Distribution Agreement

In presenting this thesis as a partial fulfillment of the requirements for a degree from Emory University, I hereby grant to Emory University and its agents the non-exclusive license to archive, make accessible, and display my thesis in whole or in part in all forms of media, now or hereafter now, including display on the World Wide Web. I understand that I may select some access restrictions as part of the online submission of this thesis. I retain all ownership rights to the copyright of the thesis. I also retain the right to use in future works (such as articles or books) all or part of this thesis.

Wenhui (Jessie) Zhao April 1, 2019

Federal Reserve Communication and Its Implications on Inflation Expectations

by

Wenhui (Jessie) Zhao

Tao Zha
Adviser

Economics

Tao Zha

Adviser

Vivian Zhanwei Yue

Committee Member

Karen Ton

Committee Member

2019

Federal Reserve Communication and Its Implications on Inflation Expectations

By

Wenhui (Jessie) Zhao

Tao Zha

Adviser

An abstract of
a thesis submitted to the Faculty of Emory College of Arts and Sciences
of Emory University in partial fulfillment
of the requirements of the degree of
Bachelor of Arts with Honors

Economics

2019

Abstract

Federal Reserve Communication and Its Implications on Inflation Expectations
By Wenhui (Jessie) Zhao

In this study, I investigate the relationship between the 12-month ahead inflation expectations and the tone used in the Federal Open Market Committee (FOMC) statements from the past decade (January 2009 – December 2018). The tone index is built using two different approaches: the dictionary method with a phrase-based lexicon and the Structural Topic Model. This index is then incorporated as one of the independent variables in regressions, along with the macroeconomic variables. I demonstrate that there is a positive relationship between inflation expectations and the tone of the FOMC statements. Throughout the last decade, an overall increasing trend of hawkishness in the FOMC statements is discovered; in addition, a hawkish tone is accompanied by the FOMC committee's decision of increasing the target federal funds rate. It is also worth noting that the results generated from both approaches strongly support the conclusions mentioned above.

Federal Reserve Communication and Its Implications on Inflation Expectation

By

Wenhui (Jessie) Zhao

Tao Zha

Adviser

A thesis submitted to the Faculty of Emory College of Arts and Sciences
of Emory University in partial fulfillment
of the requirements of the degree of
Bachelor of Arts with Honors

Economics

2019

Acknowledgements

I would like to express my gratitude to my advisors, Joshua C. Fjelstul, and Dr. Robert O'Reilly from Emory Center for Digital Scholarship for their extensive support and encouragement. Additionally, I am very thankful for Dr. Marco Vega from the Central Reserve Bank of Peru and Luciana Figueroa for their inspirations and feedback on this project.

Table of Contents

I. Introduction

Language is a vehicle that provides a source of communication as well as self-expression, and more specifically, a way to understand others and be understood. People often talk about language in a linguistic context, limiting themselves to each other's word choices, syntax, and tones used in the speech. Yet, it has become more widely accepted that language can be analyzed in a multitude of other contexts, such as political sciences, sociology, etc. In this study, I explore the question of what language means in an economics context. In particular, I analyze the tone – hawkish or dovish – used in the Federal Open Market Committee (FOMC) statements published by the Federal Reserve, and investigate its connections to inflation expectations, which is one of the most important macroeconomic variables that concern the general public.

In order to extract the tone used in each statement, I apply the techniques of text mining and sentiment analysis, which are typically used in computational linguistics research to analyze the use of natural language processing. However, both techniques have been gradually adopted in economics research during recent years to incorporate textual variables. Many previous literature has analyzed the communication of the press releases from major central banks around the world. For example, Kahveci and Odabaş (2016) focus on whether there is a change in tone along with a greater transparency of the communication within the Federal Reserve, European Central Bank, and the Central Bank of the Republic of Turkey. In addition to analyzing the content itself, economists have studied the impact of the central bank's communication on markets and macroeconomic variables; for instance, Rohlfs et al. (2016) focus on daily interest rate movements, and Grimaldi (2011) measures the financial stress. However, no such studies speak to the impact of communication on inflation expectations, probably one of the most intuitive variables that people use to gauge the state of a country's economy – whether there is a

solid growth or a depressive recession. Therefore, this study fills the literature gap by investigating how the tone (i.e. categorized in either hawkish or dovish) delivered through the FOMC statements affects population's inflation expectations.

Different from most methodologies described in previous literature that use established online databases, my approach is to build my own tone-indicating lexicon by reading through all the FOMC statements and manually categorizing phrases to capture the tone in a more comprehensive sense. In addition to the subjective dictionary method, I adopt another approach to objectively extract the textual themes from all the statements using the Structural Topic Model, a variant of the Latent Dirichlet Allocation. A tone index is built in both approaches and incorporated as one of the independent variables along with macroeconomic variables in regression models. The data needed for this analysis can be found and downloaded from the Federal Reserve website and other economics databases.

My hypothesis is the following: there is a direct relationship between the tone used in the FOMC statements and the 12-month ahead inflation expectations. In other words, an overall hawkish tone indicates an increase in inflation expectations, and vice versa. The reason I adopt two approaches to capture the tone is to conduct a comparative analysis. If the results are similar, then a firm conclusion can be reached. On the flip side, if the results are in opposite directions, then it is highly possible that there are some issues in the lexicon-building process due to subjectivity. It may therefore infer that the objective approach provides more reliable analysis.

Taken together, the results derived from this study intend to fill the current literature gap and mainly contribute to the literature that studies the communication through the Federal Reserve, its monetary policy stance, and the general impact on macroeconomic variables. Since the inflation expectations will directly affect consumer's behaviors in the current period,

researchers who are interested in Behavioral Economics could take a psycho-behavioral approach that incorporates textual variables to analyze this relationship. Overall, this study aims to bring researchers a new angle and to consider using a cross-disciplinary approach to evaluate economic problems. As more economists have started to realize the importance and uniqueness of the textual variables, it is certain that integrating these variables into research will bring important perspectives to future research.

The rest of the paper is organized as follows: Section II summarizes the main findings and methodologies used in previous literature to study the central bank's communication and its impact on the general economy. Section III discusses the data used in this study. Section IV describes two approaches I employ to quantify the relationship between the tone used in the FOMC statements and inflation expectations. Section V presents the regression outputs and graphs, followed by a detailed discussion of results in section VI. Section VII provides a concluding remark and suggestions for future research topics.

II. Review of Literature

There has been an emerging amount of research that employs textual variables to study the effect of the central bank's communication. As mentioned in Haan et al. (2007), central banks affect the overall economy primarily through influencing the market expectations of the future path of core variables, such as overnight interest rates. Therefore, communication through the press releases or official publications helps central banks to achieve their goals. Textual analysis has also quickly developed into one of the heated research topics for economists due to the following two reasons: communication is a direct way to influence expectations, and thus the overall economy can be affected; communication is used to reduce uncertainties in financial

markets and increase transparency (Haan et al. 2007). In a similar scope of interest, Rohlfs et al. (2016) measure the effects of FOMC information delivery and its potential guided directions of daily movements on both short-term and medium-term interest rate. Based on the empirical results derived from the study, Rohlfs et al. (2016) obtain 93% and 64% of accuracies when predicting the target federal funds rate and effective rate respectively.

Most academic literature utilizes text mining techniques to quantitatively account for the central bank's communication, which has been demonstrated to be one of the most essential factors that influence the markets and macroeconomic variables. For example, Lucca and Trebbi (2009) present an objective scoring technique that measures the content of the central bank's communication through analyzing the FOMC statements starting in 1999 to capture future interest rate decisions. The study concludes that the content of the statements and the changes in communication measured by semantic scores are significant in determining longer-term nominal yields (Lucca and Trebbi 2009). The FOMC statements are also proved to contain important information about both the predicted and the residual component of policy rate decisions. The importance of textual variables and semantic measures are analyzed not only through the FOMC statements, but also through the official press releases from other major central banks. For example, Kazinnik (2017) employs dictionary-based measures and the Wordscore algorithm to investigate whether these sentiment measures have a direct implication on the direction of Bank of Israel's future monetary policy. She discovers that incorporating these sentiment measures, along with other macroeconomic variables, advances the prediction of short-term policy interest rate, which explains an additional 30% of the variances (Kazinnik 2017). Kazinnik (2017) also suggests expanding the dictionary (i.e. lexicon) to further improve the accuracy of the sentiment indicator, which essentially inspires me to build my own lexicon of including not only single

words, but also phrases and expressions that capture sentiments. In addition, Kahvecia and Odabaş (2016) conduct a semantic research that examines statements released from the Federal Reserve, the European Central Bank (ECB), and the Central Bank of the Republic of Turkey (CBRT) starting from 2001, focusing on the relationship between the change in tone and the central bank's movement toward transparency in communication before and after the Financial Crisis in 2008. They demonstrate that the Federal Reserve incorporates a relatively certain tone about the economy through the forward guidance statement placed toward the end of each press release; the tone of certainty has become more significant transferring from the pre-crisis to post-crisis period (Kahvecia & Odabaş 2016). In the case of ECB and CBRT, there are no pronounced changes in tone; however, only a slight increase in tone (i.e. more certainties) is discovered in the statements of CBRT (Kahvecia & Odabaş 2016). Thus, it is clear that the tone embedded in each statement and its change over time entail substantial information to anticipate the central banks' future policy stance.

In terms of the quantitative approach used to extract the textual themes and construct the tone index, Hansen and McMahon (2016) as well as Mazis and Tsekrekos (2017) represent some mainstream algorithms widely used in current literature. In Hansen and McMahon (2016), they explore the impact of FOMC information delivery on both market and real economic variables by applying the Latent Dirichlet Allocation (LDA; Blei 2003) within a Factor-Augmented Vector Autoregressive (FAVAR) framework. Simply put, FAVAR is a two-step approach that uses principle components to estimate factors. In this case, the estimated variables are F_t that captures the unobserved factors and Y_t that can be observed, both of which are assumed to drive the economy. Mazis and Tsekrekos (2017) use the Latent Semantic Analysis (LSA), which is believed to be a precursor of the LDA, to identify the main themes retained in the FOMC

statements between 2003 and 2014. They find that "multiple, multifaceted and recurring" textual themes are integrated in the FOMC statements; these themes are all statistically significant in explaining 3-month, 2-year, 5-year, and 10-year treasury yields (Mazis and Tsekrekos 2017). The LSA is efficient in identifying significant textual themes within statements, and its algorithm is constructed to be objective and free of bias. Another popular model adopted by researchers is the Structural Topic Model (STM), which builds off of the LDA (Roberts et al. forthcoming). The STM is proved to be useful in Roberts et al. (2016) that studies the media coverage of China's rise using multiple news sources. Although not frequently applied in economics research, the STM is shown to provide more accurate estimation when compared to the LDA (Roberts et al. 2014). In their study analyzing the open-ended survey responses, Roberts et al. (2014) show that the LDA tends to weaken covariate relationships on topical prevalence, while the STM provides a much better output. In addition, Roberts et al. (forthcoming) indicate that the STM allows researchers to incorporate metadata, i.e. information embedded in each document, into the model.

It is not hard to notice that there is an increasing number of literature that studies the effects of textual themes as a proxy for the central bank's communication to the public. Most literature chooses the central bank's official press releases as the main subject of matter, mainly because these releases serve as the most reliable vehicles for the public to gauge the general state of economy. Studies have presented the relationship between the textual themes or sentiment measures embedded in the official central bank statements and interest rate movements, future monetary policy, market returns, etc. However, I have not found any literature that considers the impact of the central bank's communication and its influence on population's inflation expectations, which in my opinion are the most straightforward and easy-to-understand

measures; even a layperson would understand the economic meaning behind the inflation expectations. Because of this gap in literature, I decide to conduct a study that primarily examines this relationship. As mentioned earlier, I will construct my own lexicon to complement the online databases that only entail single words in order to improve accuracies, as suggested in Kazinnik (2017). In addition, I will adopt the STM presented in Roberts et al. (forthcoming) as my objective approach. Further details of each method are presented in section IV.

III. Data

The FOMC statements from January 2009 to December 2018 are downloaded from the Federal Reserve's website in their html format, with a total of 80 statements for analysis. I purposely avoid the 2008 Financial Crisis, intending to eliminate some of the noises potentially originated from the sensitive period.

The dependent variable of all regressions is the mean of the 12-month ahead inflation expectations. These data are directly extracted from the Consensus Economics survey, which is sent to a multitude of analysts and financial institutions monthly to obtain their forecasts and perspectives. Other reasonable alternatives include the University of Michigan's monthly inflation forecasts (FRED code: MICH) and the Survey of Professional Forecasters distributed by the Federal Reserve Bank of Philadelphia.

The independent variables include a set of macroeconomic variables (i.e. real GDP, CPI, and target federal funds rate), the lagged variables, the total word count of each statement, and interaction terms. Real GDP and CPI are downloaded from the Federal Reserve Economic Data (FRED) website. Since the real GDP (FRED code: A191RL1Q225SBEA) is measured quarterly in its percent change from preceding period, I assume the real GDP remains constant throughout

all three months within a quarter. As for the CPI, I use its monthly growth rate that is composed of total all items for the U.S. (FRED code: CPALTT01USM657N); this dataset is not seasonally adjusted to better reflect the true price changes over the months. The target federal funds rate is manually extracted from each statement and compiled into an Excel file. The lagged variables can be easily constructed using the built-in R functions. Lastly, I also integrate a few interaction terms, such as total word count with target federal funds rate and tone index with target federal funds rate. The goal is to indicate that the effect of one independent variable on the inflation expectations is different in the presence of the other independent variable within the interaction term.

IV. Methodology

The programming tool used in this study is R because of its multiple built-in programs for text mining and sentiment analysis. I adopt two different approaches to build the tone index, and the regression outputs derived from both approaches are compared and analyzed to reach a conclusion. The first approach is to construct my own lexicon containing tone-indicating phrases. The second approach is to apply the STM algorithms and extract the top themes embedded in each statement. Then, a traditional Ordinary Least Squares (OLS) model is adopted to run multiple regressions.

i. Approach I: dictionary method

Instead of using the established online databases that only extract single words into an either hawkish or dovish category, I build my own lexicon by reading through all 80 statements and extracting tone-indicating phrases. I take this relatively time-consuming approach, because single words cannot precisely capture the tone reflected in each statement. Then, I build an R

program to match the phrases in the lexicon with the main text of each statement to get a total count of both hawkish and dovish phrases in each statement. After that, a naïve classifier of monetary policy stance, called *indposmon*, is built such that,

$$\text{indposmon} = \frac{H - D}{H + D}$$

where H is the total number of hawkish phrases appeared in each statement, whereas D is the total number of dovish phrases.

Readers can argue that the lexicon may vary among different researchers, depending on their selection criteria of hawkish versus dovish words and phrases. Thus, the study bears the risk of subjectivity and can possibly lead to an inaccuracy of test results and conclusions. To address this issue and reach a more persuasive conclusion, I adopt the STM algorithms known for their objectivity as my second approach. The detailed descriptions are presented as below.

ii. Approach II: STM

I closely follow the steps of the STM algorithms described in Roberts et al. (forthcoming). As mentioned earlier, R has many built-in packages, such as "quanteda", "topicmodels", and "stm", that can be used directly to run the model.

I will not delve deep into the mathematical derivations of the STM algorithms in this paper; however, a more comprehensive review can be found in Roberts et al. (2016). The input of the model is a corpus of all 80 statements categorized by the date. It is important to pre-process the statements by eliminating the stop words, punctuations, separators, and symbols, as well as initiating the stemming process. Once the corpus is constructed, I create a document-feature matrix to describe the frequency of terms that occur in each statement. The rows of the matrix correspond to each document in the corpus, while the columns correspond to each term. Another input is a number of topics formed by the STM algorithms; I use a two-topic model due

to a relatively small dataset in this study. The top words generated from the STM reflect either hawkishness or dovishness of the overall tone in each statement. The model forms the θ (theta) parameter, known as the topic mixing proportions (Roberts et al. 2016), that captures the percentage of each statement associated with an either hawkish or dovish topic. For example, it might suggest that a statement is 3% hawkish while 97% dovish. This measure is included as one of the independent variables in my regressions, similar to the naïve classifier *indposmon* constructed in the dictionary method.

**The following sections – incorporation of macroeconomic variables and regressions – are applied in both approaches.

The lagged variables of the inflation expectations, real GDP, and CPI are built using the raw data downloaded from the Consensus Economics survey and the FRED database. After pre-processing the macroeconomic variables, multiple regressions are performed to explain inflation expectations as a function of the policy rate and the tone index. In addition, multiple graphs of interest are presented in the next section, including the graphs of tone index versus dates, FOMC committee's decision on target federal funds rate versus tone index, and 12-month ahead inflation expectations versus tone index.

V. Results

i. Miscellany

Variable Name	Description
ec.nexty	12-month ahead inflation expectations (collected from Consensus Economics survey)
ec.nexty.l1	lagged inflation expectations

tasaref	target federal funds rate
dp. l1	lagged CPI (1-period prior)
gdp.realtime.l2	lagged real GDP (2-period prior)
nword	number of words in a statement
indposmon	naïve classifier of monetary policy stance
tasaref:nword	interaction term of target federal funds rate and number of words
tasaref:indposmon	interaction term of target federal funds rate and naïve classifier

Table 1. A list of abbreviated variables used in regressions and their respective interpretation.

ii. Regressions

The dependent variable is the mean of the 12-month ahead inflation expectations in all regressions.

	Dependent variable:		
	ec.nexty		
	(1)	(2)	(3)
tasaref	0.054	0.053	0.011
	(0.034)	(0.039)	(0.134)
ec.nexty.l1	0.670***	0.671***	0.667***
	(0.086)	(0.088)	(0.089)
dp.l1	0.005	0.004	0.005
	(0.063)	(0.064)	(0.064)
gdp.realtime.l2	-0.018*	-0.018*	-0.017
	(0.010)	(0.010)	(0.010)
nword		-0.00001	-0.00004
		(0.0001)	(0.0002)
tasaref:nword			0.0001
			(0.0003)
Constant	0.672***	0.675***	0.696***
	(0.162)	(0.174)	(0.187)
Observations	78	78	78
R2	0.550	0.550	0.551
Adjusted R2	0.525	0.519	0.513
Residual Std. Error	0.156 (df = 73)	0.157 (df = 72)	0.158 (df = 71)
F Statistic	22.289*** (df = 4; 73)	17.588*** (df = 5; 72)	14.492*** (df = 6; 71)
Note:			*p<0.1; **p<0.05; ***p<0.01

Figure 1. The independent variables are mostly macroeconomic variables except for the total word count of each statement.

(1) Approach I: dictionary method

	Dependent variable:		
	ec.nexty		
	(1)	(2)	(3)
tasaref	0.054	0.024	0.064
	(0.034)	(0.036)	(0.060)
ec.nexty.l1	0.670***	0.610***	0.596***
	(0.086)	(0.087)	(0.089)
dp.l1	0.005	0.013	0.016
	(0.063)	(0.061)	(0.062)
gdp.realtime.l2	−0.018*	−0.022**	−0.022**
	(0.010)	(0.010)	(0.010)
indposmon		0.126**	0.156**
		(0.054)	(0.066)
tasaref:indposmon			−0.076
			(0.093)
Constant	0.672***	0.811***	0.824***
	(0.162)	(0.169)	(0.170)
Observations	78	78	78
R2	0.550	0.581	0.585
Adjusted R2	0.525	0.552	0.550
Residual Std. Error	0.156 (df = 73)	0.151 (df = 72)	0.152 (df = 71)
F Statistic	22.289*** (df = 4; 73)	19.984*** (df = 5; 72)	16.688*** (df = 6; 71)
Note:			*p<0.1; **p<0.05; ***p<0.01

Figure 2. In addition to macroeconomic variables, the tone index is also incorporated as one of the independent variables in regression (2) and (3). The variable *indposmon* is statistically significant at the 95% level in both regressions.

(2) Approach II: STM

	Dependent variable:		
	ec.nexty		
	(1)	(2)	(3)
tasaref	0.054	0.017	6.521
	(0.034)	(0.039)	(74.630)
ec.nexty.l1	0.670***	0.606***	0.605***
	(0.086)	(0.091)	(0.092)
dp.l1	0.005	0.038	0.038
	(0.063)	(0.064)	(0.065)
gdp.realtime.l2	-0.018*	-0.018*	-0.018*
	(0.010)	(0.010)	(0.010)
hawkish		0.102*	1.732
		(0.054)	(18.704)
tasaref:hawkish			-6.521
			(74.827)
Constant	0.672***	0.767***	-0.858
	(0.162)	(0.167)	(18.642)
Observations	78	78	78
R2	0.550	0.571	0.571
Adjusted R2	0.525	0.541	0.535
Residual Std. Error	0.156 (df = 73)	0.153 (df = 72)	0.154 (df = 71)
F Statistic	22.289*** (df = 4; 73)	19.186*** (df = 5; 72)	15.769*** (df = 6; 71)
Note:			*p<0.1; **p<0.05; ***p<0.01

Figure 3. Instead of the tone index *indposmon*, the variable *hawkish* is integrated. It is statistically significant at the 90% level in regression (2).

iii. Graphs

(1) Approach I: dictionary method

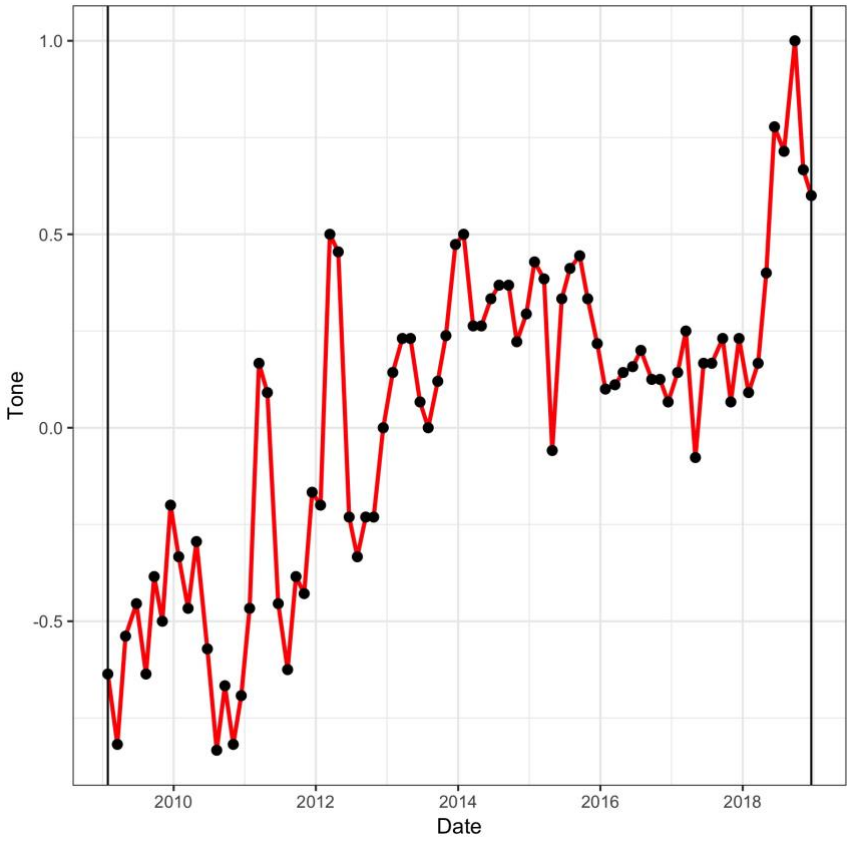

Figure 4. The change in tone from January 2009 to December 2018. The tone on the y-axis is captured through the naïve classifier *indposmon*.

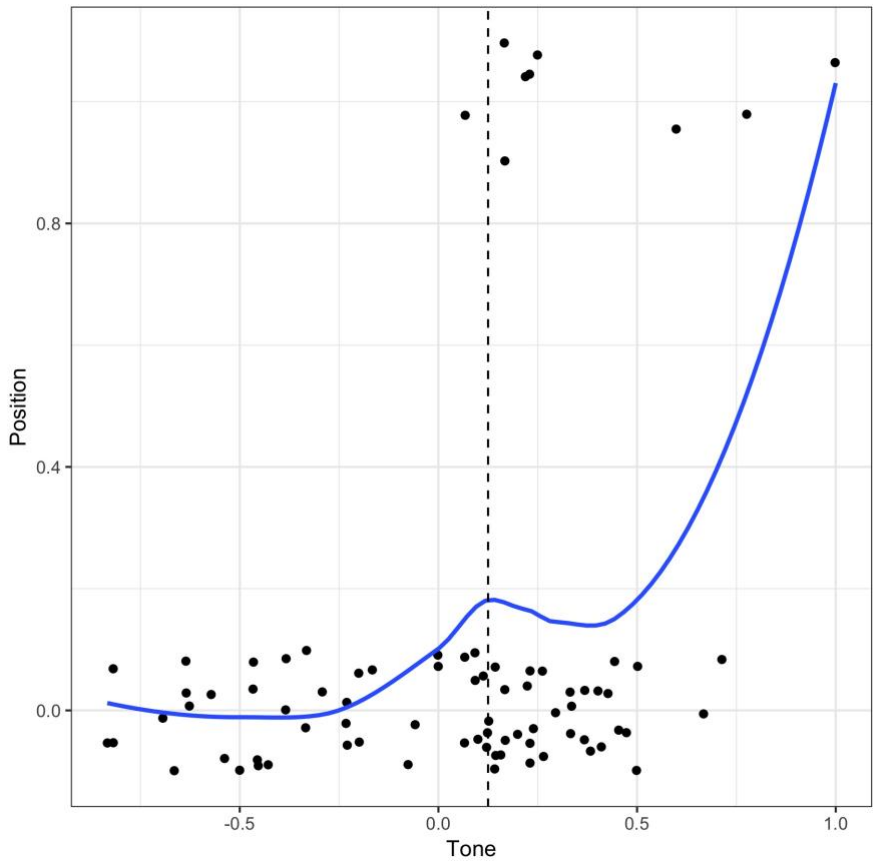

Figure 5. The y-axis is the position of the target federal funds rate indicated in each statement, and the x-axis is the naïve classifier *indposmon* of each statement calculated using the following formula: (H-D)/(H+D). The intercept of the dashed line on the x-axis indicates the median of *indposmon*. The blue line is the best fitted line of movements.

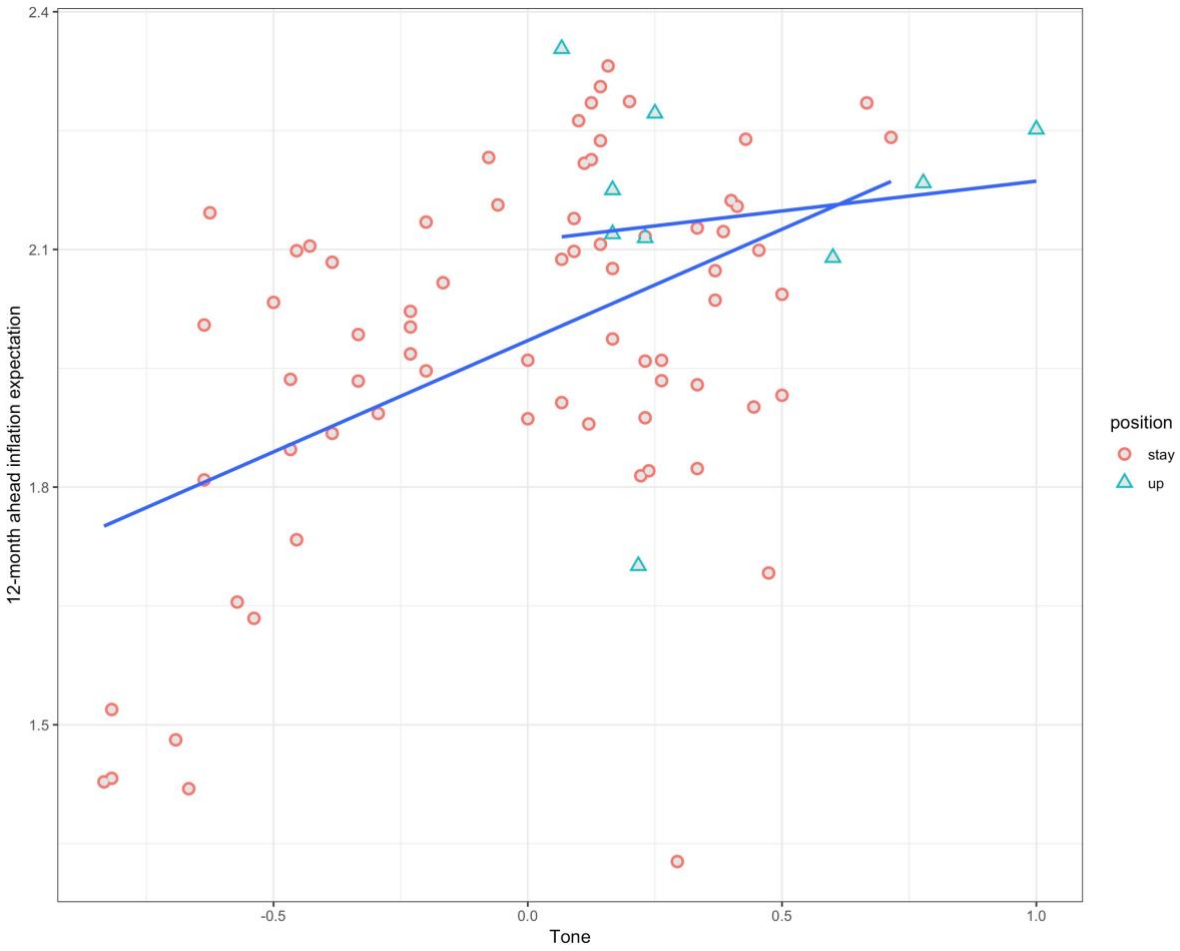

Figure 6. The relationship between the 12-month ahead inflation expectations and the tone of each statement. The sidebar *position* indicates the FOMC committee's decision to increase, decrease, or keep the target federal funds rate as is. The blue lines are best fitted lines of movements.

(2) Approach II: STM

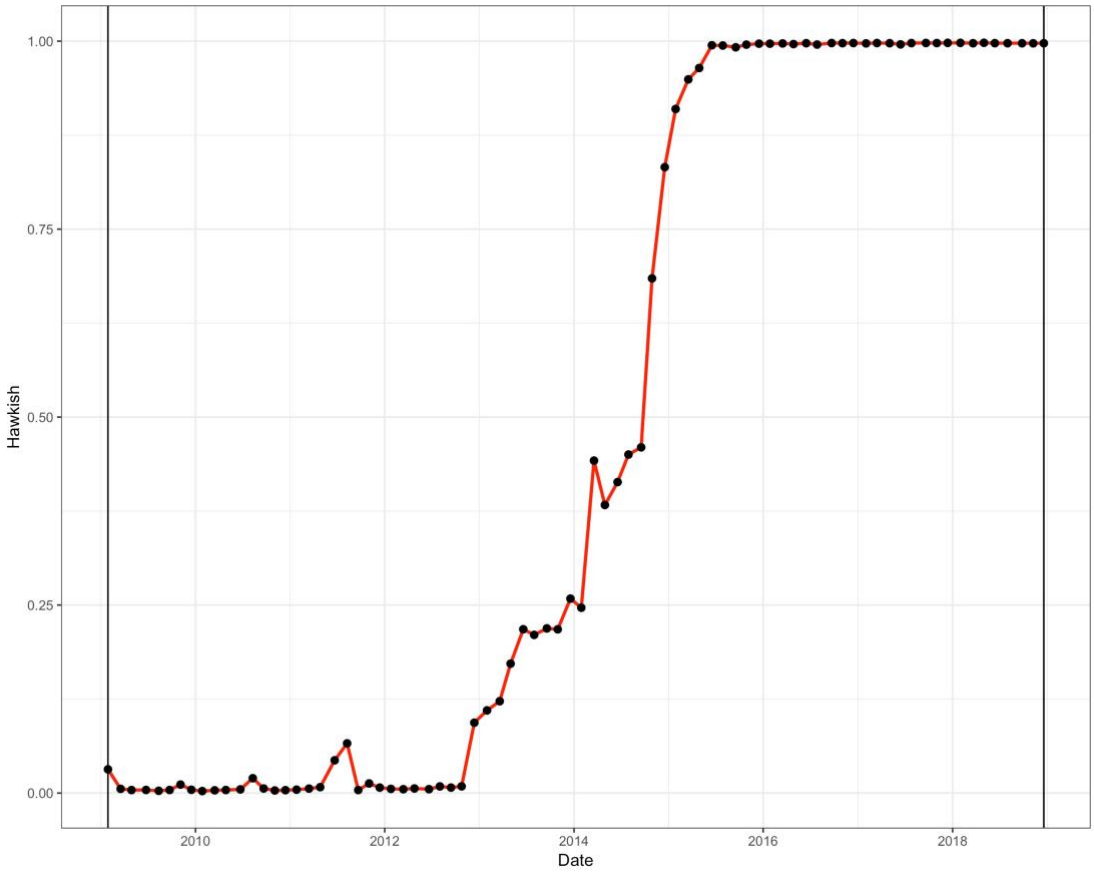

Figure 7. A gradual hawkishness is adopted in the FOMC statements from January 2009 to December 2018. The variable *hawkish* on the y-axis is captured through the theta parameter in the STM.

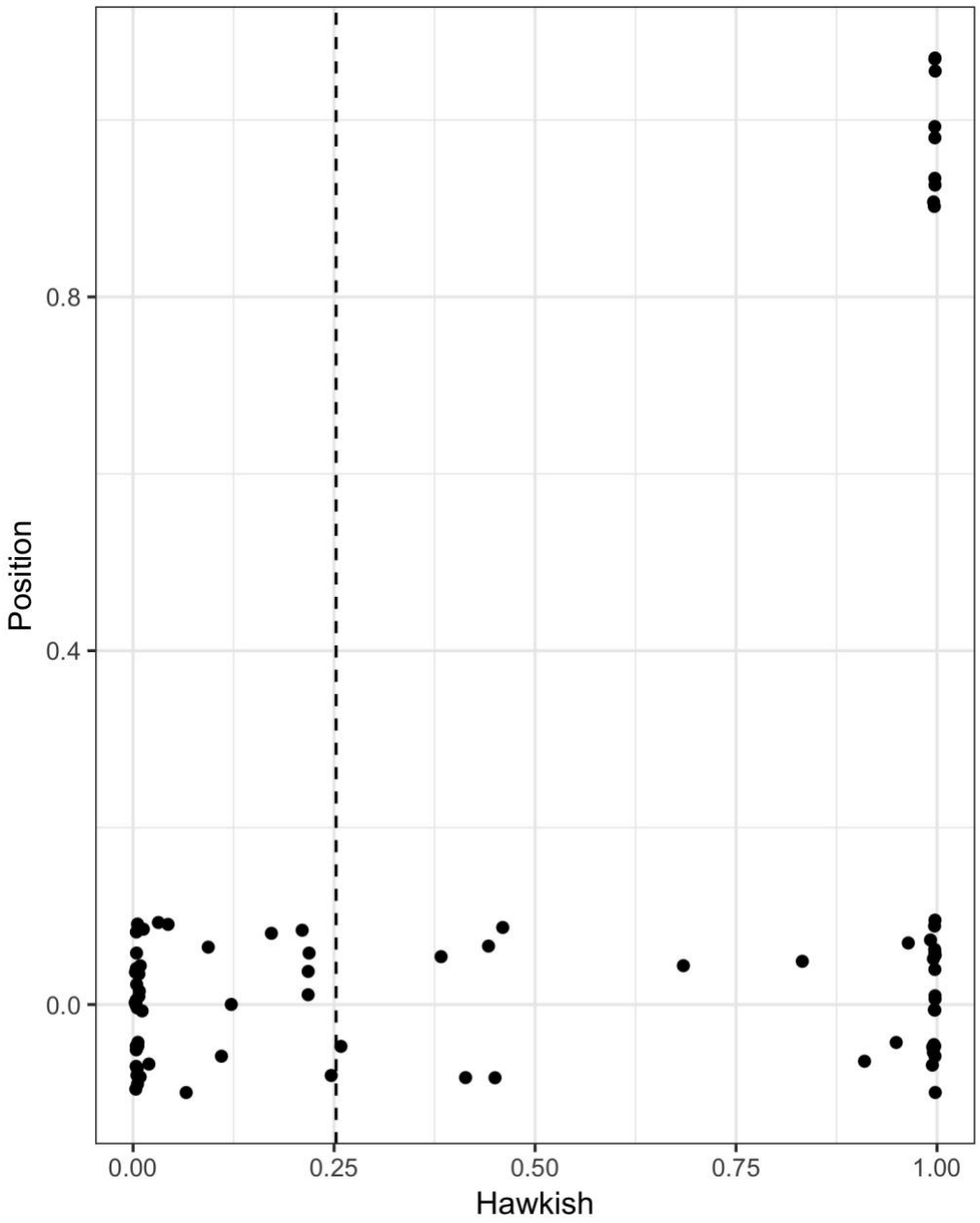

Figure 8. The y-axis is the position of the target federal funds rate indicated in each statement, and the x-axis is the variable *hawkish* captured through the theta parameter in the STM. The value 0 means that the overall tone of the statement is dovish, whereas the value 1 indicates that the overall tone is hawkish. The intercept of the dashed line on the x-axis designates the median of *hawkish*.

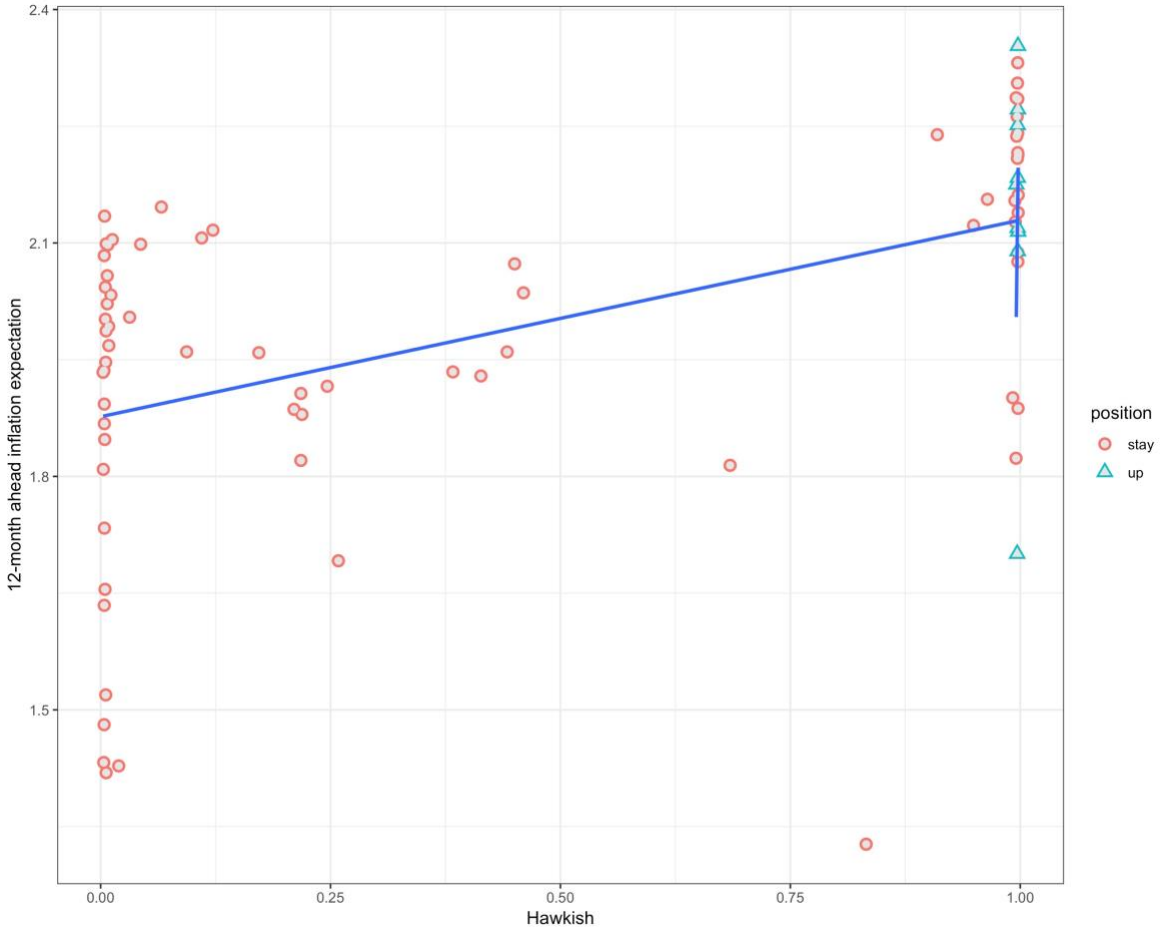

Figure 9. The relationship between the 12-month ahead inflation expectations and the hawkishness of each statement. The sidebar *position* indicates the FOMC committee's decision to increase, decrease, or keep the target federal funds rate as is. The blue lines are best fitted lines of movements.

VI. Discussion

The two main objectives of this study are: (1) whether there is a direct relationship between the tone used in the FOMC statements and inflation expectations, and (2) whether both approaches (i.e. dictionary method and STM) generate similar regression outputs and graphs that lead to the same conclusion.

The dependent variable for all regressions is the mean of the 12-month ahead inflation expectations collected from the Consensus Economics survey. In Figure 1, the independent variables are mostly the lagged terms of the macroeconomic variables, with an exception of the total word count of each statement; tone-indicating variables are not incorporated. Clearly, the lagged term of the inflation expectations (i.e. *ec.nexty.l1*) is significant at 99% level in all three regressions. Keeping all other variables constant, the dependent variable would increase 0.67 as the lagged term increases one unit.

Figure 2 and 3 present the regression outputs using the dictionary method and the STM respectively. The variable *ec.nexty.l1* is statistically significant at 99% level in both approaches. In Figure 2, I replace the variable indicating the total word count with the tone index *indposmon*. The variable *indposmon* is statistically significant at the 95% level in regressions (2) and (3). Specifically, the next year's inflation expectations would increase 0.126 and 0.156 respectively, as *indposmon* increases one unit. It demonstrates that the tone index *indposmon* positively affects the 12-month ahead inflation expectations. In Figure 3, the variable *indposmon* generated using the dictionary method is replaced by the variable *hawkish* that indicates the percentage of each statement associated with a hawkish topic. This value is captured through the theta parameter generated using the STM. It is worth mentioning that the variable *hawkish* is significant at the 90% level in regression (2), and it has a positive impact on the 12-month ahead inflation expectations (0.102). Therefore, it can be concluded that there is a positive relationship between the inflation expectations and the tone used in the FOMC statements, which is captured by either the naïve classifier *indposmon* in the dictionary method or the variable *hawkish* using the STM.

Figure 4 and 7 illustrate the change in tone adopted by the FOMC committee over the past ten years, using the dictionary method and the STM respectively. An overall increasing

trend of hawkishness with cyclical variations is discovered in Figure 4, while the trend appears to be smoother in Figure 7. This difference may be attributed to the fact that the top words captured by the STM are not fully aligned with those in the self-constructed lexicon used in the dictionary method. However, both figures display a similar hawkish trend in the FOMC statements. To further support this claim, there are more data points scattered on the right of the dashed line that signifies the median of the tone index as shown in both Figure 5 and 8, moving toward the hawkish side. It is also worth noting that the FOMC committee's decision on the target federal funds rate, which is captured by the y-axis variable *position* in both figures, is dependent on the tone used in the statement. Specifically, when the tone communicated in the statement becomes more hawkish, the target federal funds rate increases as well. This is because the Federal Reserve tries to keep the economy growing at a more moderate pace, i.e. guard against the excessive inflation, by raising the target federal funds rate.

In addition to the regression outputs, Figure 6 (dictionary method) and 9 (STM) serve as graphical evidence to further support the conclusion that there is a positive relationship between the tone index and the inflation expectations. In both figures, the sidebar *position* has two categories. The "stay" category suggests that the FOMC committee decides to keep the target federal funds rate as indicated in the previous meeting, whereas the "up" category means that the committee decides to increase the target rate. In Figure 6, the best fitted line portrays a clear, positive relationship for both categories. In Figure 9, a similar, positive relationship is observed for the "stay" category; yet, the best fitted line for the "up" category is a vertical line at the value of 1.00. This difference may, again, be due to the disparities in categorizing tone-indicating words and phrases in two approaches.

VII. Conclusion

In this study, I have discovered a positive relationship between the mean of the 12-month ahead inflation expectations and the tone adopted in the FOMC statements captured using two different approaches. Although the first approach may be argued for its subjectivity, I follow closely the different examples of expansion and contraction words used in various literature when constructing my phrase-based lexicon. Therefore, my lexicon should not deviate too much from the mainstream. The second approach, the STM, serves as a robust check for the results derived from the first approach. It is also worth mentioning that there is an overall increasing trend of hawkishness in the FOMC statements throughout the past decade. In addition, a hawkish tone is associated with an increase in the target federal funds rate.

Due to the scope of this study, I only focus on the FOMC statements from January 2009 to December 2018. It will be interesting to also analyze the statements prior to or during the Financial Crisis period. Researchers can explore whether there are any differences in the tone used in statements before and after the major events occurred in the U.S. history. Other suggested future research questions include: Do most central banks around the world gradually adopt a hawkish tone in their press releases? Among those that adopt the hawkish (or dovish) tone, are their intentions or ultimate goals similar? It is also worthwhile to extend the analysis to other forms of the Federal Reserve releases, such as speeches, FOMC minutes, etc.

References

Blei, D. M., A. Y. Ng, and M. I. Jordan (2003), "Latent Dirichlet Allocation," *Journal of Machine Learning Research*, 3, 993-1022.

Federal Reserve Bank (2019), "Federal Open Market Committee," Accessed on 15 January, 2019. Available at: https://www.federalreserve.gov/monetarypolicy/fomccalendars.htm.

Grimaldi, M. B. (2011), "Up for count? Central bank words and financial stress," *Sveriges Riksbank Working Paper Series*, No. 252, Sveriges Riksbank, Stockholm.

Haan, J., S. Eijffinger, and K. Rybinski (2007), "Central bank transparency and central bank communication: Editorial introduction," *European Journal of Political Economy*, 23, 1-8.

Hansen, S., M. McMahon (2016), "Shocking language: Understanding the macroeconomic effects of central bank communication," *Journal of International Economics*, Vol. 99, Supplement 1, 114-133.

Kahveci, E. and A. Odabaş (2016), "Central Banks' Communication Strategy and Content Analysis of Monetary Policy Statements: The Case of Fed, ECB, and CBRT," *Procedia-Social and Behavioral Sciences*, 235, 618-629.

Kazinnik, S. (2017), "Essays on Monetary Policy in Israel," *University of Houston*.

Lucca, D. and F. Trebbi (2009), "Measuring Central Bank Communication: An Automated Approach with Application to FOMC Statements," *NBER Working Papers*, No. 15367.

Mazis, P. and A. Tsekrekos (2017), "Latent semantic analysis of the FOMC statements," *Review of Accounting and Finance*, Vol. 16, Issue 2.

Organization for Economic Co-operation and Development, Consumer Price Index: Total All Items for the United States [CPALTT01USM657N], retrieved from FRED, Federal

Reserve Bank of St. Louis; https://fred.stlouisfed.org/series/CPALTT01USM657N, January 18, 2019.

Roberts, M., B. Stewart, and D. Tingley (forthcoming), "stm: R Package for Structural Topic Models," *Journal of Statistical Software*.

Roberts, M., B. Stewart, D. Tingley, C. Lucas, J. Leder-Luis, S. Gadarian, B. Albertson, and D. Rand (2014), "Structural Topic Models for Open-Ended Survey Responses," *American Journal of Political Science*, Vol. 00, No. 0, 1-19.

Roberts, M., B. Stewart, and E. Airoldi (2016), "A Model of Text for Experimentation in the Social Sciences," *Journal of the American Statistical Association*, Vol. 111, No. 515, 988-1003.

Rohlfs, C., S. Chakraborty, and L. Subramanian (2016), "The Effects of the Content of FOMC Communications on US Treasury Rates," *Proceedings of the 2016 Conference on Empirical Methods in Natural Language Processing*, 2096-2102, Austin, Texas.

U.S. Bureau of Economic Analysis, Real Gross Domestic Product [A191RL1Q225SBEA], retrieved from FRED, Federal Reserve Bank of St. Louis; https://fred.stlouisfed.org/series/A191RL1Q225SBEA, January 18, 2019.