Reading Fed Monetary Policy Communications with an Implication Decision Model Using Zero-Shot Text Classification

Analysis of the Fed's communication by using textual entailment model of Zero-Shot classification

Tomofumi Sawaki[1] Yasuhiro Nakayama [2]
Tomochika Sawaki1, Yasuhiro Nakayama2 Yasuhiro Nakayama2

[1] Mizuho Bank, Ltd.
1Mizuho Bank, Ltd.
[2] Mizuho Research & Technologies
[2] Mizuho Research & Technologies, Ltd.

Abstract: In this study, we propose a method for evaluating the tone of central bank policy by analyzing documents published by central banks using text mining techniques. Since the monetary policies of major central banks have a significant impact on financial market movements, pricing of risky assets, and the real economy in general, market participants are attempting to better understand changes in the central bank's future monetary policy outlook. Since the documents published by the Central Bank are an important tool for the Bank to communicate with the market, they have been elaborated in detail in terms of grammatical syntax and expressions, and investors are expected to read and understand the Bank's policy stance more precisely. Although sentiment analysis of the Bank's documents has been conducted for a long time, it has been difficult to accurately interpret the meaning of the text and to explicitly capture the intentional changes in nuance. In this study, we attempted to continuously evaluate the meaning of the Fed's statements, minutes of FOMC meetings, written transcripts of meetings, and statements by Fed officials, using the same model for an unknown economic environment, by employing an implication judgment method based on zero-shot text classification. The comparison is made for the case where the data sources are the statements and minutes of the FOMC meetings, the written minutes of the meetings, and the statements by Fed officials. In addition, we conducted a cross-sectional analysis of changes in the FOMC meeting minutes by policy stance since 1971.

1 Introduction

Since the monetary policies of the central banks of major countries have a wide-ranging and significant impact on financial market movements, pricing of risky assets, and spillovers to the real economy, market participants are attempting to better understand changes in the future monetary policy outlook of the central banks. In particular, the monetary policy of the U.S. central bank (Federal Reserve System, hereinafter referred to as "Fed") is regarded as the most important one because it determines the behavior of the reserve currency, the dollar. One of the means for the central banks to hold a dialogue with the market and conduct smooth policy management is the publication of various documents, such as statements and minutes released after policy meetings, as well as speeches and testimonies given by senior officials to Congress. The Federal Open Market Committee (FOMC), the meeting at which the monetary policy of the U.S. is set, is a meeting of the Federal Reserve Board (FRB) and the Federal Open Market Committee (FOMC).
(The statement is published on the website immediately after the meeting, and a press conference is held by the Chairman of the meeting.

(Transcript) is released. Then, about three weeks later, the minutes of the FOMC meeting are released (three days until December 2004). The Transcript is a relatively short document of about two pages that summarizes the current economic assessment, the monetary policy decisions, and the names of the voters. The transcript of the meeting consists of the transcript read by the chairman at the beginning of the meeting and the question-and-answer session with reporters, and is about 20 to 30 pages in length. In some cases, the transcripts include information that is not included in the statement but is of interest to market participants (specific details and future prospects). The minutes of the FOMC meeting are around 10 to 20 pages in length and contain a report of the Fed economist's economic analysis followed by the process of the discussions leading to the policy decisions and the variation of opinions among the members. Outside of the FOMC meetings, transcripts of speeches, interviews, and congressional testimonies by FOMC members (Fed officials) are released each time the meetings are held. Although the topics are not necessarily related to the monetary policy, the Fed officials' own economic perceptions and the monetary policy outlook are described, and the statements may not contain any significant changes from the past statements.

In the case of a change in the Fed's policy, it may have an impact on the market. Statements by Fed members are posted on their Web sites, and those by regional Fed presidents are posted on the Web sites of their respective Feds. In most cases, the content is around a few pages. The Beige Book (the regional Fed's business reports) is the other report that attracts the attention of market participants. The Beige Book is a report compiled by the Federal Reserve Bank of each of the 12 districts on their regional economies. The Beige Book is released on Wednesday two weeks before the FOMC meeting and is used as a reference material for the FOMC meeting. The volume of the report is about 30 pages.

The characteristics of each type of document as a form are as follows: voice

While the text, the minutes, and the Beige Book are documents with a standardized structure of contents and chapters, the transcripts of FOMC meetings and statements by officials are not standardized documents. In terms of implications for monetary policy, it is generally considered that the transcripts of officials' statements have precedence over the FOMC meeting in that the officials' perceptions can be obtained in advance of the FOMC meeting (however, there is a blackout period immediately before the FOMC meeting and the officials' statements are not made). It is also considered that more information can be obtained from the minutes or transcripts of the meetings, which contain more text, than from the statements, which contain less content.

Therefore, in this study, the statement, the minutes of the meeting, the transcript of the meeting, and the

Text mining was conducted using each of the statements made by high-ranking officials, and their usefulness as information sources was compared.

2 Related Studies

As an application of text mining techniques to the financial and economic fields, many studies have been conducted on Central Bank of Japan (CBM) documents as well as corporate earnings and news analysis. Most of the studies are active ones, using text mining techniques to score the sentiment of the Central Bank of Japan (CBM) and to predict future monetary policy, markets, and economic indicators.

Ito et al. extracted sentiment by topic for the FOMC minutes by combining an expert-created FOMC-specific dictionary and a dependency parsing, and showed that the sentiment has explanatory power against macroeconomic indicators[1]. Wang calculated sentiment by topic using sentence-level embedding vectors obtained by FinBERT, and showed that they have explanatory power for macroeconomic indicators[2].

Granziera et al. also calculated the sentiment toward inflation for the speech texts of FOMC members and district Fed chairmen[3]. Five years after the FOMC meeting

[1] https://www.federalreserve.gov/default.htm

Some studies have analyzed the sentiment of each member by using Transcript, which is a publicly available record of all the participants' comments[4].

The purpose of this study is to use natural language processing techniques to decipher the intentions of the writers and speakers of the documents as correctly as possible and to be the most correct recipient of the Fed's intended communication. Therefore, we analyzed the characteristics of the Fed's policy tone over a long period of time by dividing it into phases for each trend of the policy rate, by looking back to 1993, when the minutes were first published in the current format, and further back to 1971, when the minutes were published in a similar format.

3 Proposed Method

3.1 Data Acquisition and Preprocessing

We obtained the FOMC statement, the minutes of the meeting, the written transcript of the meeting, and the statements of senior officials, which are available on the FRB's website[1] . The minutes of the meetings are not analyzed in terms of the names of the participants, descriptions of the specifics of the policies, and other paragraphs that are not related to the stance of the FOMC. (However, all paragraphs were included in the analysis for the old minutes for which the paragraph structure could not be obtained.) The transcripts of the meetings are published only in the form of the transcript read by the chairperson on the day of the meeting, and are published the next day in the form of the transcript including the question-and-answer part by the reporters. However, only the responses of the chairperson are included in the transcripts, while the questions and answers of the reporters are excluded from the transcripts.

The retrieved documents are divided into sentence units after pre-processing such as removal of footers and annotations.

3.2 Topic judgment by implication judgment model

Topic judgments are made using a zero-shot based[6] implication judgment model. In this case, we use publicly available trained models. In selecting a model, members with domain knowledge compared the outputs of several models and selected the one that best fits our senses.[2] . The implication judgment model returns a score from 0 to 1 indicating whether or not the target sentence implies the meaning of the hypothetical sentence. A score close to 1 means that the sentence implies the hypothetical sentence, while a score close to 0 means that the sentence contradicts the hypothetical sentence. Since this is a zero-shot model, the hypothetical sentence can be set freely. In this case, the hypothetical sentence is "This sentence is related to the topic of {}. and the topic to be judged is set in parentheses.

The topics of this issue are "Inflation," "Employment

[2] https://huggingface.co/facebook/bart-large-mnli

(The three categories are "Employment", "Economic Activity", and "Economic Activity", and the judgments are made for each sentence. Since a single sentence may contain multiple topics, we judged whether or not the sentence belonged to each topic. Since the model returns an implication score, we set a threshold and judged a sentence as belonging to a topic if the score is higher than the threshold. The threshold is set at 0.9, based on the judgment scores of some sentences extracted as samples and confirmed by members who have domain knowledge.

3.3 Setting up categorical hypothesis statements

In order to determine the connotations of detailed nuances in each category, we create categorical hypothetical sentences. The hypothetical sentences are formulated by the following process.

First, from the past statements released by the Federal Reserve Board, we extracted keywords from the three categories of "Inflation," "Job Gain," and "Economic Growth" were extracted from the past statements released by the Fed on a keyword basis. From the extracted expressions for each category, frequently used expressions were mainly selected and set as hypothetical sentences. When selecting the expressions, we also considered the balance of directions. For example, for "Inflation," the expressions "defined," "diminished," "edged down," and "up" were selected for the downward direction. edged down," and "increased," "moved up," " elevated" in the upward direction.

3.4 Implication judgment for each category

Based on the hypothetical sentences for each category, the implication judgment is performed for each category. The target sentences for implication judgments for each category are as follows, based on the results of the topic judgments in 3.2. For the category "Inflation," we use "inflation (Inflation)" topic, "Job Gain" category is "Employment (Employment)" topic, and the "Economic Growth" category was analyzed for sentences that were determined to belong to the The "Economic Growth" category includes sentences judged to belong to the "Economic Activity" topic. The model for judging the implication of each category is the same as the one used for judging the topic in 3.2. We perform implication judgments on the hypothetical sentences for each category set in 3.3, count the number of sentences for which the implication judgment score is higher than the threshold value (0.9), and aggregate the results for each document. In each category, the number of implied sentences is counted for each direction of expression, and the difference is taken and divided by the total number of implied sentences in both directions to calculate the stance score $s_c(d)$.

□□ (□) is the number of sentences determined to contain the expression □ in document □.

Figure 1: Analysis Flow

4 Results and Discussion

4.1 data

Data on FOMC statements, minutes, transcripts of meetings, and statements by officials were obtained from the FRB website and evaluated. Only the minutes were obtained from February 1971, and the other data were obtained from December 2018. Since policy changes after February 2023, the time when this study was conducted, cannot be verified, the analysis was conducted until March 2022.

4.2 Analysis from December 2018 to March 2022

Using the data of FOMC statements, minutes, transcripts, and officials' remarks from the December 2018 meeting to the March 2022 meeting, we conducted text mining for the post-2021 phase (the phase in which the Fed shifts

$$S_C(d) = \left(\sum_{e \in E_{C,p}} C_e(d) - \sum_{e \in E_{C,n}} C_e(d) \right) \Big/ \sum_{e \in E_{C,p} \cup E_{C,n}} C_e(d)$$

where □□ , □ is the upward directional expression in the category □ set, □□,□ is the set of downward representations in the category □, and

from an accommodative policy of zero interest rate policy and quantitative easing to a tight policy of tapering, raising policy rates, and reducing the B/S amid

rising inflation).), we conducted text mining and compared the data sources. The comparison is limited to the period starting from December 2018 because of the change in the FOMC meeting schedule, which requires the Chairman to hold a press conference after each FOMC meeting.

The reason for this is to match the timing.

Figure 2 plots the stance score on inflation. The following is a hypothetical expression of the upward direction of inflation. Hypothetical expression of the upward direction of inflation ("increased", "picked up", "moved up", "elevated "increased", "picked up", "moved up", "elevated", etc.), the percentage of sentences that were judged to imply inflation

The percentage of sentences judged to contain a hypothetical expression in the downward direction ("declined", "diminished", "moved lower", "edged down", etc.) is subtracted and scored. edged down", etc.) are subtracted from the score.

With regard to the FOMC statement, the score was such that the perception of a downward trend in inflation was indicated until the March 2021 meeting, and from the April 2021 meeting the perception of upward pressure on inflation turned out to be increasing. Thereafter, a similar trend continued.

Next, as for the score when the FOMC minutes are used as the data source, the score moved from the perception of low inflation to near neutral from the September 2020 meeting, and the perception of increasing upward pressure on the inflation rate did not shift until the April 2021 meeting, as was the case with the statement. The same trend has continued since then. It should be noted, however, that the minutes are released three weeks after the meeting.

Third, looking at the score transition when the FOMC meeting transcript is used as the data source, as in the case of the FOMC meeting minutes, the score has gradually returned to the neutral range since the September 2020 meeting and turned to the perception that upward pressure on the inflation rate has strengthened at the March 2021 meeting. This is the most significant change from the statement and the minutes. This was the earliest turnaround since the statement and minutes. At the March 2021 meeting, there was no major change in the inflation perception in the statement, but the result suggests the possibility that the Chairman's press conference after the meeting indicated his recognition of the strengthening of inflationary pressures. The same trend continued thereafter. Given that the minutes of the meeting are released three weeks after the meeting while the written transcript of the meeting is released the next day, this is considered to be the fastest turnaround among the three data sources.

Finally, for the case where the source is a high-ranking official's statement
The average of the five most recent FOMC statements is provided for comparison with the other documents. In order to compare with other documents at the same point in time, the average scores of the five statements by senior officials immediately before the FOMC meeting are shown. As in the minutes, the score gradually turned upward from September 2020, and then sharply increased in April 2021.

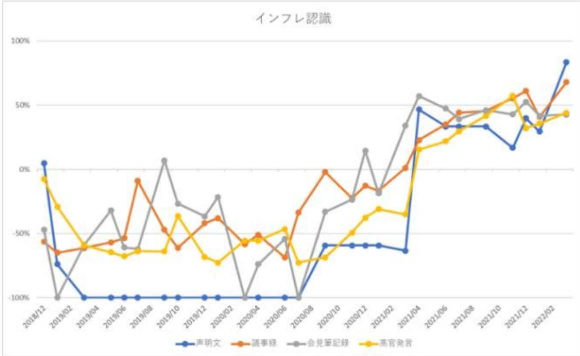

Figure 2: Stance on Inflation

Next, we review the results for the same analysis on employment presented in Figure 3. As for the transition to an accommodative monetary policy following the shutdown in the wake of the global spread of the novel coronavirus after March 2020, all information sources confirm a sharp drop in scores in April 2020, partly due to the emergency interest rate cut, and a shift to a perception of an accommodative employment market. . Thereafter, the written and recorded minutes of the meetings gradually returned to the neutral direction, while the statement shifted to the perception of a tight employment environment with a near-digital change in the April 2021 meeting.

Figure 3: Stance on Employment

Finally, Figure 4 shows the economic perceptions. In the statement and minutes of the February 2019 meeting, economic perceptions turned toward deceleration, confirming the trend that preceded the policy rate cut from July 2019 onward. Around March 2020, as in the case of employment perceptions, there was no significant difference in the timing of the score decline due to differences in data sources, partly because the economic slowdown after the shutdown was abrupt. As for the subsequent changes, the scores based on the minutes of meetings, written transcripts of press conferences, and statements by senior officials returned to the same level as before March 2020 in the summer of 2020, while the score for the statements returned to a neutral level in the spring of 2021.

Figure 4: Stance on the economy

4.3 Analysis using minutes from February 1993

Table 1 uses the minutes of the FOMC meetings from February 1993 to March 2022 as the data source and generates scores for the communication in the same way as in Section 4.2, and compares the scores for each phase of the policy stance. The policy rate (Federal Fund Rate) is obtained from the FRB website and Bloomberg.

First, with respect to the perception of inflation, the Fed's communication also showed a low perception of upward pressure on inflation on average over the entire period, as inflation (annualized core PCE) has remained low and subdued since 1993, ranging between 1% and 3% until 2021. On the other hand, the score was higher than the all-period average during periods when the policy rate was raised, and lower than the all-period average during periods when the policy rate was lowered or zero interest rate was applied. These results are consistent with the idea that the policy rate is changed to achieve the inflation target or to stabilize prices.

Next, employment perceptions are also as shown in Table 1.
The score is higher than the average for all periods when the policy rate is raised. The score is higher than the average of all periods when the policy rate is raised, and lower than the average of all periods when the policy rate is lowered or when the policy rate is zero. These results are consistent with the idea that the policy rate is changed as a means to achieve the maximum employment, and thus the results are considered to be a correct reading of the Fed's communication.

Finally, a similar trend was observed for the score on the perception of the economy. The average score for the whole period is positive, i.e., the perception that the economy is strong, and the score is higher than the average score for the whole period during the interest rate hike phase. On the other hand, the average score for the period of interest rate cuts was negative, suggesting that the time when the Fed perceives the economy to be in a recessionary phase and the time when the policy rate is actually cut generally coincide with each other.

Table 1: Period Averages of Scores Using FOMC

Minutes	(i)	(2)	(iii)	4)
inflation	-0.23	+0.00	-0.40	-0.34
employm ent (long	-0.08	+0.26	-0.45	-0.24

(1) Full term, (2) Interest rate hike period, (3) Interest rate cut period, (4) Zero interest rate period (2009/1-2015/10, 2020/4-2022/1)
In the case where the interest rate is raised every two meetings, the meeting at which the interest rate is kept unchanged is also included in the rate hike period.

Welch's t-test with the alternative hypothesis that (2) the mean value of the interest rate hike phase > (3) the mean value of the interest rate cut phase, all of inflation, employment, and the economy were determined to be significantly different at the 1% significance level.

4.4 Analysis using minutes from February 1971 to December 1992

Prior to December 1992, minutes in the current format did not exist.
Although the minutes of the FOMC meetings from February 1971 to December 1992 are not available, there exists a document similar to the minutes of the meetings called Minutes of Action. Table 2 compares the Minutes of Action of the FOMC meetings from February 1971 to December 1992 as a data source for each phase of the policy stance using the same method as in 4.3. The same trend as in 4.3 is observed, although the level of the scores is different. Comparing the averages of Table 1 and Table 2 over the whole period, the score of inflation perception is higher in Table 2, while the score of employment perception is lower and the score of business cycle perception is consistent with each other. This is consistent with the larger-than-smaller relationship between the mean values of the inflation rate (core PCE) and the unemployment rate in each period. In addition, in Table 2, the average values of all three items, inflation, employment, and the economy, tend to be higher during the rate hike phase than during the rate cut phase, as in Table 1, indicating that the Fed's reaction stance to the inflation and employment data is consistent over a long period of the past 50 years and can be quantified using the model in this study. This suggests that the Fed's reaction stance to inflation and employment data is consistent over a long period of the past 50 years and may be quantified using the model in this study.

Table 2: Period Averages of Scores Using FOMC Minutes

	(i)	(2)	(iii)
inflation	-0.05	+0.17	-0.34
employm ent (long	-0.31	-0.13	-0.63

(1) Full term, (2) Rate hike period, (3) Rate cut period

Welch's t-test with the alternative hypothesis that (2) the mean value of the interest rate hike phase > (3) the mean value of the interest rate cut phase, all of inflation, employment, and the economy were determined to be significantly different at the 1% significance level.

5 Conclusion

In this study, we attempted to read the changes in Fed's communication by utilizing the implication judgment model based on zero-shot text classification. Since zero-shot classification is used, it is possible to deal with other themes without additional learning.

Using data for the period from December 2018 through March 2022, we attempted to make comparisons for each document published by the Fed. For the statements, the scores generated by the model used in this study show a somewhat digital-like movement

On the other hand, a more gradual change in stance was observed in the case of minutes and transcripts of press conferences. The scores for the statements of high-ranking officials remained around the neutral level, partly due to the fact that the average of the past five statements was taken.

Next, we consider the long term (from February 1971 to March 2022). We conduct an aspect analysis using the minutes of the FOMC meetings for the period from April 1, 2008 to March 31, 2009, and our model allows us to interpret the Fed's communication on inflation, employment, and economic perceptions as consistent with actual policy changes.

thanks

This research was inspired by and developed from the report of the Data Science school of the University of Tokyo in the first half of fiscal year 2022. We would like to express our gratitude to all the students who worked diligently on this project and came up with excellent ideas, as well as to the TAs and supervisors who guided us.

matters that require attention

The contents and views expressed in this paper are those of the author alone and do not represent the official views of his/her company.

References

[1] Ryo Ito, Shintaro Suda, Kiyoshi Izumi: Evaluation of FOMC minutes using topic-specific polarity value assignment method, SIG-FIN-017, (2016)

[2] Sarah-Yifei-Wang: Aspect-based Sentiment Analysis in Document - FOMC Meeting Minutes on Economic Projection, arXiv:2108.04080, (2021)

[3] E Granziera, V. H. Larsen, G Meggiorini: Fed Sentiment and Expectations: Evidence from Speeches by the FOMC Members, (2022)

[4] S Cannon, Sentiment of the FOMC: Unscripted, economic Revie, 2015, issue Q IV, 5-31, (2015)

[5] L Xiao: Using Sentiment Analysis to Understand Monetary Policy Uncertainty, (2022)

[6] W Yin, J Hay, D Roth: Benchmarking Zero-shot Text Classification: Datasets, Evaluation and Entailment Approach, In Proceedings of the 2019 Conference on Empirical Methods in Natural Language Processing and the 9th International Joint Conference on Natural Language Processing, EMNLPIJCNLP 2019, Hong Kong, China, November 3-7, 2019, pp. 3914-3923, (2019)