IIMB-WP N0. 622

WORKING PAPER NO: 622

Who moved the Market? Analyzing the role of Central Bank 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 – September 2020

"Who moved the Market"? Analyzing the role of Central Bank speeches.

Abhinav Anand, Sankarshan Basu, Jalaj Pathak, Ashok Thampy

Abstract

We quantify the sentiment from central bank speeches of five leading developed nations (US, UK, Japan, France, and Germany) and analyze their role in explaining the return of stock market indices for the respective nations. In this study we improve upon existing sentiment quantification techniques by introducing two innovations: (i) by introducing the sentence as the unit of analysis, and (ii) by introducing "valence shifters", which assign appropriate weights to adjectives and adverbs. We demonstrate that our modified sentiment extractor is a more effective explanatory variable as compared to both direct measures (Consumer Confidence Indices) as well as indirect measures (Baker and Wurgler Index).

1 Introduction

The role of central banks in moderating financial stability and assisting economic growth has been considered of prime importance and has been a focus of an array of studies in the past. Hence, central bank communication in terms of speeches by governors is watched very closely by all market participants. The studies on central bank communication can be broadly classified into two categories. The first category is the set of studies in which the central bank's communications' reaction is quantified into a dummy classification (e.g., +1, 0, -1) based on the subjective assessment of its content by the researcher. For example, Guthrie and Wright [2000] use central bank communication to show how central bank statement (rather than open market operations) can be used to implement monetary policy in New Zealand. The communication is classified into objective categories (+1,0,-1) based on the authors' subjective assessment and it is shown that the communication, rather than open market operations causes the large changes in interest rates. The second category includes studies that analyze the importance of speech days based on a dummy variable for the presence/absence of the speech. For example, Savor and Wilson [2013] show that the average market return and Sharpe ratio are significantly higher on important macroeconomic announcement days.

However, there are two drawbacks to both categories of studies. Concerning the first category, the classification of the communication is subjective and thus can vary depending upon the researcher as well as on the objective of the study. Thus, the results of such studies cannot be agreed upon to be standard. Whereas the second category of studies focus just on the event of speech, ignoring its content and hence its impact.

We propose a way to overcome these drawbacks by specifying an objective method to quantify sentiment. The proposed methodology also marks an improvement in the current methods of textual analysis in finance including "bag-of-words"(Tetlock [2007], Li [2008], Tetlock et al. [2008]) and the "ngram" approach, as well as the Loughran and McDonald's (LM) dictionary (Loughran and McDonald [2011]).

We introduce two new innovations which improve text analysis in finance:

1. We introduce the sentence as a new unit of analysis for sentiment extraction. Prior methods ("bag of words"/"ngram") have overlooked this aspect.

2. We introduce the concept of "valence shifting" in financial text analysis

(Kennedy and Inkpen [2006], Polanyi and Zaenen [2006], Schulder et al. [2018]) which are adjectives and adverbs (such as "almost", "but" etc.) which modify the meaning of text but were classified as stop words and hence removed from the content in the parsing process.

We divide a speech into a set of sentences and extract the sentiment for each sentence considering both the polar words (negative/positive) as well as the adverbs and adjectives (valence shifters) (Kennedy and Inkpen [2006], Polanyi and Zaenen [2006], Schulder et al. [2018]) surrounding the polar words. Since valence shifters have not been used yet to classify financial texts, fifty-two of these (for example, "ain't","although","almost") have been classified as stopwords in the Loughran and McDonald dictionary.[1] Thus, we also improve the existing dictionary by taking these words out of the stopwords' list and giving them appropriate weightage as they can modify and/or alter the meaning of the sentence.[2] For example, for the sentence *"The economy has recovered although slowly"*, the sentiment using LM dictionary and "bag-of-words" approach is -0.40, whereas using the modified method and valence shifters is -0.91, as the word "although" is not given appropriate weigtage in the existing method and LM dictionary. To the best of our knowledge this is the first instance of the usage of valence shifting in financial text analysis.

Also, since a whole sentence is considered as a single unit to quantify sentiment, this solves the question of how many words should be considered as a cluster for sentiment extraction and thus gives a valid alternative to"bag-of-words" (one word at a time) and ngram (n-words at a time) approach.[3]

We show with an example, how the "bag-of-words" approach along with the LM dictionary can understate/overstate the sentiment. Additionally we

[1]Since they were classified as stopwords, they were removed from the content in the parsing process

[2]Full list of valence shifters is presented in 9.

[3]This is based on the assumption that a sentence is a complete unit in itself.

also show that, in cases where the negative valence shifters are not taken into consideration, the LM dictionary and “bag-of-words” approach can lead to incorrect sentiment.[4]

Further, using the updated process and dictionary we find a significant effect of speech sentiment in explaining returns for the U.S., U.K., and Japan. We also show that due to its unique properties, and the drawbacks in the existing sentiment variables, speech sentiment is a more effective explanatory variable for market returns in comparison to the existing sentiment variables.

With respect to the existing sentiment variables, they can be categorized as per Figure 1.

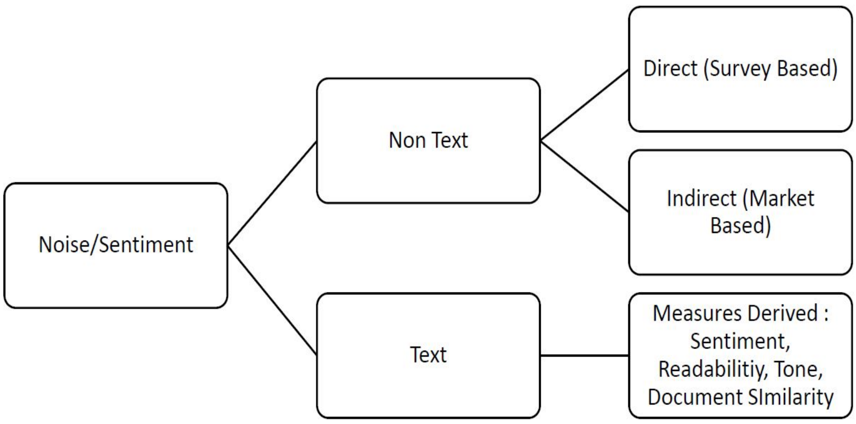


Figure 1: The existing sentiment varibles can be divided into two categories, Text and Non-Text, the Non-Text category can further be divided into Direct (e.g. Consumer Confidence Index) and Indirect (BW Index).

However there are certain drawbacks to these measures, for example, the direct measures (survey-based: e.g. Consumer Confidence Index) might be outdated by the time they are published, especially if markets are volatile

[4]There are 19 such negators in the list of 52 valence shifters which were classified as stopwords in the LM dictionary.

Simon and Wiggins III [2001]. On the other hand, speeches and communication from the central banks are available almost immediately and hence can be analyzed in real-time. Another issue is that the responses in surveys are equally weighted regardless of the investment position of the respondents. This issue of uniform versus variable weighting does not apply to our speech sentiment extractor.

Similarly, indirect measures (such as Baker and Wurgler Index), are derived from market variables and in turn are used to predict other market variables hence there can be a case of reverse causality. This issue is resolved in case of speeches since a vast majority of these are confirmed to be delivered in advance and not in a response to a specific event or crisis.[5] This is reflected in the finding that other sentiment extractors become insignificant in explaning returns in the presence of our modified sentiment measure.

We ensure the robustness of the results in two ways, firstly we include EPUI (Economic Policy Uncertainty Index - Baker et al. [2016]), which is based on news coverage about policy-related economic uncertainty. Since central bank communication is bound to make news in most circumstances, this is an important control variable. Secondly, we analyze the effect of speech sentiment on two of the largest and fastest-growing emerging markets (India and China) to ensure the effect is not due to the special attributes of the developed markets.

The paper is organized as follows, Section 2 is the Literature Review for central bank speeches as well as existing measures of sentiment, Section 3 specifies the methodology for sentiment calculation and analysis followed by Section 4 which lists the data sources. Section 5 is for analysis and results followed by Section 6 for the discussion of the results. Section 7 is for robustness analysis and finally, Section 8 offers concluding remarks.

[5] We also test this conjecture by analyzing the impact of return on speech.

2 Literature Review

The literature can be divided into three categories, two of which are based on the existing categories of sentiment variables (Direct and Indirect) and the third is based on the new variable we introduce (based on central bank communication). All three categories are discussed below:

2.1 Central Bank Communication

Due to the perceived economic and financial importance of the central bank, the work centering around them has been ample as well as diverse. 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. On the other hand, Kohn et al. [2003], Demiralp and Jorda [2004], Ehrmann and Fratzscher [2004] and Jansen and De Haan [2006] are among the studies which categorize days as a dummy variable based on the presence or absence of central bank communication. Jansen and De Haan [2006] also study the comments by central bankers on the interest rate, inflation, and economic growth in Eurozone. The statements are categorized into dummies based on subjective analysis by the authors. Similarly, Gerlach et al. [2007] discusses the interest rate related statements made by ECB and their respective impact using subjective dummy classification of the statement by the authors. 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. On related lines, Gentzkow et al. [2019] analyze the trend in partisanship of congressional speeches using machine learning and find that partisanship has increased since early 1990s.

2.2 Text based Measures

With respect to quantification of sentiment from financial text, Antweiler and Frank [2004] extract sentiment from message activity in chat rooms and analyze its impact on trading volume. Tetlock [2007], Engelberg [2008], Li [2008], Tetlock et al. [2008], Li [2010] are some of the other important studies which have used "bag-of-word" as well as Machine Learning approach to classify financial texts as positive or negative. These studies have used 10-K reports, newspaper articles, message boards, and press releases as sources of the text. Loughran and McDonald [2011] specify a new dictionary and show its importance in comparison to the Harvard IV dictionary for analyzing financial texts. On similar lines, Garcia [2013] and Jegadeesh and Wu [2013] study the impact of sentiment, calculated from news stories and use of term weighing for sentiment calculation respectively. Kearney and Liu [2014] provides a survey of methods in text sentiment in finance. Solomon et al. [2014] shows how media coverage of fund holdings affects investors' fund allocation. Kim and Kim [2014] study the relationship between investment sentiment calculated from message postings in Yahoo! Finance and stock returns. Chen et al. [2014] analyze the impact of social media calculated sentiment on stock returns and earnings surprises. Further, Loughran and McDonald [2015] study the different dictionaries and their respective suitability for analyzing financial documents. Loughran and McDonald [2016] do a survey of the textual analysis in Accounting and Finance.

2.3 Direct and Indirect Measures

With respect to direct and indirect sentiment measures, the literature has evolved over the last three decades. It can be traced back to Lee et al. [1991] and Neal and Wheatley [1998] using closed-end fund rates among other variables as a proxy for an indirect measure of sentiment. Following this Baker and Stein [2004] analyze the viability of liquidity as an indirect measure. Baker and Wurgler [2006] and Baker and Wurgler [2007] are the first studies that use proxy market variables and form a sentiment index using these variables. On similar lines, Kim and Ha [2010]; Liao et al. [2011]; Baker et al. [2012], Greenwood and Shleifer [2014]; Yang and Gao [2014]; Yang and Zhou [2015] use market variables as proxies to form an indirect measure of sentiment index and test its impact on varying market variables.

Similarly, for direct variables, Otoo [1999] and Charoenrook use the University of Michigan consumer survey sentiment index. Lee et al. [2002] calculate sentiment from the Investor Intelligence (II) survey. Jansen and Nahuis [2003] use survey data of the European Commission and study the relationship between sentiment and returns and report significant results. Similarly, Lemmon and Portniaguina [2006], Bergman and Roychowdhury [2008], Schmeling [2009], Zouaoui et al. [2011], Spyrou [2012], Arif and Lee [2014], Aristei and Martelli [2014] and, Szu et al. [2015] use Consumer Confidence Index (CCI) in their analysis.

3 Methodology

3.1 Sentiment Quantification

We calculate the sentiment for each speech by classifying it as a collection of sentences. Also, for instances where there are multiple speeches on the same day, the content for all is analyzed as one. 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 three categories, valence shifters (adjectives and adverbs), polar words (positive/negative sentiment words) and stop words.[6] The polar words are taken from the LM dictionary. An important distinction here from the LM process is that words such as "ain't", "although" and "almost" which are classified as stopwords in LM, are classified as valence shifters in this study. This is so because these words do add to the meaning of the respective sentence. Fifty-two such words are taken from the stopwords list in LM and are included in the valence shifters list. The valence shifters are further classified into four categories, i.e., amplifiers ("absolutely", "acutely", "very"), de-amplifiers ("barely", "faintly", "few"), negators ("ain't", "aren't", "cannot") and adversative conjunction ("despite", "however", "but"). The amplifiers, de-amplifiers, and adversative conjunction are given a weight a 0.8 (positive for an amplifier, negative for a de-amplifier and negative for the words before adversative conjunction and positive for the words after adversative conjunction).[7] This is done 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. Thus, for each sentence, first, the updated stop words (updated by removing valence shifters from the LM stopwords list) are removed from a sentence. After that polar words are identified and given the weight of +1/-1, following which a set of words are identified around each polar word from the beginning till the end of the sentence. This is classified as a word cluster for each polarized word.

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 sentiment in two ways. Firstly, by quantifying the sen-

[6]The list of valence shifters contain each word in all possible forms (for e.g. "understand", "understandable", "understandably"), thus it is not required to stem the words to their basic form ("understand" in this case).

[7]The weight, 0.8, is as per the existing literature (Kennedy and Inkpen [2006], Polanyi and Zaenen [2006], Schulder et al. [2018]), we verify the results by varying the weight of valence shifters from 0.5 to 0.9 and our results still holds.

[8]E.g. "The economy is doing well but the rising prices are a concern."

timent with the correct sign but under/overstating the coefficient. Secondly, by quantifying with an incorrect sign (by missing the 19 negators (valence shifters) classified as stopwords). We present examples of both categories.

3.1.1 Category 1 - Correct Sign but Incorrect Coefficient

The below sentence is taken from one of the speeches from our sample:

The central bank considers the most important thing during the recent economic-financial situation to not abandon the financially unstable firms and to be able to accomplish smooth financing operations.

From the above sentence, stop words are removed and it becomes:

central bank considers most important thing recent economic-financial situation not abandon financially unstable firms and able accomplish smooth financing operations.

From the above sentence, polarized words are identified (using the LM dictionary) which are "abandon" (-1), "accomplish" (+1), and "smooth"(+1). Further, word cluster around these polar words are identified thus, there will be three clusters in this sentence:[9]

1. *central bank considers most important thing recent economic-financial situation not abandon*

2. *financially unstable firms, and able accomplish*

3. *smooth financing operations.*

[9]In case of adversative conjunction clusters are formed both before and after the polar word.

Now valence shifters are identified in each of these clusters, for example, "most", "not" are valence shifters in the first cluster. The sentiment of the cluster is then defined by the valence shifter/s in combination with the polar word. For example, in the first cluster "most" will increase the weight of abandon by 0.8 and "not" will multiply the resulting weight by -1, hence the sentiment of the cluster will be 0.8*-1*-1 = +1.8. Thus, the sentiment of the sentence will be defined by overall sentiment from all clusters divided by the number of words in that sentence. The sentiment is then averaged across sentences to get the sentiment of each speech.

The sentiment calculated using the above method turns out to be

$$\frac{1.8^{*}\text{first cluster} + 1^{*}\text{“smooth”} + 1^{*}\text{“accomplish”}}{19} = 0.14$$

Whereas using the LM method it turns out to be[10]

$$\frac{-1^{*}\text{“abandon”} + 1^{*}\text{“smooth”} + 1^{*}\text{“accomplish”}}{17} = 0.05$$

Thus the sentiment is understated by 1.8 times in comparison to the LM dictionary and "bag-of-words" approach.

3.1.2 Category 2 - Incorrect Sign

As an example of miscalcuation of the sentiment, consider the below sentence taken from another speech in our sample:

The taxpayers shouldn't worry that they would be burdened with bailouts like before.

From the above sentence, stop words are removed and it becomes:

[10]The number of total words are two less as the stop word list is updated in this study (as specified earlier).

taxpayers shouldn't worry burdened bailouts.

From the above phrase, polarized words are identified (using the LM dictionary), which is this case is "worry" (-1). Since, there is only one polar word, word cluster constitutes everything around it.

1. *taxpayers shouldn't worry burdened bailouts.*

Now valence shifters are identified in the cluster, which is "shouldn't" (-1). Thus, the sentiment calculated is:

$$\frac{-1^{*}-1}{5} = 0.20$$

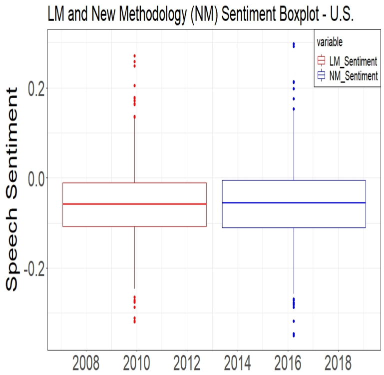
Whereas using the LM method it is[11]:

$$\frac{-1}{4} = -0.25$$

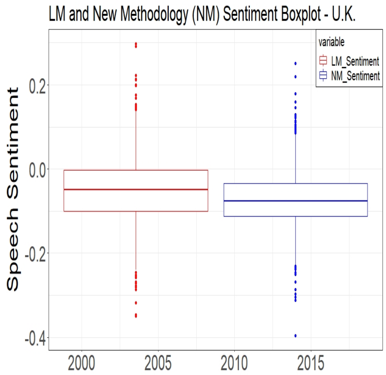
Thus, in this case the sentence is considered as negative as per LM dictionary and existing process even though it is is positive (as classified by updated process and dictionary).

The overall difference in speech sentiment calculated using LM and new methodology for all nations is presented in figures 2a, 2b, 2c, 3a and 3b. It can seen that the for U.K., France, and Germany there is a difference in the boxplot of LM sentiment (red) versus our methodology (blue) in terms of mean, size of the box (50% of observations), and the tail values. For these three nations we also find that the difference of mean sentiment, tested using bi-variate t-test, between LM and new methology is also significant at 1%. Also, the modified sentiment extractor also yields lower medians for the three nations. These are indications of the difference in sentiment quantification between the existing methodolgy and the one introduced in this study.

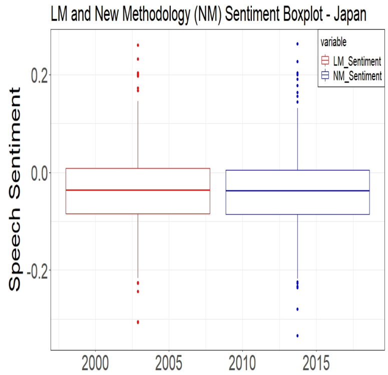
[11] The number of total words is one less as the stop word in the LM dictionary.



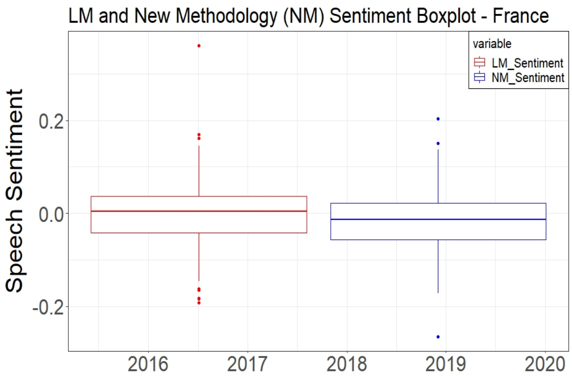
(a) The boxplots represent the distribution of speech sentiment for LM dictionary and “bag-of-words” approach (in red) in comparison to the methodology and modified dictionary used in this paper (in blue).



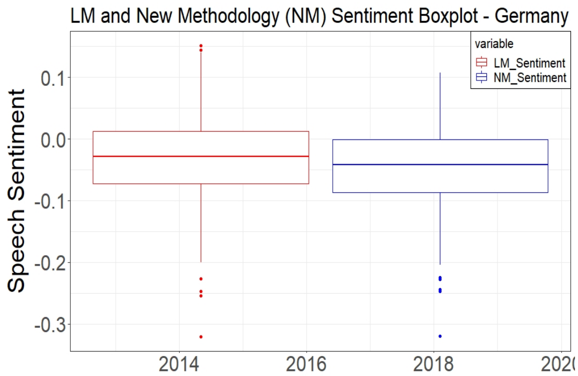
(b) The boxplots represent the distribution of speech sentiment for LM dictionary and “bag-of-words” approach (in red) in comparison to the methodology and modified dictionary used in this paper (in blue).



(c) The boxplots represent the distribution of speech sentiment for LM dictionary and “bag-of-words” approach (in red) in comparison to the methodology and modified dictionary used in this paper (in blue).



(a) The boxplots represent the distribution of speech sentiment for LM dictionary and “bag-of-words” approach (in red) in comparison to the methodology and modified dictionary used in this paper (in blue).



(b) The boxplots represent the distribution of speech sentiment for LM dictionary and “bag-of-words” approach (in red) in comparison to the methodology and modified dictionary used in this paper (in blue).

3.2 Empirical design

Return is calculated as per the below formula for both daily as well as monthly frequency.

$$R_i = (P_i - P_{i-1})/P_i$$

Where i denotes the respective day or month.

We use central bank speech sentiment to explain the stock market returns of five major developed nations: US, UK, Japan, France and Germany by testing their respective market indices as well as by their respective small cap indices as specified in Table 3.

An important point here is that for the U.S. we follow the process of Tetlock [2007] for return calculation of the smallcap index. Thus, we orthogonalize the Rusell 2000 Index with respect to the SP500 Index and use the residual returns. We do not do this process of orthogonalization for the other four nations. This is because only in the case of the U.S. there are more than one contenders for main Index (DJIA, SP500).

Also, the analysis is done for both daily as well as monthly frequency. This is done to ease comparisons with all existing sentiment variables (direct and indirect) all of which are of monthly frequency. The monthly sentiment is calculated by summing over the sentiment for all speech days of a particular month.

A number of past studies, including Tetlock [2007], while analyzing the relationship between sentiment and Index return have used VAR (Vector Autoregression). However, we do not employ VAR and use OLS as the speeches are spread intermittently. Hence, there are days as well as months when there are no speeches. Thus, in this case, if we use VAR the number of observations reduces drastically (less than 30 for all nations except the U.S. and U.K.).

Also, since the impact of sentiment can be delayed due to socio-economic reasons it is tested for up to three lags. The lags are kept in accordance with Tetlock [2007].

Thus, the below equation is tested for both daily and monthly frequency:

$$R_t = a_0 + a_1 Sent_{t-n} + a_2 R_{t-1} + a_3 R_{t-2} + a_3 R_{t-3} + +a_5 Controls + \gamma_t \quad (1)$$

Where n ranges from 0 to 3 and controls include the day of the week and month dummy (for daily analysis).

For monthly analysis (for equation 1), we first check at which lag/s is the sentiment coefficient significant, and then the existing sentiment variables are added as controls at the same lag.[12]

4 Data

The data are acquired from an array of sources due to the varied nature of variables. All the speeches are downloaded automatically using a link extractor from the official website of each country's central bank.[13] Except for Germany, all other speeches are available in English from the official website. For Germany, the speeches are converted into English using Google Translate. Five variables are extracted from the speeches: date of delivery, place of delivery, speaker, the title of the speech, and content. The index data for all countries is downloaded from Bloomberg.

Sources of the existing measures of sentiment are listed below:

1. Direct Measure: We use the Economic Sentiment Index for Germany and France since the Consumer Confidence Index is not available for

[12] We also check the existing sentiment measures as controls at all other lags and the results are the same.

[13] One of the reasons why speeches are downloaded from the official website and not as reported in the news articles (from Reuters or Bloomberg News) is to ensure that the content is in its original form. This is so because, in most cases, news articles, in addition to the reported speech, also have the journalists' opinion which could affect the reader's perception.

the two nations. For the rest of the countries, we use the Consumer Confidence Index (Michigan Sentiment Index - U.S.).

2. Indirect Measure: Baker and Wurgler Index[14]

   Among the six variables for the BW Index (Dividend Premium, Turnover, IPO Frequency, IPO first-day return, Closed-End Fund Discount (CEFD) and, Share of Equity Issued), the data for all except "Debt Issued" is obtained from Bloomberg. Debt Issued variable for all five nations is taken from CEIC – CDM database.

3. EPUI Index: The Economic Policy Uncertainty Index (Baker et al. [2016], Arbatli et al. [2017]) is available on the official website https://www.policyuncertainty.com/.

5 Analysis and Results

We first look at the summary statistics for return as well as speech variables for all nations. Table 1 and Table 2 specify the speech statistics for each country. We get the speeches from the earliest period available for each countries' official website. The longest time period of availability is for the U.K. and Japan. The U.K. also has the highest number of daily speeches whereas Germany has the highest number of average speeches per month. Also, France has the lowest number of average speeches per month. The mean and median for daily as well as monthly sentiment are negative for all developed nations (although positive for China, an emerging market).

Table 3 below shows the Index and return statistics for each country. The average number of trading days is broadly the same for all nations except

[14]The BW Sentiment Index for UK, Japan, Germany, and France are replicated from the same process as Baker et al. [2012].

Table 1: Speech Statistics

Variable/Country	Time Period	Total Number of Speeches	Daily Speeches after combining for same day	No. of Positive Sentiment Speeches (D)	No. of Negative Sentiment Speeches (D)	Total no. of months Covered by speeches	No. of Positive Sentiment (M)	No. of Negative Sentiment (M)	Avg. No. of Speeches per month
USA	Jan 2006 - Feb 2020	797	693	146	547	165	21	144	4.1
UK	Apr 1996 - May 2020	1074	648	62	586	255	16	239	3.4
Japan	Apr 1996 - May 2020	696	637	178	459	239	58	181	2.2
France	Jan 2015 - May 2020	156	146	80	66	54	25	29	1.2
Germany	Jan 2012 - May 2020	599	480	130	350	99	13	86	4.7
India	May 2009 - Mar 2020	695	573	190	382	142	31	111	1.6
China	Feb 2002 - Apr 2020	295	280	189	90	151	107	44	1.1

Note: This table presents the summary statistics for speech frequency with respect to daily and monthly levels for the five nations. The data is obtained from official central bank website for each nation. The 4th column shows the number of speeches after combining all speeches in a day into one.

Table 2: Speech Sentiment Statistics

Country	Time Period	Mean (Daily)	Max(daily)	Min (Daily)	Mean (Monthly)	Max (Monthly)	Min (Monthly)
USA	Jan 2006 - Feb 2020	-0.0569	0.2974	-0.3497	-0.2716	0.4823	-1.2252
UK	Apr 1996 - May 2020	-0.0685	0.2889	-0.3655	-0.2633	0.2447	-1.1026
Japan	Apr 1996 - May 2020	-0.0372	0.2605	-0.3061	-0.1019	0.4293	-0.7853
France	Jan 2015 - May 2020	-0.0005	0.3607	-0.1924	-0.0036	0.3409	-0.2250
Germany	Jan 2012 - May 2020	-0.0307	0.1510	-0.3202	-0.1678	0.1446	-0.7561
India	May 2009 - Mar 2020	-0.0373	0.2218	-0.3576	-0.1527	0.2108	-0.7797
China	Feb 2002 - Apr 2020	0.0343	0.3379	-0.3188	0.0608	0.4754	-0.3640

Note: This table presents the summary statistics for speech sentiment with respect to daily and monthly levels for the five nations. The data is obtained from official central bank website for each nation. The daily variables are reported after combining all speeches on the same day into one.

Japan and China.[15]

Table 3: Index Return Statistics

Country	Main Index	Smallcap Index	Mean Return Main Index (Monthly)	Mean Return Smallcap Index (Monthly)	Mean Return Main Index (Daily)	Mean Return Smallcap Index (Daily)	Trading days per year
USA	DJIA Index	Rusell 2000	0.004670	0.005461	0.0002339	0.0003264	251
UK	FTSE Index	FTSE Smallcap	0.007686	0.003048	0.0000548	0.0001216	252
Japan	Nikkei Index	Topix Smallcap	0.000608	0.004111	0.0001451	0.0002373	245
France	CAC Index	CAC Smallcap	0.005495	0.005495	0.0000709	0.0001786	255
Germany	DAX Index	DAX Smallcap	0.003994	0.006870	0.0002279	0.0003205	253
India	Nifty Index	Nifty Smallcap	0.009854	0.008333	0.0004380	0.0004301	258
China	Shanghai Composite Index	CSI smallcap 500 Shanghai Index	0.005348	0.003146	0.0002696	0.0001111	242

Note: This table presents the summary statistics for return with respect to daily and monthly levels for the five nations. The data is obtained from Bloomberg for each nation.

Figures 4, 5, 6, 7 and 8 show the movement of speech sentiment and main index return across time for all five nations. It can be seen that for the U.S., U.K., and Japan, the variables are mimicking each other's movement closely. However, the same cannot be said for Germany and France as the movement is not synchronized (especially, for France). Also, in the case of the U.S., U.K., and Japan, the movement of both variables is mostly contemporaneous

[15] The reason for China having fewer trading days is due to frequent financial/trading lockdowns imposed by the government.

with speech sentiment leading at a few places.

Thus, we expect to see a significant relationship between speech sentiment and return for the U.S., U.K., and Japan. Although, the same cannot be expected for Germany and France.

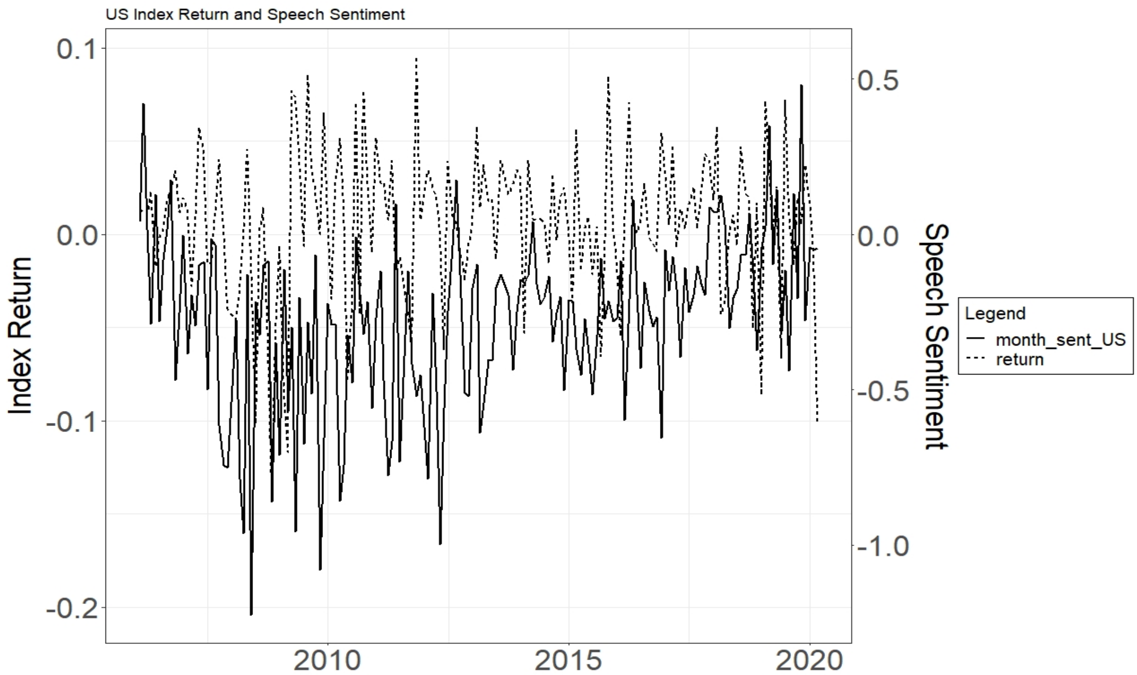


Figure 4: The monthly return (dotted line) is for the DJIA Index (U.S.) whereas the speech sentiment (solid line) is calculated by summing up the speeches over a month and then extracting sentiment using the specified methodolgy in this study. The return is represented by the primary Y axis and the speech sentiment by the secondary Y axis. This is because the scales are different for both variables.

5.1 Impact of Speech Sentiment

To verify the patterns in the graphs above, first, we do the daily analysis for each of the five nations.[16] The results are presented in Table 4. For the smallcap index, the speech sentiment affects the return with a lag of 2 days for US and UK; and with a lag of 3 days for Japan. However, there

[16] All coefficients reported in this study are HAC robust.

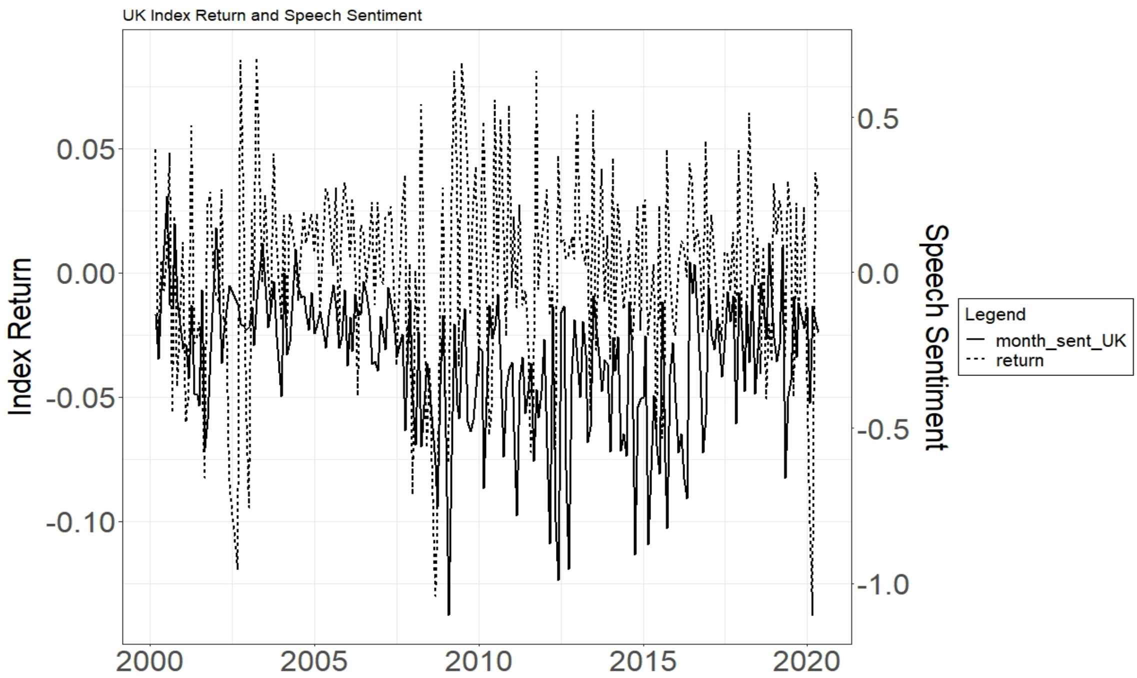


Figure 5: The monthly return (dotted line) is for the FTSE Index (U.K.) whereas the speech sentiment (solid line) is calculated by summing up the speeches over a month and then extracting sentiment using the specified methodolgy in this study. The return is represented by the primary Y axis and the speech sentiment by the secondary Y axis. This is because the scales are different for both variables.

is no significant result for France and Germany.[17] Also, for the main index, the results follow similar patterns except for the U.S. where the results are observed on the same day as the delivery of the speech.

Similarly, table 5 below sheds light on the monthly effect of speech sentiment for all five developed nations. The results for the US and UK are in line with the daily results. We find that speech sentiment affects the return for the same (immediate) month for these two nations for the smallcap index

[17]For Germany and France, since the results are insignificant for all lags (including 0) we present the results for the equation which is in line with significant results (the U.S. and/or the U.K. and/or Japan). This is done to ensure the comprehensibility of the regression tables.

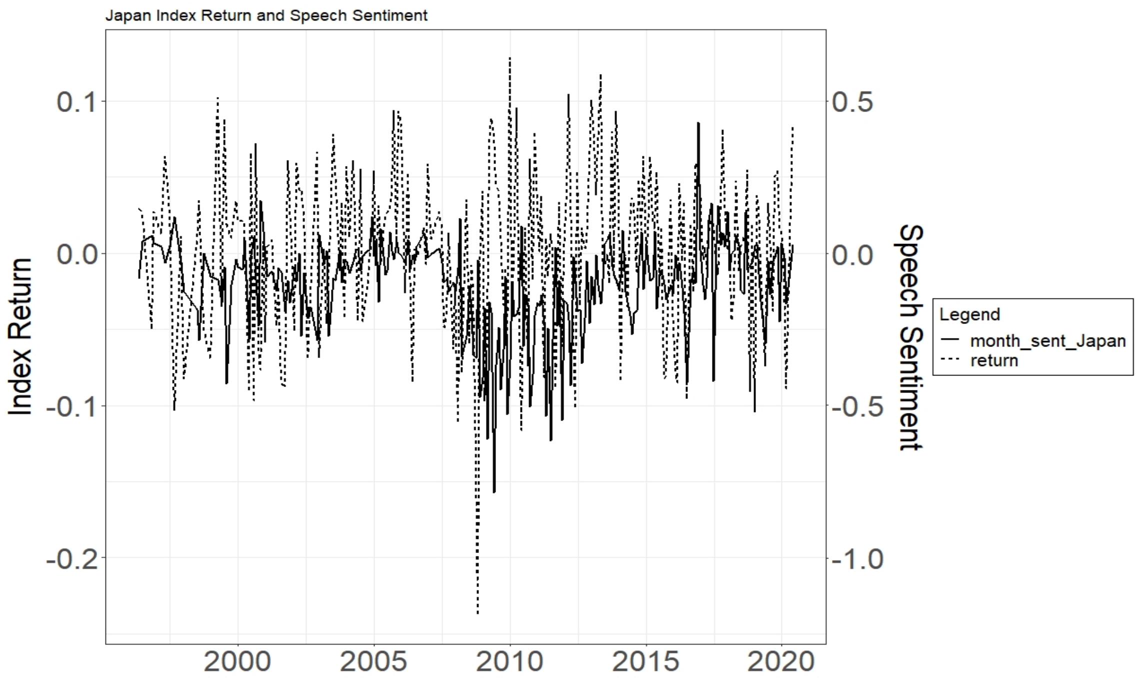


Figure 6: The monthly return (dotted line) is for the Nikkei Index (Japan) whereas the speech sentiment (solid line) is calculated by summing up the speeches over a month and then extracting sentiment using the specified methodolgy in this study. The return is represented by the primary Y axis and the speech sentiment by the secondary Y axis. This is because the scales are different for both variables.

and with a lag of 2 months for the U.S. for the main index. An important distinction here is that there are no significant results for Japan, the reason could be that in case of Japan, there are positive and negative sentiment speeches in the same month (for 84 out of 240 months) and as daily results are significant, we believe they could nullify the effect when looked at the monthly level. The results for Germany are in line with daily results, however, France smallcap Index shows significant results for the immediate month.

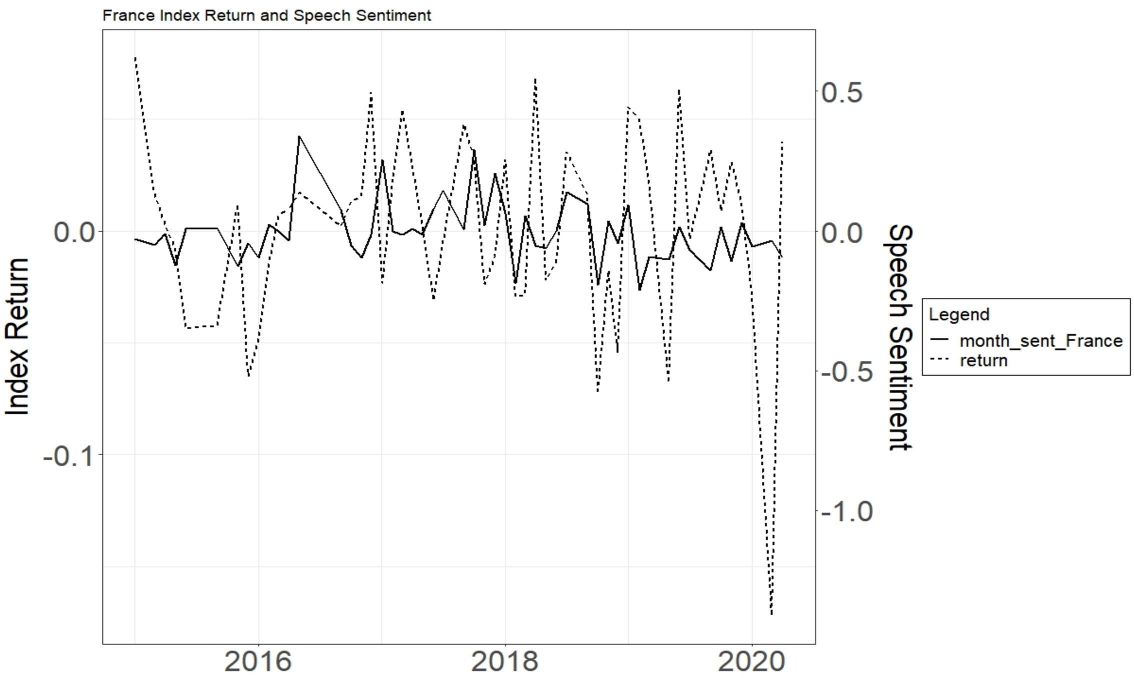


Figure 7: The monthly return (dotted line) is for the CAC Index (France) whereas the speech sentiment (solid line) is calculated by summing up the speeches over a month and then extracting sentiment using the specified methodolgy in this study. The return is represented by the primary Y axis and the speech sentiment by the secondary Y axis. This is because the scales are different for both variables.

5.2 Impact of Speech Sentiment - Additional Controls

Next, we check the impact of speech sentiment with existing sentiment variables as additional controls.[18] The results are presented in table 6 and 7.

We find that the results for speech sentiment remain the same. An important point to note in these results is that all three control variables (BW, EPUI, and CCI) are not significant in the presence of speech sentiment index. However, we checked the same in the absence of the speech variable (with the same specification) and they are a significant predictor of return. Hence,

[18]We do the analysis for all 3 lags and current month, however, we report the same as Table 5 for comprehensibility.

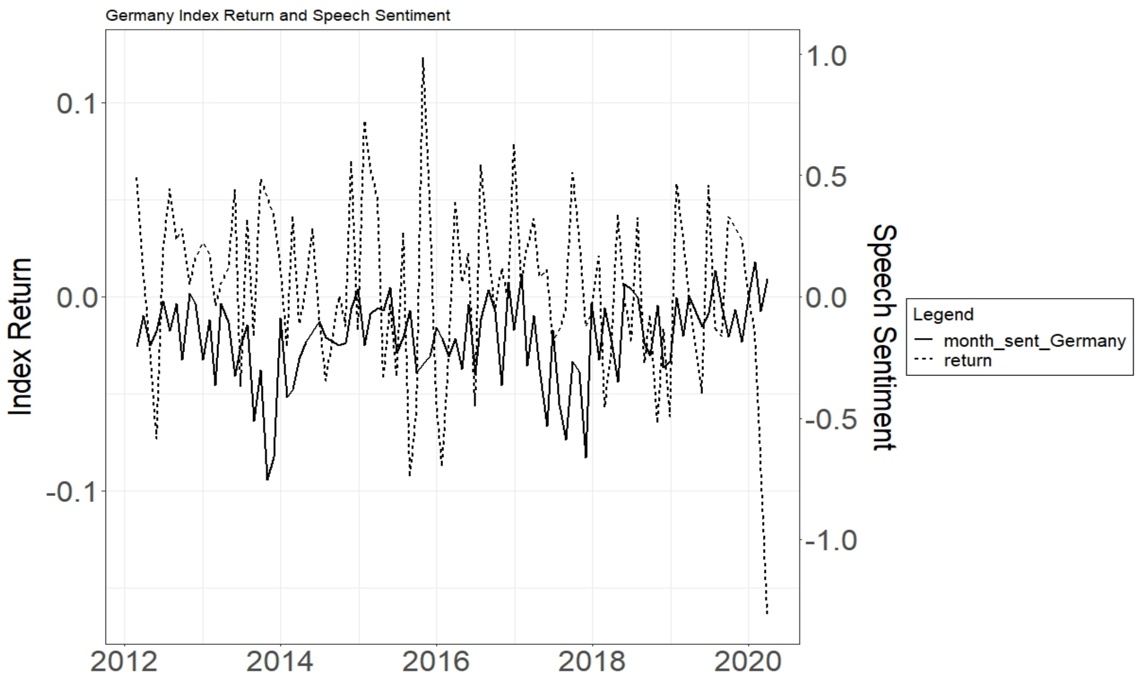


Figure 8: The monthly return (dotted line) is for the DAX Index (Germany) whereas the speech sentiment (solid line) is calculated by summing up the speeches over a month and then extracting sentiment using the specified methodolgy in this study. The return is represented by the primary Y axis and the speech sentiment by the secondary Y axis. This is because the scales are different for both variables.

it can be emphasized that the presence of sentiment derived from the speech of central banks renders the existing sentiment variables insignificant in U.S. and U.K. However, the same cannot be said in case of Japan, France and Germany as we find that the BW Index is significant in case of Japan and Germany and the Consumer Confidence Index in case of Japan and France.

5.3 Impact of Index Return

Also, to ensure that speeches are not impacted by the index returns, we analyze the impact of return on speech sentiment for both daily and monthly frequency. We find that for both main as well as smallcap index there is no

Table 4: Daily Analysis

	Smallcap Index		Main Index		
Country/Variable	Speech Sent Lag 2	Speech Sent Lag 3	Speech Sent Lag 0	Speech Sent Lag 2	Speech Sent Lag 3
USA	0.005*** (0.002)		0.008* (0.005)		
UK	0.008* (0.005)			0.017** (0.007)	
Japan		0.012* (0.007)			0.023** (0.009)
France	0.003 (0.011)		0.011 (0.015)		
Germany	−0.001 (0.007)		−0.009 (0.008)		
Controls	Yes	Yes	Yes	Yes	Yes

Note: This table presents the results from daily regression on speech sentiment for the smallcap and main index. The dependent variable is the daily return from the Smallcap and Main Index. The results are reported in line with equation 1. Thus, for each nation we present the lag for which the speech sentiment is significant. Thus, for smallcap index, the significance is observed at lag 2 for U.S.,U.K. and at lag 3 for Japan. For Germany and France none of the lags are significant (including lag 0), thus we present lag 2 to ensure comparability. The number of observations are the same as number of speech-days for each country. The coefficients reported are all Heteroskedasticity and Autocorrelation (HAC) robust. The controls include three lags of return, day of the week and month dummy. ***, ** and * indicate that the coefficient estimate are significantly different from zero at the 1 percent, 5 percent and 10 percent levels respectively.

significant impact of return on speech sentiment. We present the result for main index in Table 8.

6 Discussion of Results

We offer three possible explanations for the daily and monthly results. They are is explained in detail below:

6.1 Speeches with contrasting sentiments

For France and Japan, the reason could be opposing speeches in the same month (as specified in section 5.1). Thus, the presence of both positive as

Table 5: Monthly Analysis

Smallcap Index		Main Index	
Country/Variable	Speech Sent Lag 0	Speech Sent Lag 2	Speech Sent Lag 0
USA	-0.013^{*} (0.007)	0.019^{*} (0.010)	
UK	0.028^{**} (0.014)		0.024^{*} (0.013)
Japan	-0.004 (0.023)		-0.005 (0.023)
France	0.086^{*} (0.049)		0.055 (0.042)
Germany	-0.032 (0.023)		-0.022 (0.024)
Controls	Yes	Yes	Yes

Note: This table presents the results from monthly regression on speech sentiment for the smallcap and main index. The dependent variable is the monthly return from the Smallcap and Main Index. The results are reported in line with equation 1. Thus, for each nation we present the lag for which the speech sentiment is significant. The number of observation are the same as number of speech-months for each country.The coefficients reported are all Heteroskedasticity and Autocorrelation (HAC) robust. The controls include three lags of return. ***, ** and * indicate that the coefficient estimate are significantly different from zero at the 1 percent, 5 percent and 10 percent levels respectively.

well as negative speeches in the same month, for majority of months, could nullify the impact of daily speeches, rendering monthly results insignificant.[19]

6.2 Strength and Weight Argument

There is one theory that gives a possible explanation for the results of all five developed nations. The idea can be traced back to Griffin and Tversky [1992] attempt to reconcile conservatism (Edwards [1968]) and representativeness (Tversky and Kahneman [1974]). In Griffin and Tversky's framework, people update their beliefs based on the "strength" and the "weight" of the evidence (Barberis et al. [1998]). Griffin and Tversky use an example of a recommendation letter to explain both the attributes. The "strength" of the

[19]The number of months for which both positive and negative sentiment speeches are present in the same month are 84 (out of total 240) for Japan and 42 (out of total 54) for France.

Table 6: Monthly Analysis with Baker and Wurgler (BW) Index as Control

	Smallcap Index		**Main Index**		
Country/Variable	Speech Sent Lag 0	BW	Speech Sent Lag 2	Speech Sent Lag 0	BW
USA	-0.014^{*} (0.007)	0.001 (0.008)	0.024^{**} (0.011)		-0.020 (0.012)
UK	0.029^{**} (0.014)	0.001 (0.001)		0.028^{**} (0.014)	0.001 (0.001)
Japan	0.010 (0.023)	0.0002^{*} (0.0001)		-0.0015 (0.025)	0.0002 (0.0001)
France	0.079 (0.049)	-0.001 (0.001)		0.047 (0.044)	-0.001 (0.001)
Germany	-0.011 (0.019)	-0.001^{***} (0.001)		0.005 (0.018)	-0.001^{***} (0.001)
Controls	Yes	Yes	Yes	Yes	Yes

Note: This table presents the results from monthly regression on speech sentiment for the smallcap and main index. The dependent variable is the monthly return from the Smallcap and Main Index. The results are reported in line with equation 1. Thus, for each nation we present the lag for which the speech sentiment is significant. Also, the coefficient for BW is reported in accordance with the lag of the speech sentiment. The number of observation are the same as number of speech-months for each country.The coefficients reported are all Heteroskedasticity and Autocorrelation (HAC) robust. The controls include three lags of return and Baker Wurgler (BW) Index. ***, ** and * indicate that the coefficient estimate are significantly different from zero at the 1 percent, 5 percent and 10 percent levels respectively.

letter refers to how positive and warm its content is and the “weight”, on the other hand, measures the credibility and stature of the letter writer.

Both ideas are in accordance with the World Economic Forum’s Trustworthiness and Confidence Index (based on Soundness of Banks, Regulation of securities Exchange and Legal Rights Index) as shown in Figure 9.[20] It can be seen that Germany’s score has been steady or falling in most of the period of the analysis and is well below the U.S. and U.K. Similar trend can be noted in the case of France. However, the score for the U.S. and U.K. for their respective period has been steady or rising. Additionally, the financial market development index of the world economic forum shows a similar trend for all five nations. It can be implied from the above argument that market participants have placed low “weight” on the speeches in France and

[20] The data is only available till 2016.

Table 7: Monthly Analysis with Consumer Confidence Index (CCI) as Control

	Smallcap Index		**Main Index**		
Country/Variable	Speech Sent Lag 0	CCI	Speech Sent Lag 2	Speech Sent Lag 0	CCI
USA	-0.012^{*} (0.007)	0.091^{**} (0.036)	0.018^{*} (0.010)		-0.033 (0.075)
UK	0.030^{*} (0.016)	0.003 (0.003)		0.025^{*} (0.015)	-0.001 (0.003)
Japan	0.001 (0.023)	0.139 (0.108)		-0.0007 (0.024)	0.094 (0.123)
France	0.092^{*} (0.051)	-0.237 (0.224)		0.056 (0.043)	-0.056 (0.142)
Germany	-0.028 (0.021)	0.003 (0.003)		-0.018 (0.022)	0.003 (0.003)
Controls	Yes	Yes	Yes	Yes	Yes

Note: This table presents the results from monthly regression on speech sentiment for the smallcap and main index. The dependent variable is the monthly return from the Smallcap and Main Index. The results are reported in line with equation 1. Thus, for each nation we present the lag for which the speech sentiment is significant. Also, the coefficient for CCI is reported in accordance with the lag of the speech sentiment. The number of observation are the same as number of speech-months for each country.The coefficients reported are all Heteroskedasticity and Autocorrelation (HAC) robust. The controls include three lags of return and Consumer Confidence Index (CCI) . ***, ** 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 Return on Speech

	Daily Analysis			**Monthly Analysis**		
Country/Variable	Return Lag 1	Return Lag 2	Return Lag 3	Return Lag 1	Return Lag 2	Return Lag 3
USA	-0.143 (0.405)	-0.242 (0.305)	-0.227 (0.337)	0.513 (0.541)	0.331 (0.651)	-0.154 (0.629)
UK	-0.073 (0.233)	0.172 (0.196)	0.360 (0.236)	0.562 (0.423)	-0.134 (0.311)	0.051 (0.346)
Japan	-0.234 (0.241)	0.116 (0.180)	-0.315 (0.218)	0.148 (0.222)	0.241 (0.187)	0.342^{*} (0.174)
France	-0.488 (0.368)	1.386 (0.910)	0.651 (0.627)	-0.213 (0.332)	0.700^{*} (0.245)	0.271 (0.359)
Germany	0.415 (0.266)	0.243 (0.262)	-0.208 (0.257)	-0.047 (0.376)	0.149 (0.332)	0.129 (0.323)
Controls	Yes	Yes	Yes	Yes	Yes	Yes

Note: This table presents the results from daily and monthly regression on return for the main index. The dependent variable is the daily and monthly speech sentiment. The number of observation are the same as number of speech-days and speech-months for each country.The coefficients reported are all Heteroskedasticity and Autocorrelation (HAC) robust. ***, ** and * indicate that the coefficient estimate are significantly different from zero at the 1 percent, 5 percent and 10 percent levels respectively.

Germany even though the “Strength” is high and this could be the reason for statistical insignificance. On the other hand, for U.S., U.K., and Japan, the “weight” is so high that it has rendered the “Strength” to be less effective than it is intended to be. This could explain the positive sign of statistically significant coefficient in case of U.S., U.K. and Japan.[21]

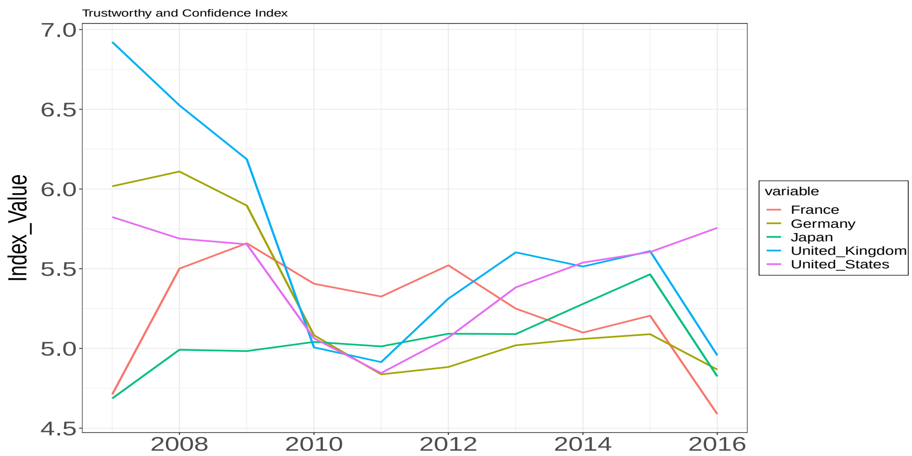


Figure 9: The figure presents the Trusthworthy and Confidence Index of the five nations from 2007 to 2016. Source : World Bank

6.3 Lost In Translation

For Germany, one important consideration for insignificant results can be that it is the only country in the sample for which the content had to be converted to English. This might have to lead to loss of meaning and hence incorrect quantification of sentiment. Also, for other non-English speaking nations, such as Japan and France, where the speeches were originally in the native language, a part of meaning might have been lost even though the official translation is provided by the respective central bank. To verify this conjecture further, we repeat the speech sentiment analysis for two English

[21] As the magnitude of speech sentiment is especially low in case of daily sentiment, its “strength” can be more prone to ignorance.

speaking nations, i.e. Canada and New Zealand and find that speech sentiment is a significant variable.[22] It is also verified in case of India, where speeches are given in English.

7 Robustness

7.1 Developed Market Features

It can be the case that the speech sentiment explains stock returns significantly only in the case of developed markets. In such a scenario, the results can be attributed to the economic, social, and technological advancements predominant in the developed markets. To ensure the results are not due to the characteristics of developed markets we analyze the impact of speeches in two of the largest and fastest-growing emerging economies i.e. India and China.

There have been very few studies that have formulated a Sentiment Index for emerging markets (Dash and Mahakud [2013], Kumari and Mahakud [2015] – India and Zhu and Niu [2016] – China), and none have done an exact replication of BW Index for Emerging Market. Thus, we only present the daily and monthly analysis for India and China without the Direct and Indirect sentiment controls. The results are presented in table 9.

The speech sentiment significantly affects return at lag 1 for India at both daily and monthly frequency (both smallcap as well as the main index). However, the results for China are not significant at any frequency for any of the indices.

There are two probable explanations for the insignificant in the speech

[22] The results are not presented to ensure brevity.

Table 9: Emerging Market Analysis

Smallcap Index		Main Index
Country/Variable	Speech Sent Lag 1	Speech Sent Lag 1
Daily Analysis		
India	-0.015^{***} (0.007)	-0.017^{***} (0.007)
China	-0.008 (0.014)	-0.006 (0.012)
Controls	Yes	Yes
Monthly Analysis		
India	-0.093^{**} (0.040)	-0.060^{**} (0.030)
China	-0.037 (0.079)	-0.023 (0.052)
Controls	Yes	Yes

Note: This table presents the results from daily and monthly regression on speech sentiment for the smallcap and main index with respect to the emerging markets. The dependent variable is the daily and monthly return from the Smallcap and Main Index respectively. The results are reported in line with equation 1. Thus, for each nation we present the lag for which the speech sentiment is significant.. The number of observation are the same as number of speech-days and speech-months for each country.The coefficients reported are all Heteroskedasticity and Autocorrelation (HAC) robust. The controls include three lags of return (for monthly analysis) and day of the week as well as the month dummy (for daily analysis) . ***, ** and * indicate that the coefficient estimate are significantly different from zero at the 1 percent, 5 percent and 10 percent levels respectively.

sentiment variable for China. Firstly, the day trading restrictions can be a possible explanation for the insignificant in the daily frequency. Also, the speeches are updated with a delay of anything between 3 to 20 days on the official website and the delayed date is also what is specified in the speeches. For the study, we manually check the event description in the speech and matched with the corresponding date from the web. Hence, due to the untimely dissemination of information, the effect might have been lost.

7.2 Economic Policy Uncertainty Index

Secondly, we check the robustness of our results in the presence of EPUI (Economic Policy Uncertainty Index). The index is based on news coverage about policy-related economic uncertainty. Since central bank communication is bound to make news in most circumstances, this is an important control variable. Thus, the EPUI is expected to cover the impact of speeches to a certain extent. The process is similar to BW and CCI as we add EPUI as an additional control variable. The results, presented in table 10 , are the same for U.S. and U.K., and additionally, we find that the speech sentiment is significant for the smallcap index in France and the main index in Germany.

Table 10: Monthly Analysis with EPUI Index as Control

	Smallcap Index		**Main Index**		
Country/Variable	Speech Sent Lag 0	EPUI	Speech Sent Lag 2	Speech Sent Lag 0	EPUI
USA	-0.012^{*} (0.007)	0.005 (0.006)	0.019^{*} (0.010)		-0.0002 (0.010)
UK	0.028^{*} (0.014)	-0.00001 (0.00002)		0.024^{*} (0.013)	-0.0001 (0.00001)
Japan	0.002 (0.023)	-0.041^{*} (0.022)		0.001 (0.022)	-0.064 (0.022)
France	0.081^{*} (0.048)	-0.053^{*} (0.029)		0.054 (0.040)	-0.041 (0.025)
Germany	-0.024 (0.022)	-0.040^{***} (0.009)		-0.017 (0.024)	-0.030^{***} (0.009)
Controls	Yes	Yes	Yes	Yes	Yes

Note: This table presents the results from monthly regression on speech sentiment for the smallcap and main index. The dependent variable is the monthly return from the Smallcap and Main Index. The results are reported in line with equation 1. Thus, for each nation we present the lag for which the speech sentiment is significant.Also, the coefficient for EPUI is reported in accordance with the lag of the speech sentiment. The number of observation are the same as number of speech-months for each country. The coefficients reported are all Heteroskedasticity and Autocorrelation (HAC) robust. The controls include three lags of return and Economic Policy Uncertainity Index (EPUI) . ***, ** and * indicate that the coefficient estimate are significantly different from zero at the 1 percent, 5 percent and 10 percent levels respectively.

8 Conclusion

The study attempts to quantify the sentiment from the speeches of the central bank of five developed nations (US, UK, Japan, France, and Germany). The method of sentiment quantification overcomes the existing drawbacks in the Loughran and McDonald Dictionary as well as the "bag-of-words" and ngram approach. The analysis is done for both daily and monthly frequency to ensure comparison with the existing direct (Consumer Confidence Index) and indirect sentiment variables (BW Index). It is found that speech sentiment significantly affects returns on both smallcap as well as the main Index for the US, UK, and Japan in case of daily frequency. Whereas the results are not significant for France and Germany. Also, for monthly frequency, the results are the same for the US and UK, however, the speech sentiment does not affect Index returns in Japan, due to the presence of negative and positive sentiment speeches in the same month (which nullify the impact). We also find that due to the unique properties of central bank speeches and the drawbacks in the existing sentiment measures, the speech sentiment turns out to be a better explanatory variable for index return. With respect to the explanation for the results, we have proposed three theories, nullifying effect for Japan and France; lost in translation argument for Germany and finally, the strength and weight argument for all five countries. All three can be investigated further in the future. For example, the methodology used in this study can be used in the native language for non-English speaking nations (such as Japan, Germany, and France) to resolve the lost in translation argument.

9 List of Valence Shifters

The tables 11 and 12 below specifies all the valence shifters used in this study including the 52 previously classified as stopwords in the LM dictionary.

Table 11: List of Valence Shifters

Word	Classification	Weight	Word	Classification	Weight
absolutely	Amplifier	0.8	despite all that	Adversative Conjunction	0.8
acute	Amplifier	0.8	despite all this	Adversative Conjunction	0.8
acutely	Amplifier	0.8	despite that	Adversative Conjunction	0.8
ain't	Negator	-1	despite this	Adversative Conjunction	0.8
aint	Negator	-1	didn't	Negator	-1
almost	De-amplifier	0.8	didnt	Negator	-1
although	Adversative Conjunction	0.8	doesn't	Negator	-1
aren't	Negator	-1	doesnt	Negator	-1
arent	Negator	-1	don't	Negator	-1
barely	De-amplifier	0.8	dont	Negator	-1
but	Adversative Conjunction	0.8	enormous	Amplifier	0.8
can't	Negator	-1	enormously	Amplifier	0.8
cannot	Negator	-1	especially	Amplifier	0.8
cant	Negator	-1	extreme	Amplifier	0.8
certain	Amplifier	0.8	extremely	Amplifier	0.8
certainly	Amplifier	0.8	faintly	De-amplifier	0.8
colossal	Amplifier	0.8	few	De-amplifier	0.8
colossally	Amplifier	0.8	greatly	Amplifier	0.8
considerably	Amplifier	0.8	hadn't	Negator	-1
couldn't	Negator	-1	hadnt	Negator	-1
couldnt	Negator	-1	hardly	De-amplifier	0.8
daren't	Negator	-1	hasn't	Negator	-1
darent	Negator	-1	hasnt	Negator	-1
decidedly	Amplifier	0.8	haven't	Negator	-1
deep	Amplifier	0.8	havent	Negator	-1
deeply	Amplifier	0.8	heavily	Amplifier	0.8
definite	Amplifier	0.8	heavy	Amplifier	0.8

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

Table 12: List of Valence Shifters

Word	Classification	Weight	Word	Classification	Weight	Word	Classification	Weight
high	Amplifier	0.8	needn't	Negator	-1	really	Amplifier	0.8
highly	Amplifier	0.8	neednt	Negator	-1	seldom	De-amplifier	0.8
however	Adversative Conjunction	0.8	neither	Negator	-1	serious	Amplifier	0.8
huge	Amplifier	0.8	never	Negator	-1	seriously	Amplifier	0.8
hugely	Amplifier	0.8	no	Negator	-1	severe	Amplifier	0.8
immense	Amplifier	0.8	nobody	Negator	-1	severely	Amplifier	0.8
immensely	Amplifier	0.8	none	Negator	-1	shan't	Negator	-1
incalculable	Amplifier	0.8	nor	Negator	-1	shant	Negator	-1
incalculably	Amplifier	0.8	not	Negator	-1	shouldn't	Negator	-1
incredibly	De-amplifier	0.8	only	De-amplifier	0.8	shouldnt	Negator	-1
isn't	Negator	-1	oughtn't	Negator	-1	significant	Amplifier	0.8
isnt	Negator	-1	oughtnt	Negator	-1	significantly	Amplifier	0.8
kind of	De-amplifier	0.8	particular	Amplifier	0.8	slightly	De-amplifier	0.8
kinda	De-amplifier	0.8	particularly	Amplifier	0.8	somewhat	De-amplifier	0.8
least	De-amplifier	0.8	partly	De-amplifier	0.8	sort of	De-amplifier	0.8
little	De-amplifier	0.8	purpose	Amplifier	0.8	sorta	De-amplifier	0.8
majorly	Amplifier	0.8	purposely	Amplifier	0.8	sparsely	De-amplifier	0.8
massive	Amplifier	0.8	quite	Amplifier	0.8	sporadically	De-amplifier	0.8
massively	Amplifier	0.8	rarely	De-amplifier	0.8	sure	Amplifier	0.8
mightn't	Negator	-1	real	Amplifier	0.8	surely	Amplifier	0.8
mightnt	Negator	-1	very	Amplifier	0.8	that being said	Adversative Conjunction	0.8
more	Amplifier	0.8	very few	De-amplifier	0.8	totally	Amplifier	0.8
most	Amplifier	0.8	very little	De-amplifier	0.8	true	Amplifier	0.8
much	Amplifier	0.8	wasn't	Negator	-1	truly	Amplifier	0.8
mustn't	Negator	-1	wasnt	Negator	-1	uber	Amplifier	0.8
mustnt	Negator	-1	weren't	Negator	-1	vast	Amplifier	0.8
whereas	Adversative Conjunction	0.8	wont	Negator	-1	wouldnt	Negator	-1
won't	Negator	-1	wouldn't	Negator	-1	werent	Negator	-1

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

References

Werner Antweiler and Murray Z Frank. Is all that talk just noise? the information content of internet stock message boards. *The Journal of Finance*, 59(3):1259–1294, 2004.

Ms Elif C Arbatli, Steven J Davis, Arata Ito, Naoko Miake, and Ikuo Saito. *Policy Uncertainty in Japan*. International Monetary Fund, 2017.

Salman Arif and Charles MC Lee. Aggregate investment and investor sentiment. *The Review of Financial Studies*, 27(11):3241–3279, 2014.

David Aristei and Duccio Martelli. Sovereign bond yield spreads and market sentiment and expectations: Empirical evidence from euro area countries. *Journal of Economics and Business*, 76:55–84, 2014.

Malcolm Baker and Jeremy C Stein. Market liquidity as a sentiment indicator. *Journal of Financial Markets*, 7(3):271–299, 2004.

Malcolm Baker and Jeffrey Wurgler. Investor sentiment and the cross-section of stock returns. *The Journal of Finance*, 61(4):1645–1680, 2006.

Malcolm Baker and Jeffrey Wurgler. Investor sentiment in the stock market. *Journal of Economic Perspectives*, 21(2):129–152, 2007.

Malcolm Baker, Jeffrey Wurgler, and Yu Yuan. Global, local, and contagious investor sentiment. *Journal of Financial Economics*, 104(2):272–287, 2012.

Scott R Baker, Nicholas Bloom, and Steven J Davis. Measuring economic policy uncertainty. *The Quarterly Journal of Economics*, 131(4):1593–1636, 2016.

Nicholas Barberis, Andrei Shleifer, and Robert Vishny. A model of investor sentiment. *Journal of Financial Economics*, 49(3):307–343, 1998.

Nittai K Bergman and Sugata Roychowdhury. Investor sentiment and corporate disclosure. *Journal of Accounting Research*, 46(5):1057–1083, 2008.

Anchada Charoenrook. Does sentiment matter.

Hailiang Chen, Prabuddha De, Yu Jeffrey Hu, and Byoung-Hyoun Hwang. Wisdom of crowds: The value of stock opinions transmitted through social media. *The Review of Financial Studies*, 27(5):1367–1403, 2014.

Saumya Ranjan Dash and Jitendra Mahakud. Impact of investor sentiment on stock return: evidence from india. *Journal of Management Research*, 13(3):131–144, 2013.

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

Ward Edwards. Conservatism in human information processing. *Formal Representation of Human Judgment*, 1968.

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

Joseph Engelberg. Costly information processing: Evidence from earnings announcements. In *AFA 2009 San Francisco meetings paper*, 2008.

Diego Garcia. Sentiment during recessions. *The Journal of Finance*, 68(3): 1267–1300, 2013.

Matthew Gentzkow, Jesse M Shapiro, and Matt Taddy. Measuring group differences in high-dimensional choices: method and application to congressional speech. *Econometrica*, 87(4):1307–1340, 2019.

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

Robin Greenwood and Andrei Shleifer. Expectations of returns and expected returns. *The Review of Financial Studies*, 27(3):714–746, 2014.

Dale Griffin and Amos Tversky. The weighing of evidence and the determinants of confidence. *Cognitive Psychology*, 24(3):411–435, 1992.

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

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.

W Jos Jansen and Niek J Nahuis. The stock market and consumer confidence: European evidence. *Economics Letters*, 79(1):89–98, 2003.

Narasimhan Jegadeesh and Di Wu. Word power: A new approach for content analysis. *Journal of Financial Economics*, 110(3):712–729, 2013.

Colm Kearney and Sha Liu. Textual sentiment in finance: A survey of methods and models. *International Review of Financial Analysis*, 33:171–185, 2014.

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

Soon-Ho Kim and Dongcheol Kim. Investor sentiment from internet message postings and the predictability of stock returns. *Journal of Economic Behavior & Organization*, 107:708–729, 2014.

Taehyuk Kim and Aejin Ha. Investor sentiment and market anomalies. In *23rd Australasian Finance and Banking Conference*, 2010.

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

Jyoti Kumari and Jitendra Mahakud. Does investor sentiment predict the asset volatility? evidence from emerging stock market india. *Journal of Behavioral and Experimental Finance*, 8:25–39, 2015.

Charles MC Lee, Andrei Shleifer, and Richard H Thaler. Investor sentiment and the closed-end fund puzzle. *The Journal of Finance*, 46(1):75–109, 1991.

Wayne Y Lee, Christine X Jiang, and Daniel C Indro. Stock market volatility, excess returns, and the role of investor sentiment. *Journal of Banking & Finance*, 26(12):2277–2299, 2002.

Michael Lemmon and Evgenia Portniaguina. Consumer confidence and asset prices: Some empirical evidence. *The Review of Financial Studies*, 19(4): 1499–1529, 2006.

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

Feng Li. The information content of forward-looking statements in corporate filings—a naïve bayesian machine learning approach. *Journal of Accounting Research*, 48(5):1049–1102, 2010.

Tsai-Ling Liao, Chih-Jen Huang, and Chieh-Yuan Wu. Do fund managers herd to counter investor sentiment? *Journal of Business Research*, 64(2): 207–212, 2011.

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.

Tim Loughran and Bill McDonald. The use of word lists in textual analysis. *Journal of Behavioral Finance*, 16(1):1–11, 2015.

Tim Loughran and Bill McDonald. Textual analysis in accounting and finance: A survey. *Journal of Accounting Research*, 54(4):1187–1230, 2016.

Robert Neal and Simon M Wheatley. Do measures of investor sentiment predict returns? *Journal of Financial and Quantitative Analysis*, pages 523–547, 1998.

Maria Ward Otoo. Consumer sentiment and the stock market. 1999.

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

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. Investor sentiment and stock returns: Some international evidence. *Journal of Empirical Finance*, 16(3):394–408, 2009.

Marc Schulder, Michael Wiegand, Josef Ruppenhofer, and Stephanie Köser. Introducing a lexicon of verbal polarity shifters for english. 2018.

David P Simon and Roy A Wiggins III. S&p futures returns and contrary sentiment indicators. *Journal of Futures Markets: Futures, Options, and Other Derivative Products*, 21(5):447–462, 2001.

David H Solomon, Eugene Soltes, and Denis Sosyura. Winners in the spotlight: Media coverage of fund holdings as a driver of flows. *Journal of Financial Economics*, 113(1):53–72, 2014.

Spyros Spyrou. Sentiment changes, stock returns and volatility: evidence from nyse, amex and nasdaq stocks. *Applied Financial Economics*, 22(19): 1631–1646, 2012.

Wen-Ming Szu, Yi-Chen Wang, and Wan-Ru Yang. How does investor sentiment affect implied risk-neutral distributions of call and put options? *Review of Pacific Basin Financial Markets and Policies*, 18(02):1550010, 2015.

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.

Amos Tversky and Daniel Kahneman. Judgment under uncertainty: Heuristics and biases. *Science*, 185(4157):1124–1131, 1974.

Chunpeng Yang and Bin Gao. The term structure of sentiment effect in stock index futures market. *The North American Journal of Economics and Finance*, 30:171–182, 2014.

Chunpeng Yang and Liyun Zhou. Investor trading behavior, investor sentiment and asset prices. *The North American Journal of Economics and Finance*, 34:42–62, 2015.

Bo Zhu and Feng Niu. Investor sentiment, accounting information and stock price: Evidence from china. *Pacific-Basin Finance Journal*, 38:125–134, 2016.

Mohamed Zouaoui, Genevieve Nouyrigat, and Francisca Beer. How does investor sentiment affect stock market crises? evidence from panel data. *Financial Review*, 46(4):723–747, 2011.