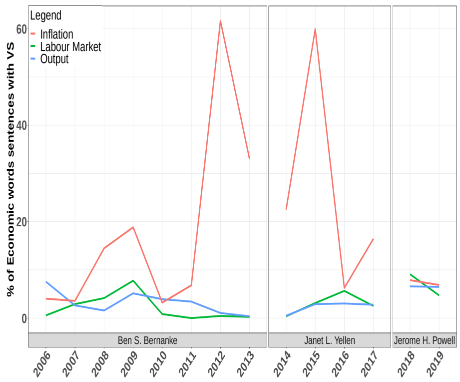
(a) Negator



(b) Amplifier



(c) De-Amplifier



(d) Adversative Conjunction

Figure 7: Graphs above present the movement distributions of four type of valence shifters as a percentage for the sentences containing economic words for Chairpersons over 2006 to 2019. The three sections present the tenure of Ben Bernanke, Janet Yellen and Jerome Powell respectively.

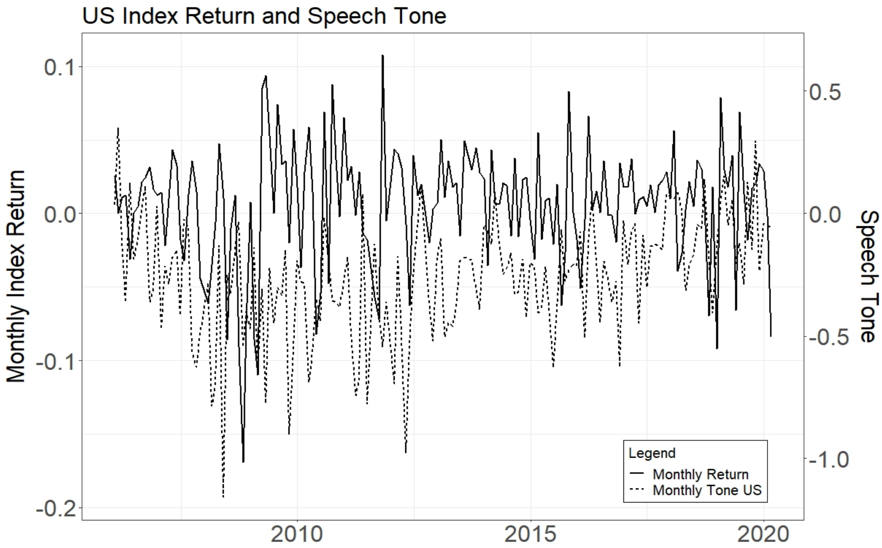


Figure 8: The time series of monthly S&P 500 index returns on the primary y axis; and the monthly speech tone on the secondary y-axis.

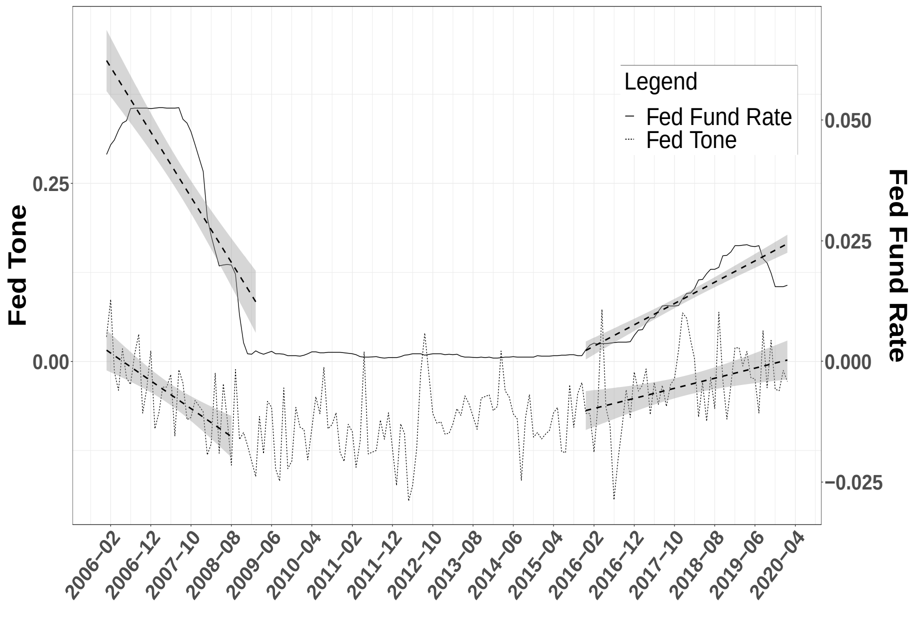


Figure 9: The time series of monthly Fed Tone on the primary y axis; and the monthly Fed Rate on the secondary y-axis.

A List of Valence Shifters

The table A.1 below specifies the valence shifters encountered in the speeches analyzed in this study.

Table A.1: 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

Continued on next page

Table A.1 – continued from previous page

Word	Classification	Weight	Word	Classification	Weight
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
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 A.2: Impact on S&P 500 daily returns with valence shifter weight = 0.5

Index	Day 0	Day 1	Day 2	Day 3	Day 4	Day 5
S&P 500	0.011*	0.007	−0.0007	0.0009	−0.002	0.0001
	(0.006)	(0.006)	(0.006)	(0.005)	(0.007)	(0.005)
Controls	Yes	Yes	Yes	Yes	Yes	Yes
Macro Controls	Yes	Yes	Yes	Yes	Yes	Yes
Speech Controls	Yes	Yes	Yes	Yes	Yes	Yes

Note: This table presents the results from regressing daily index returns on speech tone (and controls). The valence shifter weight is 0.5 as compared to 0.8 in the earlier table 6. 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 and percentage of complex words) along with macroeconomic controls: real exchange rate and Bloomberg Economic Surprise Index. ***, ** and * indicate that the coefficient estimate are significantly different from zero at the 1 percent, 5 percent and 10 percent levels respectively.

References

Adrian, T., Crump, R. K. & Moench, E. (2013), 'Pricing the term structure with linear regressions', *Journal of Financial Economics* **110**(1), 110–138.

Anand, A., Basu, S. & Pathak, J. (2022), 'A new proxy of financial texts' readability'. IIM Bangalore Research Paper No 648.

Andreevskaia, A. & Bergler, S. (2008), When specialists and generalists work together: Overcoming domain dependence in sentiment tagging, *in* 'Proceedings of ACL-08: HLT', pp. 290–298.

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

Apel, M. & Grimaldi, M. B. (2014), 'How informative are central bank minutes?', *Review of Economics* **65**(1), 53–76.

Apergis, N. & Pragidis, I. (2019), '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.

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

Bennani, H., Fanta, N., Gertler, P. & Horvath, R. (2020), '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.

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

Blei, D. M., Ng, A. Y. & Jordan, M. I. (2003), 'Latent Dirichlet allocation', *Journal of Machine Learning Research* **3**, 993–1022.

Bodilsen, S., Eriksen, J. N. & Grønborg, N. S. (2021), 'Asset pricing and FOMC press conferences', *Journal of Banking & Finance* **128**, 106–163.

Born, B., Ehrmann, M. & Fratzscher, M. (2014), 'Central bank communication on financial stability', *The Economic Journal* **124**(577), 701–734.

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

Brockman, P., Cicon, J. E., Li, X. & Price, S. M. (2017), 'Words versus deeds: Evidence from post-call manager trades', *Financial Management* **46**(4), 965–994.

Brusa, F., Savor, P. & Wilson, M. (2020), 'One central bank to rule them all', *Review of Finance* **24**(2), 263–304.

Bundick, B., Herriford, T. & Smith, A. (2017), 'Forward guidance, monetary policy uncertainty, and the term premium'. Research Working Paper RWP 17-7, Federal Reserve Bank of Kansas City.

Carpenter, P. A. & Just, M. A. (1975), 'Sentence comprehension: a psycholinguistic processing model of verification.', *Psychological review* **82**(1), 45.

Christensen, K. R. (2009), 'Negative and affirmative sentences increase activation in different areas in the brain', *Journal of Neurolinguistics* **22**, 1–17.

Cieslak, A., Morse, A. & Vissing-Jorgensen, A. (2019), 'Stock returns over the FOMC cycle', *The Journal of Finance* **74**(5), 2201–2248.

Cieslak, A. & Schrimpf, A. (2019), 'Non-monetary news in central bank communication', *Journal of International Economics* **118**, 293–315.

Demiralp, S. & Jorda, O. (2004), 'The response of term rates to Fed announcements', *Journal of Money, Credit and Banking* pp. 387–405.

Dossani, A. (2018), 'Central bank tone and currency risk premia'. Working Paper.

DuBay, W. (2007), *Unlocking Language: The classic studies in readability*, BookSurge Publishing.

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

Ehrmann, M. & Fratzscher, M. (2004), 'Taking stock: Monetary policy transmission to equity markets', *Journal of Money, Credit and Banking* pp. 719–737.

Ehrmann, M. & Fratzscher, M. (2007), 'The timing of central bank communication', *European Journal of Political Economy* **23**(1), 124–145.

Fischler, I., Bloom, P. A., Childers, D. G., Roucos, S. E. & Perry Jr, N. W. (1983), 'Brain potentials related to stages of sentence verification', *Psychophysiology* **20**(4), 400–409.

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

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

Gilchrist, S., Yue, V. & Zakrajšek, E. (2019), 'US monetary policy and international bond markets', *Journal of Money, Credit and Banking* **51**, 127–161.

Gonzalez, M. & Tadle, R. C. (2021), 'Monetary policy press releases: An international comparison'. Working Papers Central Bank of Chile 912.

Gu, C., Kurov, A. & Wolfe, M. H. (2018), 'Relief rallies after fomc announcements as a resolution of uncertainty', *Journal of Empirical Finance* **49**, 1–18.

Gunning, R. (1952), *Technique of clear writing*, McGraw-Hill.

Guthrie, G. & Wright, J. (2000), 'Open mouth operations', *Journal of Monetary Economics* **46**(2), 489–516.

Hansen, S. & McMahon, M. (2016), 'Shocking language: Understanding the macroeconomic effects of central bank communication', *Journal of International Economics* **99**, S114–S133.

Hansen, S., McMahon, M. & Prat, A. (2018), 'Transparency and deliberation within the FOMC: a computational linguistics approach', *The Quarterly Journal of Economics* **133**(2), 801–870.

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

Kennedy, A. & Inkpen, D. (2006), 'Sentiment classification of movie reviews using contextual valence shifters', *Computational Intelligence* **22**(2), 110–125.

Kohn, D. L. & Sack, B. (2003), 'Central bank talk: does it matter and why?'. Finance and Economics Discussion Series, Divisions of Research & Statistics and Monetary Affairs, Federal Reserve Board.

Leombroni, M., Vedolin, A., Venter, G. & Whelan, P. (2021), 'Central bank communication and the yield curve', *Journal of Financial Economics* **141**(3), 860–880.

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

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

Lucca, D. O. & Moench, E. (2015), 'The pre-FOMC announcement drift', *The Journal of Finance* **70**(1), 329–371.

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

Malmendier, U., Nagel, S. & Yan, Z. (2021), 'The making of hawks and doves', *Journal of Monetary Economics* **117**, 19–42.

Máté, Á., Sebők, M. & Barczikay, T. (2021), 'The effect of central bank communication on sovereign bond yields: The case of Hungary', *PLoS ONE* **16**(2), e0245515.

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

Neuhierl, A. & Weber, M. (2019), 'Monetary policy communication, policy slope, and the stock market', *Journal of Monetary Economics* **108**, 140–155.

Pennebaker, J. W., Mehl, M. R. & Niederhoffer, K. G. (2003), 'Psychological aspects of natural language use: Our words, our selves', *Annual Review of Psychology* **54**(1), 547–577.

Picault, M. & Renault, T. (2017), 'Words are not all created equal: A new measure of ECB communication', *Journal of International Money and Finance* **79**, 136–156.

Polanyi, L. & Zaenen, A. (2006), Contextual valence shifters, *in* 'Computing Attitude and Affect in Text: Theory and Applications', Springer, pp. 1–10.

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

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

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

Schmeling, M. & Wagner, C. (2019), 'Does central bank tone move asset prices?'. Working Paper.

Schulder, M., Wiegand, M., Ruppenhofer, J. & Köser, S. (2018), Introducing a lexicon of verbal polarity shifters for English, *in* 'Proceedings of the Eleventh International Conference on Language Resources and Evaluation (LREC 2018)', European Language Resources Association (ELRA), Miyazaki, Japan.

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

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

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

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