IIMB-WP N0. 622/2021

WORKING PAPER NO: 622

Who Moved The Market? Analyzing The Role Of Federal Reserve Speeches

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
Doctoral student
Finance & Accounting
Indian Institute of Management Bangalore
Bannerghatta Road, Bangalore – 5600 76
jalaj.pathak17@iimb.ac.in

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

Year of Publication – October 2021

Who Moved The Market? Analyzing The Role Of Federal Reserve Speeches

Abhinav Anand*, Sankarshan Basu†, Jalaj Pathak‡
Ashok Thampy§

Abstract

We show that the tone of speeches delivered by Federal Reserve's senior functionaries is significantly associated with movements in the US stock markets and term premiums. Negative (positive) speeches are associated with depressed (elevated) returns, amplified (lowered) volatilities and higher term premium spreads. The US stock market reacts strongly to forward-looking Fed speeches, and the impact of negative speeches is higher than that of positive speeches. To establish our results, we introduce a new method of tone quantification for Federal Reserve speeches which additionally incorporates i) "valence shifters": adjectives and adverbs (such as "massive", "although", "faintly" etc.) which alter the tone of speeches but have been given no weights in the current methods, and ii) ngrams derived from using the sentence as a unit of analysis which enables our technique to quantify the tone of multi-clausal phrases (e.g., "slight slowdown in productivity") more accurately than is currently possible.

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

JEL Classification: G14, G18, G28, G41

*IIM Bangalore
†IIM Bangalore
‡Great Lakes Institute of Management
§IIM Bangalore

"Monetary policy is 98% talk and only 2% action."

Former Fed chairman Ben Bernanke on the show '60 Minutes'.[1]

1 Introduction

Due to its prime importance in conducting monetary policy and maintaining financial stability, all aspects of Federal Reserve communication are watched very closely by market participants. Although a large collection of papers have been published analyzing the impact of press releases and FOMC statements [Lucca and Trebbi, 2009, Hansen and McMahon, 2016, Gonzalez and Tadle, 2021], we examine a very important yet understudied tool in the Federal Reserve communication toolkit: the role of speeches delivered by senior functionaries of the Federal Reserve [Neuhierl and 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 and Sack, 2003, Demiralp and Jorda, 2004, Lucca and Trebbi, 2009, Smales and Apergis, 2017]; money supply [Gerlach, 2007]; currency markets [Dossani, 2018] as well as stock return and volatility [Ehrmann and Fratzscher, 2004, Apergis and Pragidis, 2019, Brusa et al., 2020, Bodilsen et al., 2021]. According to Schmeling and Wagner [2019], central bank communication impacts market expectations and thus can be associated with market return which could be due to the information content of the communication impacting asset prices [Savor and 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 senior Fed functionaries, it is essential that such communication in the form of speeches to its audiences is conveyed accurately and is interpreted in the manner envisaged

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

by the Fed. Any miscommunication or misinterpretation in this regard can prove quite costly to the financial markets in particular and to the economy in general. Hence, a study of how Federal Reserve speeches 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 and Wright [2000], any evidence which connects the impact of Fed speeches to movements in the markets helps the central banker in gauging whether it is successfully conveying its message.[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 Fed officials extremely closely.

Hence, it is quite likely that speeches by Fed officials 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 content of such hypotheses by examining whether the tone of speeches by the Federal Reserve impacts movements in the US stock market indices, both in terms of returns as well as volatility.

In order to analyze the tone of speeches delivered by functionaries of the Federal Reserve, we subject the speeches to financial text analysis. The

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

Loughran and McDonald dictionary [Loughran and McDonald, 2011] along with the "bag-of-words" and ngram approach have been among the key tools in this field.[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 and Bergler, 2008], and

2. by using "valence shifters"—adjectives and adverbs such as "but", "large", "barely" etc.—which modify the meaning of the sentence [Kennedy and Inkpen, 2006, Polanyi and Zaenen, 2006, Schulder et al., 2018] but have not been granted any weight in the LM dictionary.[4]

To the best of our knowledge this is the first instance of the usage of valence shifting in financial text analysis. This process of tone quantification, along with using the sentence as a base unit of measurement is, in our opinion, a noteworthy advance in the framework of tone quantification of text.

The LM dictionary list as well as ngram phrases/word list (derived from the speech text itself) have certain shortcomings. For example, the LM polar words' list does not characterize certain words, such as "increase" since their meaning is dependent upon the noun form they are used with. For example, "increase" used with "unemployment" has a negative connotation while its usage with "growth" carries a positive connotation. The positive/negative categorization of such verb-noun combinations is addressed by the usage of ngram polar phrases. For example, taking two words at time ($n = 2$) we

[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.

can correctly assign a positive value to "increased growth"; and a negative value to "increased unemployment". On the other hand, creating new dictionaries from content derived word lists (such as ngram phrases) is ill-advised according to Schmeling and Wagner [2019] since it can lead to hindsight bias because the same data are used twice: first to build a dictionary and subsequently to analyze the tone and its impact.

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 and Grimaldi [2014] which characterizes text with respect to central bank communication.[5] Further, we extract polar phrases in line with Apergis and Pragidis [2019] 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 and Renault, 2017].

We divide a speech into a set of sentences and extract the tone for each sentence considering both the polar words/phrases (negative/positive) as well as the adverbs and adjectives (valence shifters) surrounding the polar words/phrases. The valence shifters themselves can be divided into four categories: adversative conjunction (e.g. "although", "however"), negator (e.g. "nor", "not"), amplifier (e.g. "very") and de-amplifier (e.g. "few") which alter the tone of the sentence. For example, the following sentence is taken from the speech of Stanley Fisher, Vice Chair Federal Reserve, on 19-05-2016:

> *"high inflation can destroy an economy and result in enormous hardship for everyone involved."*

[5] According to Picault and Renault [2017] there is a significant correlation between the tones calculated from the two dictionaries.

The tone using LM dictionary and "bag-of-words" approach is -0.22, whereas using our modified tone extraction method it is -0.28, since the negative effect of the word "hardship" is accentuated by the the presence of the amplifier "enormous".[6] The capability of an adjective—such as 'enormous' in this case—to alter the content of the sentence is overlooked by the LM dictionary, which leads to an inaccurate tone calculation.

We subject Federal Reserve 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

[6] We present more examples of differences in tone quantification using the current and the new methodology in table 2.

the speech is delivered on say, 1 PM, the markets will respond to its contents only after 1 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.

The paper is organized as follows, section 2 is the literature review for central bank communication as well as that for text analysis in finance, section 3 specifies the methodology for tone calculation followed by section 4 which describes the data sources. Section 5 is for results and analysis followed by section 6 for robustness analysis and finally, section 7 offers concluding remarks.

2 Literature review

Due to the perceived economic and financial importance of central banks, a diverse group of studies have investigated their impact in a variety of ways. For example, Guthrie and Wright [2000] study how central bank statement—rather than open market operations—can be used to implement monetary policy in New Zealand. Romer and Romer [2004] analyze central bank communication using subjective assessment of the content and examine its impact on monetary policy. Bhattarai and Neely [2016] provide a literature review of the analysis and impact of monetary policy. Kawamura et al. [2019] examine the content of Bank of Japan's monthly report and find that ambiguous expressions tend to appear more frequently with negative expressions and this leads to obfuscation of important information. 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. Dybowski and Kempa [2020] use a topic modelling approach to examine the stance of the European Central Bank communication and find that the ECB has shifted its focus from monetary policy towards the stability of European financial system. Correa et al. [2021] examine the relationship between the tone of the financial stability reports of the central banks and the financial cycle. They report that the movements in the financial cycle indicators—credit, asset prices, systemic risk, and monetary policy rates are significantly associated with the financial stability tone. Gonzalez and Tadle [2021] find that the press releases of most central banks converge during periods of international crises. Most recently, Nyman et al. [2021] use algorithmic analysis to classify the Bank of England's internal commentary into two emotional traits: "excitement" and "anxiety" and find that the traits are significantly associated with economic variables: "production", "employment", as well as the FTSE Index.

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 on stock returns and volatility

Savor and 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 and Moench [2015] report large average excess pre-FOMC returns on the US equities but no impact on treasury bills. Picault and 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 and Wagner [2019] and Apergis and Pragidis [2019] also quantify central bank tone and find that it is significantly associated with both return and volatility. Cieslak and Schrimpf [2019] find that the non-monetary component accounts for more than half the central bank communication and is significantly associated with treasury yields. 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. Brusa et al. [2020] investigate the impact of FOMC meetings on global equities and report significant results. Bodilsen et al. [2021] also report that FOMC meetings followed by press conferences are significantly associated with stock return.

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

Kohn and Sack [2003] analyze central bank communication using dummy categorization of the content and find that it significantly impacts the interest rates. Jansen and De Haan [2006] also study 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. Similarly, 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. Lucca and 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 and McMahon [2016] use a topic analysis approach on FOMC communication to analyze its impact on the market using a FAVAR framework and report significant association with treasury yields but not on real economic variables. On similar lines, Smales and Apergis [2017] extract the readability of monetary policy statements using the Flesch-Kincaid index and present its impact on 10 year T-bills. Dossani [2018] examines how the central bank press conferences impact risk premia in the currency markets and finds significant results. Neuhierl and Weber [2019] analyze how the tone of speeches by FOMC members correlates with the Fed funds futures. 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. Baranowski et al. [2021] use survey data to analyze the expectation channel of monetary policy and find that the impact differs for interest rate and inflation expectations. Most recently, Dossani [2021] examines the impact of central bank press conferences on the currency market and reports that hawkishness of a press conference is associated with a decline in in the variance risk premium and the subsequent decline in the variance swap return. On similar lines, Consoli et al. [2021] use the news indicators from the database Global Database of Events, Language and Tone and report that negative macroeconomic news indicators are significantly associated with the change in bond yields.

3 Methodology

3.1 Tone Quantification

We calculate the tone for each speech by classifying it as a collection of sentences. 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. After downloading the speeches, the content is parsed and converted to all lower cases. We remove references (if any) from the content and then 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. A complete speech is thus broken down into a collection of sentences. For each sentence, words are classified into two categories: valence shifters (adjectives and adverbs) and polar words/phrases (positive/negative tone words/phrases).

Polar words are taken from the LM dictionary [Loughran and McDonald, 2011] and the phrases are extracted according to Apel and Blix Grimaldi [2012] and Apergis and Pragidis [2019]. The phrases comprise two parts: a verb and a noun. The nouns are taken from the Economist's "Economics Dictionary"[7] and the verb list includes all the verb forms not classified in

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

the LM dictionary, such as "increase", "decrease", "reduce", "fall", "raise" etc. These phrases of verb-noun combinations are identified as ngram units ($2 \leq n \leq 5$) and are categorized as either positive or negative. Thus phrases with a noun and a verb form such as "raise growth", and "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 further ensure that there is no duplication of polar words/phrases between the LM dictionary and the ngram classification. We find that approximately 51% of sentences contain one or more polar words from the LM dictionary and an additional 14% sentences contain one or more of the macro-related nouns and verbs not classified in the LM dictionary. Thus, by using both the LM dictionary and verb-noun combinations, a more extensive portion of speeches can be quantified as compared to using just ngrams or 'bag-of-words' (unigram/LM dictionary).

We add another dimension to the aforementioned technique of using the sentence as the unit of analysis combined with the verb-noun phrase combination. This aspect relates to the usage of adjective and adverbs ("valence shifters") which modify the meaning of the sentence. The valence shifters and their respective weights are taken from Kennedy and Inkpen [2006], Polanyi and Zaenen [2006] and Schulder et al. [2018]. These valence shifters are further classified into four categories: amplifiers ("absolutely", "acutely", "very"), de-amplifiers ("barely", "faintly", "few"), negators ("not", "cannot") and adversative conjunction ("despite", "but"). The amplifiers, de-amplifiers, and adversative conjunction are given a weight of 0.8: positive for an amplifier, negative for a de-amplifier, negative for the words before adversative conjunction; and positive for the words after adversative conjunction. This is so because adversative conjunction such as "but" will amplify the argument after it and weight down the argument 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

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

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 'valence shifters' is quite distinct since it is used for characterizing adverbs and adjectives that modify the meaning of sentences and is adapted for use from the communications and computational linguistics literature. 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.

Thus, 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.

We show that, in comparison to the new process and updated dictionary, the existing LM dictionary and "bag-of-words" approach can lead to incorrect quantification of tone. An example is presented below from 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" approach and existing dictionary (LM) 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 three type of valence shifters, except for adversative conjunctions (such as "but", "although" etc.) since these words modify the meaning of the text both before and after them. Thus each sentence is divided into clusters with respect to polar words/phrases. In terms of the speech fragment analyzed before, our procedure can be broken down into the following steps:

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

2. *the role played by monetary policy*

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

The tone is calculated is as follows:

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

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

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 which was ignored in the existing methodology.

Comparing the sentence tones of the existing methodology ($=-0.09$) and the new methodology ($=-0.15$) reveals a stark difference between the degree of negativeness embedded in the sentence. 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.

Table 1 presents the frequency with which valence shifters appear in the speeches by functionaries of the Federal Reserve.

Insert table 1 about here.

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. A priori, we would expect that if the quantification of valence shifters in the speeches leads to a measurable change in the tone of speeches, it will be reflected in the distribution of speech tones calculated under the new methodology. To this end, 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 methodology 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 current 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, the results of which are presented in table 4. It can be seen that 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.

Insert table 4 about here.

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:

$$
\begin{aligned}
R_t = a_0 + b_n Tone_{t-n} + \sum_{i=1}^{3} c_i R_{t-i} + d_1 * Controls \\
+ d_2 * SpeechControls + d_3 * MacroControls + \gamma_t
\end{aligned}
\tag{1}
$$

$$Vol_t = a_0 + b_n Tone_{t-n} + \sum_{i=1}^{3} c_i Vol_{t-i} + d_1 * Controls$$
$$+ d_2 * SpeechControls + d_3 * MacroControls + \gamma_t \tag{2}$$

n ranges from 0 to 5 and controls include the day of the week and month dummy for time series controls; 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 two control variables to account for the readability and complexity of speeches [Gunning, 1952, Li, 2008, Biddle et al., 2009, Miller, 2010]. The lags of return as control are kept in accordance with previous studies which examine the impact of central bank communication on index returns [Ehrmann and Fratzscher, 2007, Born et al., 2014, Gertler and Horvath, 2018].

In addition, we include macroeconomic variables as control factors. These include the real exchange rate 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

There are three sources of the data used in this study: Fed speeches from the Federal Reserve website, intraday data on index returns from 'FirstRate Data';[9] and data for stock indices, VIX, controls and macro variables are taken from Bloomberg. The Fed Funds rate data are downloaded from the

[9] https://firstratedata.com/

St. Louis Fed website (FRED).[10]

All speeches are downloaded automatically using web parsing from the official Federal Reserve website. The sample used in this study only includes speeches given by the Governor/Chairperson and Deputy Governor/Vice Chairperson. It does not include press releases or FOMC announcements as these communication forms have been the focus for an array of studies in the past. The speech data are available for the Fed from January 2006 to February 2020. Overall, 77% of the speeches are delivered by the Governor/Chairperson and the remaining 23% are by the Deputy Governor/Vice Chairperson.

5 Results and Analysis

5.1 Descriptive Statistics

We first provide basic descriptive statistics for the frequency of speeches, the number of words contained therein; and for the speech tone calculated according to the methodology specified in this study. Tables 5, 6 and 7 present the relevant results.

Insert table 5 about here.

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;[11] while the shortest contains 237 words.

Insert table 6 about here.

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

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

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

in our sample is 0.29, while the lowest is -0.34. The standard deviation is 0.09.

Insert table 7 about here.

Table 8 presents daily return statistics of the two benchmark stock indices of the US: the S&P 500 and the DJIA. Both display a mean 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 8 about here.

Figure 2 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 2 about here.

Similarly figure 3 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.

Insert figure 3 about here.

Together, figures 2 and 3 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 in more detail.

5.2 Impact of Federal Reserve speech tone

In the following subsections, we present extensive evidence on the impact of speeches delivered by the Federal Reserve on the return as well as volatility of US stock markets. In order to add weight to our investigation, we conduct this analysis on both the daily as well as the intraday levels.

5.2.1 Impact on the S&P 500 daily returns

Table 9 presents the results of regressing the daily returns of the benchmark S&P 500 index on the speech tones of the Federal Reserve in line with equation (1). The methodology is that of ordinary least squares with heteroskedasticity and autocorrelation consistent (HAC) errors.[12]

Insert table 9 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 are associated with a drop in daily return; and those with positive tones are associated with an increase in the daily return. Based on our result, a one standard deviation change in the Federal Reserve speech tone is associated with 0.07 standard deviation change in daily market return for the S&P 500 index.

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

5.2.2 Impact on the S&P 500 2 day return

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 results are presented in table 10 which show the speech tone to be 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.

Insert table 10 about here.

5.2.3 Impact of speech tone from forward-looking speeches

As specified in Ehrmann and 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 are less likely to be endogenous [Gertler and 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 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.

Tables 11 and 12 present 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 tables 9 and 10 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 9. Further, for the two-day return, the impact is observed on both days—the speech-day as well as the next day.

Insert tables 11 and 12 about here.

5.2.4 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 US stock markets. 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 13 about here.

Table 13 presents our results on the effect of speech tone on the daily realized volatility of the S&P 500 index. 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 drop in daily volatility. We note that this result is in agreement with that for the Hungarian central bank impact on volatility outlined in Máté et al. [2021].

Table 14 presents the results of the effect of speech tone on changes in the daily VIX. 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 coefficient (-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.

Insert table 14 about here.

5.2.5 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 specified in Gertler and Horvath [2018], the examination of the impact of central bank communication on financial variables at shorter frequencies is less likely to suffer from endogeneity issues. The results of our examination are presented in table 15. 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 the next day (Day 1). This is because, if the speech was delivered, for example, at 1 PM, the markets will be able to react to its content only after 1 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.

Insert table 15 about here.

5.2.6 Impact on intraday changes in VIX

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.

Table 16 presents the results of our examination. 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 16 about here.

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

As described in table 6, 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 17. 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.[13] 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.[14]

Insert table 17 about here.

5.2.8 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 Latent Dirichlet Allocation (LDA) [Blei et al., 2003, Hansen et al., 2018].

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

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

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

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

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

Our main findings are presented in table 18. The results are quite similar to our benchmark results in table 9 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 9.

Insert table 18 about here.

5.2.9 Impact on the US term premium

Gilchrist et al. [2019] examine the impact of the US monetary policy on dollar denominated sovereign bonds and find that US monetary easing leads to a significant narrowing of credit spreads on these bonds. Similarly, Tillmann [2020] examines the impact of monetary policy surprises on term structure of interest rates and reports that policy tightening leads to a significantly smaller increase in long-term bond yields. On similar lines, we also examine the impact of Fed speech tone on the US term premium and the results are presented in table 23. The term premium is calculated using the methodology specified in Adrian et al. [2013]. The data for the calculated term premium is available from the New York Fed website.[16] 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 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 23 about here.

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

6 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.

6.1 Impact on DJIA

In table 19 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 19 about here.

The table indicates that the results are almost identical to that with the S&P 500 index returns outlined in table 9. 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 20 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 9.

Insert table 20 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 8.

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.

6.2 Investigating reverse causality: Impact of index return on speech tone

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 21 and we find that the daily index return does not have any significant impact on the Fed speech tone for any lag.

Insert table 21 about here.

6.3 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 22.

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 22 about here.

7 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 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.

Tables and Figures

7.1 Tables

Table 1: Valence Shifter Frequency

Entity	% of Sentences containing valence shifters	% of Adversative Conjunction	% of Amplifier	% of De-amplifier	% of Negator
Federal Reserve	37.91%	16.60%	53.26%	10.86%	19.26%

Note: This table presents the frequency of valence shifters in the speeches by the Federal Reserve.

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: KS Test

D Statistic	p-value
0.17417	1.676×10^{-9}

Note: This table presents the statistics for the Kolmogorov Smirnov test to examine the difference between the speech tone distribution calculated using the "bag-words-approach" and the LM dictionary and the new methodology specified in this study. The D statistic specifies the distance between the two tone distributions and the p-value is for the null hypothesis that there is no difference between the two distributions.

Table 5: Speech Frequency

Time Period	Total Speeches	Total after combining for same day	# Positive Speeches	# Negative Speeches	Avg / month
Jan 2006–Feb 2020	797	693	146	547	4.1

Note: This table presents the count for speech frequency. The data are obtained from the official Federal Reserve website. The 3rd column shows the number of speeches after combining all speeches delivered over a day into one composite speech.

Table 6: Speech Statistics

Time Period	Max	Min	Mean	SD
Jan 2006–Feb 2020	10923	237	3482	1662.56

Note: This table presents the summary statistics for the number of words in the speech sample. 'Max' denotes the maximum number of words in a speech, 'Min' denotes number of words in the shortest speech, 'Mean' is the average number of words per speech; and 'SD' denotes the standard deviation of words per speech.

Table 7: Speech Tone Statistics

Time Period	Max	Min	Mean	SD
Jan 2006–Feb 2020	0.2949	-0.3403	-0.0605	0.0864

Note: This table presents the summary statistics for speech tone. 'Max' denotes the maximum value of speech tone in a speech, 'Min' denotes minimum value of speech tone, 'Mean' is the average value of speech tone; and 'SD' denotes the standard deviation of the speech tone.

Table 8: Index Return Statistics

Index	Max Daily Return %	Min Daily Return %	Mean Daily Return %	SD Daily Return %	Trading days per year
S&P 500 Index	11.5800	−11.9840	0.0227	1.2574	251
DJIA Index	11.3650	−12.9265	0.0233	1.2089	251

Note: This table presents the summary statistics for daily return of the two major US indices. The data for each index are obtained from Bloomberg.

Table 9: Impact of Federal Reserve speech tone on the S&P 500 daily returns

Variable	Day 0	Day 1	Day 2	Day 3	Day 4	Day 5
S&P 500 Index	0.011* (0.006)	0.003 (0.006)	0.001 (0.006)	−0.001 (0.006)	0.0002 (0.007)	0.002 (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 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 10: Impact of Federal Reserve speech tone on the 2 day return

Variable	Interval 0	Interval 1	Interval 2	Interval 3	Interval 4	Interval 5
S&P 500 Index	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 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 11: Impact of Federal Reserve speech tone (Forward Looking) on the daily return

Variable	Day 0	Day 1	Day 2	Day 3	Day 4	Day 5
S&P 500 Index	0.028**	0.007	0.014	−0.002	−0.005	−0.003
	(0.014)	(0.009)	(0.011)	(0.010)	(0.011)	(0.011)
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 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 12: Impact of Federal Reserve speech tone (Forward Looking) on the 2 day return

Variable	Interval 0	Interval 1	Interval 2	Interval 3	Interval 4	Interval 5
S&P 500 Index	0.027*	0.024*	0.008	0.010	0.005	−0.004
	(0.016)	(0.014)	(0.013)	(0.013)	(0.012)	(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 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 13: Impact of Federal Reserve speech tone on the S&P 500 daily realized volatility

Variable	Day 0	Day 1	Day 2	Day 3	Day 4	Day 5
S&P 500	-0.0002^{*}	-0.00006	-0.0001	0.00009	0.0002	-0.0003^{*}
	(0.0001)	(0.0002)	(0.0002)	(0.0001)	(0.0001)	(0.0001)
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 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 14: Impact of Federal Reserve speech tone on changes in the daily VIX

Variable	Day 0	Day 1	Day 2	Day 3	Day 4	Day 5
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 changes in the daily VIX on speech tone (and controls). The results are reported in line with equation (2). The standard errors are reported in the parentheses and are all heteroskedasticity and autocorrelation (HAC) robust. The controls include three lags of changes in daily VIX, 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 on the S&P 500 intraday returns

S&P 500 Index Return		
Interval	Day 0	Day 1
Interval 1	0.006 (0.003)	0.012*** (0.004)
Interval 2	−0.001 (0.001)	−0.0007 (0.002)
Interval 3	−0.0004 (0.001)	−0.001 (0.001)
Interval 4	0.003*** (0.001)	0.002 (0.001)
Interval 5	0.00003 (0.001)	−0.0008 (0.001)
Interval 6	−0.0005 (0.001)	0.0001 (0.001)
Interval 7	0.0005 (0.001)	−0.0007 (0.001)
Interval 8	0.002** (0.001)	0.001 (0.001)
Interval 9	0.0001 (0.001)	−0.002* (0.001)
Interval 10	−0.0005 (0.001)	0.0004 (0.001)
Interval 11	−0.0007 (0.001)	−0.001 (0.001)
Interval 12	0.002* (0.001)	−0.002 (0.001)
Interval 13	−0.0007 (0.002)	−0.0007 (0.002)
Interval 14	0.0006 (0.0005)	−0.0002 (0.0002)
Controls	Yes	Yes
Macro Controls	Yes	Yes
Speech Controls	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 16: Impact on intraday changes in the VIX

	Changes in VIX	
Interval	Day 0	Day 1
Interval 1	-0.026 (0.026)	-0.044^{**} (0.023)
Interval 2	-0.00007 (0.009)	0.002 (0.012)
Interval 3	0.003 (0.008)	0.006 (0.008)
Interval 4	-0.019^{**} (0.007)	-0.014^{**} (0.007)
Interval 5	-0.007 (0.006)	0.008 (0.008)
Interval 6	0.002 (0.007)	-0.006 (0.005)
Interval 7	-0.004 (0.006)	0.001 (0.006)
Interval 8	-0.003 (0.006)	-0.00008 (0.006)
Interval 9	-0.002 (0.006)	-0.0006 (0.006)
Interval 10	0.004 (0.007)	-0.001 (0.006)
Interval 11	0.007 (0.007)	-0.003 (0.008)
Interval 12	-0.013^{*} (0.007)	0.010 (0.009)
Interval 13	-0.004 (0.012)	0.0007 (0.011)
Interval 14	0.006 (0.009)	-0.008 (0.010)
Controls	Yes	Yes
Macro Controls	Yes	Yes
Speech Controls	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 changes in the VIX in line with equation (2). The one difference as compared to the daily regression is that three lags of VIX 30 minute changes are not kept as additional controls. The standard errors are reported in the parentheses and are all Heteroskedasticity and Autocorrelation (HAC) robust.

Table 17: 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 18: Impact on S&P daily returns based on topic analysis

	Day 0	Day 1	Day 2	Day 3	Day 4	Day 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 19: Impact on DJIA daily returns

	Day 0	Day 1	Day 2	Day 3	Day 4	Day 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 20: Impact on DJIA intraday returns

Interval	Day 0	Day 1
Interval 1	0.004 (0.003)	0.012*** (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 (0.001)	−0.0009 (0.001)
Interval 6	−0.0002 (0.001)	−0.00005 (0.001)
Interval 7	0.00004 (0.001)	−0.0007 (0.001)
Interval 8	0.002** (0.001)	0.001 (0.001)
Interval 9	0.0003 (0.001)	−0.002** (0.001)
Interval 10	−0.0003 (0.001)	0.0002 (0.001)
Interval 11	−0.0004 (0.001)	−0.0008 (0.001)
Interval 12	0.002* (0.001)	−0.002 (0.001)
Interval 13	−0.0002 (0.002)	−0.0009 (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 21: Impact of the S&P 500 daily returns on Federal Reserve speech tone

Variable	Day 0	Day 1	Day 2	Day 3	Day 4	Day 5
S &P Return	0.319 (0.609)	−0.196 (0.790)	−1.501 (3.839)	0.219 (0.818)	0.198 (1.626)	−3.206 (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 22: 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.

Table 23: 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.

7.2 Figures

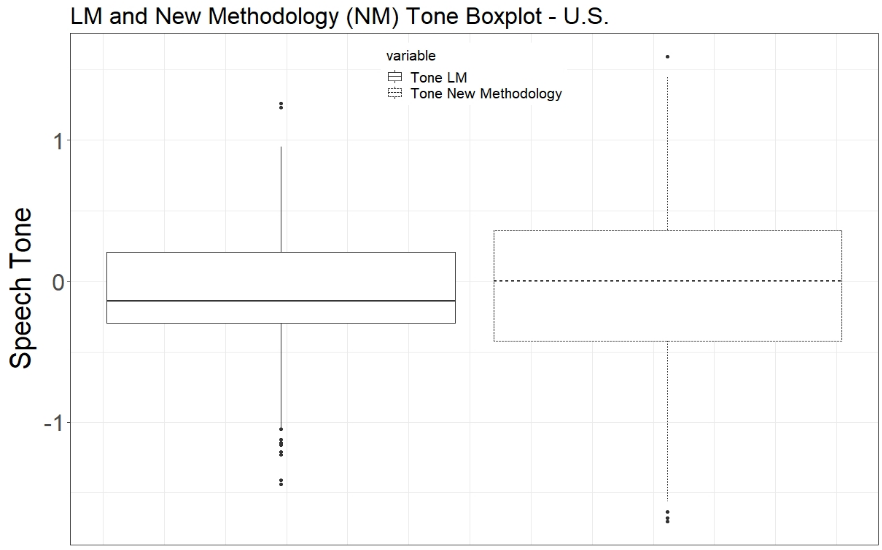

Figure 1: Boxplots of speech tones under the old and new methodologies. The LM tone (solid line) is speech tone calculated using the "bag-of-words" approach and the LM dictionary whereas the new methodology tone (dotted line) is the tone calculated by the methodology specified in this study

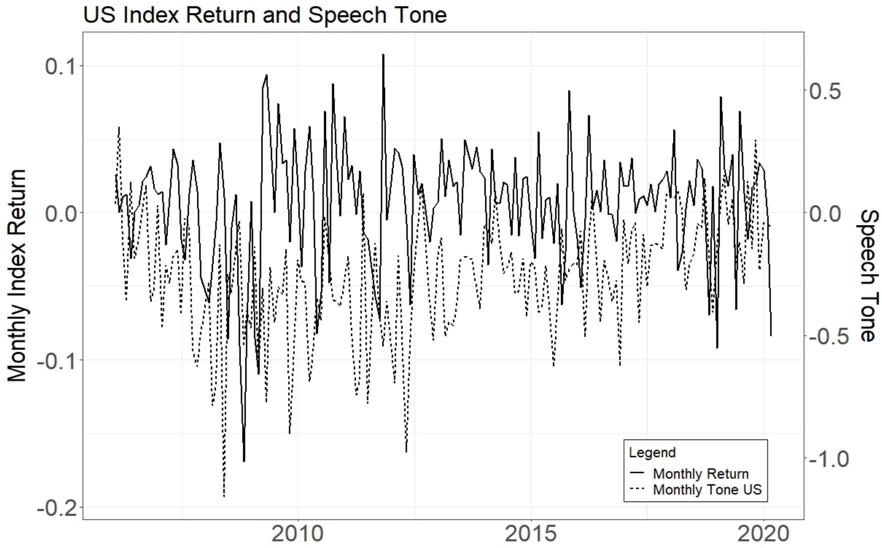

Figure 2: The time series of monthly S&P 500 index returns on the primary y axis; and the monthly speech tone on the secondary y-axis.

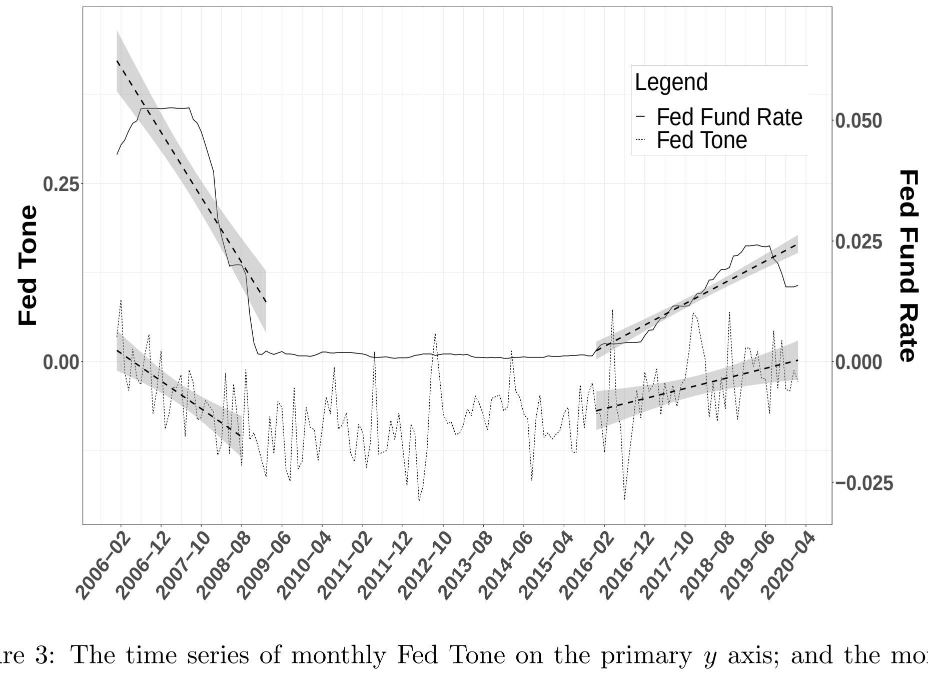

Figure 3: The time series of monthly Fed Tone on the primary y axis; and the monthly Fed Rate on the secondary y-axis.

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^{*} (0.006)	0.007 (0.006)	-0.0007 (0.006)	0.0009 (0.005)	-0.002 (0.007)	0.0001 (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 9. 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

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

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

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

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

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

Paweł Baranowski, Wirginia Doryń, Tomasz Łyziak, and Ewa Stanisławska. Words and deeds in managing expectations: Empirical evidence from an inflation targeting economy. *Economic Modelling*, 95:49–67, 2021.

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

Saroj Bhattarai and Christopher J Neely. An analysis of the literature on international unconventional monetary policy, 2016. FRB St. Louis Working Paper 2016-21.

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

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

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

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

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

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

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

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

Sergio Consoli, Luca Tiozzo Pezzoli, and Elisa Tosetti. Emotions in macroeconomic news and their impact on the European bond market. *Journal of International Money and Finance*, page 102472, 2021.

Ricardo Correa, Keshav Garud, Juan M Londono, and Nathan Mislang. Sentiment in central banks' financial stability reports. *Review of Finance*, 25 (1):85–120, 2021.

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

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

Asad Dossani. Central bank tone and currency risk premia. *Journal of International Money and Finance*, page 102424, 2021.

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

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

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

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

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

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

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

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

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

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

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

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

Kohei Kawamura, Yohei Kobashi, Masato Shizume, and Kozo Ueda. Strategic central bank communication: Discourse analysis of the bank of japan's monthly report. *Journal of Economic Dynamics and Control*, 100:230–250, 2019.

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

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

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

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

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

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

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

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

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

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

Rickard Nyman, Sujit Kapadia, and David Tuckett. News and narratives in financial systems: exploiting big data for systemic risk assessment. *Journal of Economic Dynamics and Control*, 127:104119, 2021.

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

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

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

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

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

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

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

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

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

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