The Japanese Society for Artificial Intelligence Type 2 Study Group
Materials Financial Informatics Study Group
SIG-FIN-030-12

Reading comprehension of Fed monetary policy communication by entailment model using zero-shot text classification

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

Satoshi Sawaki1Yasuhiro Nakayama2 _

Tomochika Sawaki1 , Yasuhiro Nakayama2

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

Abstract: In this study, we propose a method to evaluate the policy tone of central banks by analyzing documents published by central banks using text mining technology. Since the monetary policies of major central banks have a wide range of impacts on trends in financial markets, the pricing of risky assets, and the real economy, market participants should be able to more accurately monitor changes in the future monetary policy outlook of central banks. I am trying to capture the Since the published documents are also an important tool for the central bank to communicate with the market, the grammatical structure and expressions have been elaborated in detail so that investors can more accurately understand the central bank's policies. A reading comprehension of the stance is required. Sentiment analysis of central bank documents has been conducted for some time, but it has been difficult to accurately interpret the meaning of the text and to explicitly capture intentional changes in nuance. By using the entailment judgment method, we intend to continuously evaluate the unknown economic environment using the same model. We compared the changes in tone when each high-ranking official's remarks were used as data sources.In addition, we analyzed the changes in policy stances since 1971 in the minutes of FOMC meetings.

1 Introduction

The monetary policies of major central banks have a wide range of impacts, including trends in financial markets, the pricing of risky assets, and spillovers to the real economy. We are trying to capture more accurately. In particular, the monetary policy of the U.S. central bank (Federal Reserve System, hereinafter Fed) is of the utmost importance because it affects the movement of the key currency, the dollar. As a means of conducting dialogue and conducting smooth policy management, various documents such as statements and minutes released after policy-making meetings, and transcripts of lectures and parliamentary testimonies attended by high-ranking officials are available. The FOMC (Federal Open Market Committee), the meeting at which U.S. monetary policy is formulated, is held eight times a year with members of the Federal Reserve Board (FRB) and regional Federal Reserve Bank presidents as participants. Immediately afterward, the statement was published on the website, and a conference was held by the chairman.

After that, the minutes of the FOMC meeting will be released about three weeks later (three days later until December 2004). , It is a relatively short document of about two pages that summarizes the names of the voters. The transcript of the conference consists of a manuscript read aloud by the chairman at the beginning and a question-and-answer session with reporters. It is about 30 pages long.In some cases, matters of interest to market participants (specific details and future prospects) that are not included in the statement are recorded. Following the report on the economic analysis by the economists, the document is about 10 to 20 pages in length and covers the process of discussion that led to policy decisions and the differences in opinion among members. In addition, transcripts of speeches, interviews, and congressional testimony by FOMC participants (high-ranking Fed officials) are made public each time the meeting is held. He expressed his own perception of the economy and outlook for monetary policy, making a significant change from his past remarks.

Remarks by FRB members are published on their websites, and remarks by regional Fed presidents are published on the websites of each Fed. In many cases, it is around the front or back of the page. Another thing that attracts the attention of market participants is the Beige Book (Regional Federal Reserve Bank Business Conditions Report). It will be released on Wednesday, two weeks before the FOMC meeting, and will be the report to be used as reference material at the FOMC meeting. The characteristics are that statements, minutes, and beige books

are documents in which the structure and chapters of the content are stylized, while transcripts of interviews and remarks by senior officials are documents that are not stylized. From the perspective of policy implications, it is generally believed that the manuscripts of high-ranking officials' remarks are more advanced than the FOMC meetings, as they allow us to grasp the perceptions of high-ranking officials in advance (There is a blackout period during the meeting, and high-level statements are not made). Therefore, in this research, text mining was performed using statements, minutes, transcripts of conferences, and

remarks of high-ranking officials, and their usefulness as an information source was compared.

2 Related research

As an application of text mining technology to the financial and economic fields, there has been a great deal of research on central bank documents, along with the analysis of corporate accounts and news. There are many active methods that use technology to score central bank sentiment and predict future monetary policy, market trends, and economic

indicators. , Combining the FOMC dictionary created by specialists and dependency analysis, extracting sentiment by topic, showed that it has explanatory power for macroeconomic indicators [1]. used sentence-level embedding vectors to calculate topic-specific sentiments and showed that they have explanatory power for macroeconomic indicators [2]. Granziera et al. Inflation sentiment

was calculated for the speech texts of the FOMC [3].

There is also research that analyzes the sentiment of each member by targeting the public transcript of all participants'

utterances [4]. The aim is to read as accurately as possible the intentions of the writers and speakers, and to be the most correct recipients of the Fed's intended communications. In addition, we also traced back to 1971 the past data, which had been published in a similar format, and analyzed the characteristics of the Fed's policy tone over the long term by segmenting them according to policy interest rate trends.

3 Proposed method

3.1 Data acquisition and preprocessing

We have obtained the FOMC statements, minutes, conference transcripts, and remarks by senior officials published on the Fed's website1 . (However, in the case of old minutes for which the paragraph structure could not be acquired, all paragraphs were included in the analysis.) Only part of the article is made public, and the next day it will be published including the question-and-answer section of the reporter. The text was deleted, and only the chairman's response was extracted.

3.2 Topic Judgment by Entailment Judgment Model

Topic determination is performed using a zero-shot-based entailment decision model [6].In this study, we used a trained model that is open to the public. We compared the outputs and selected the one that best matched our senses2 . The closer it is, the more it implies the hypothetical sentence, and the closer it is to 0, the more it contradicts the hypothetical sentence. Topic classification was performed by putting the hypothetical sentence in the form of “This sentence is related to the topic of {}.” employment

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

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

We set three items, “Employment” and “Economic Activity”, and made judgments for each sentence. Since the model returns an entailment score, we set a threshold and decided that if the score is above the threshold, the topic belongs to that topic. A member with domain knowledge confirmed the judgment score of that sentence and set it at 0.9.

3.3 Setting Hypothetical Sentences by Category

In order to judge the entailment of detailed nuances for each category, we create hypothetical sentences for each

category. Hypothetical sentences were set by the following process.)”, “Job Gain”, and “Economic Growth” were extracted based on keywords. In selecting expressions, we also considered the balance of directions. We tried to extract expressions in both directions, such as ``increased'', ``moved up'', and ``elevated'', in a well-balanced manner.

3.4 Entailment test for each category

Based on the set hypothetical sentences for each category, the entailment judgment is performed for each category. Inflation” topic, “Job Gain” category belonged to “Employment” topic, and “Economic Growth” category belonged to “Economic Activity” topic. We used the same model as the topic judgment model in Section 3.2. In each category, we counted the number of implied sentences in each direction of expression, and calculated the difference between the implied sentences in each direction. Calculate the stance score ( ) by dividing by the number.

$$( ) = (\ddot{y} ( ) \quad \ddot{y} \ddot{y} ( ) \quad) \ddot{y} ( ) \ddot{y}$$

category is the set of ascending representations in where , is the set of descending representations in category ,

( ) is the number of sentences judged to entail expression in document .

Figure 1: Analysis Flow

4 Results and discussion

4.1 Data

FOMC statements, minutes, transcripts of press conferences, and data on remarks by senior officials were obtained from the FRB website and evaluated. In addition, since it was not possible to verify the changes in policy after February 2023, the time of this research, the analysis was divided up to March 2022.

4.2 Analysis from December 2018 to March 2022

From the December 2018 meeting to the March 2022 meeting Using data from FOMC statements, meeting minutes, press conference transcripts, and high-level remarks, we estimate the phase after 2021 (with inflation rising, the Fed adopting an accommodative policy of zero interest rates and quantitative easing). We conducted text mining on the transition from policy to tapering, policy interest rate hikes, and tighter policies such as the reduction of the balance sheet, and compared each data source. The reason for limiting the number of FOMC meetings to December

2018 is to match the timing of the chairman's briefing after each FOMC meeting. Figure 2 plots the stance scores on inflation. From the percentage of sentences judged to imply hypothetical expressions (“increased”, “picked up”, “n

Scores are obtained by subtracting the percentage of sentences judged to imply hypothetical expressions in the downward direction (``declined'', ``diminished'', ``moved lower'', ``edged down'', etc.).

Regarding the FOMC statement, it is recognized that the inflation rate will continue to decline until the March 2021 meeting, and that upward pressure on the inflation rate will intensify from the April 2021 meeting. After that, the same trend continued. Next, when using the

FOMC meeting minutes as a data source, the score changed from low inflation perception to near neutral from the September 2020 meeting. It was not until the April 2021 meeting that the score shifted and the perception that upward pressure on the inflation rate was increasing increased, as was the case with the Statement. It should be noted that the minutes of the FOMC meeting are published three weeks after the meeting. As in the

case of the BOJ, the score gradually returned to near neutral from the September 2020 meeting, and at the March 2021 meeting, the BOJ recognized that upward pressure on the inflation rate was intensifying. At the meeting in March 2021, there was no major change in the recognition of inflation in the statement, but at the post-meeting chairman's conference, recognition of increasing inflationary pressure was indicated. A similar trend continued thereafter. It is thought that this was the fastest conversion among the data sources in

Finally, when using high-level remarks as the source, the average value of high-level remarks for the five times immediately before the FOMC meeting is listed for comparison with other documents. It was confirmed that the score gradually turned upward from April 2021, and that it will rise sharply in April 2021.

Figure 2: Changes in Inflation Stances

Next, we confirm the results of a similar analysis on employment shown in Figure 3. The shift to monetary easing policy accompanying the shutdown following the global spread of the novel coronavirus in March 2020 With regard to the BOJ, partly because of the emergency rate cut, all information sources confirmed a sharp drop in the score in April 2020, and the job market shifted to an accommodative perception. While the transcripts and meeting minutes are gradually returning to a neutral direction, the statement at the April 2021 meeting has shifted to a tighter perception of the employment environment with a range of changes close to digital.

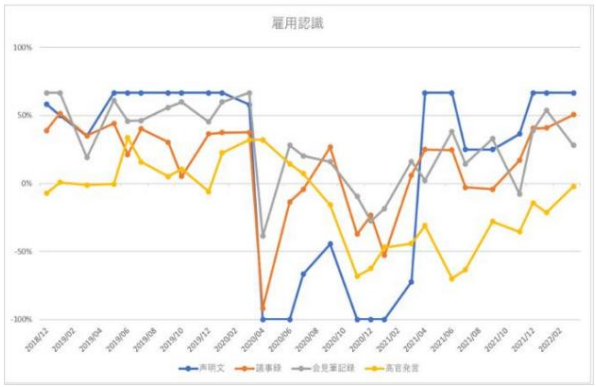

Figure 3: Employment stance trends

Finally, the BOJ's perception of the economy is as shown in Figure 4. In the statement and minutes of the February 2019 meeting, the BOJ stated that the BOJ's perception of the economy had turned to a slowdown, leading to the policy interest rate cut from July 2019 onwards. Around March 2020, as with employment perceptions, there was a sharp economic slowdown after the shutdown, and there was a large difference in the timing of score declines due to differences in data sources. After that, the score based on the minutes of meetings, transcripts of interviews, and remarks by senior officials returned to the same level as before March 2020 around the summer of 2020. , The statement regressed to a neutral score in the spring of 2021.

Figure 4: Changes in economic stances

4.3 Analysis using minutes from February 1993

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

situation where the rate hovered between 3%, the Fed's communication also recognized that upward pressure on inflation was low on average over the entire period. The score was higher than the average for the entire period in the period when the rate was cut, and the score was lower than the average for the period when the interest rate was cut and the period when the interest rate was zero. This result is consistent with the idea that the policy rate will be changed as a means of stabilizing prices.

Next, regarding employment recognition, it is also shown in Table 1. During the period when the policy interest rate was raised, the score was higher than the average for the entire period, and during the period when the interest rate was lowered and the zero interest rate. This result is consistent with the idea that policy interest rates are changed as a means of maximizing employment. It is thought that the Fed's communication was correctly read and understood. Finally, a similar trend was

confirmed for the scores regarding economic perceptions. In the interest rate hike phase, the score was higher than the average over the entire period. It was suggested that the time when the policy rate was actually lowered coincided with the time when the policy rate was actually lowered.

Table 1: Period Averages of Scores Using FOMC Minutes

	ÿ ÿ ÿ ÿ			
Inflation -0.23		+0.00	-0.40	-0.34
Employment -0.08		+0.26	-0.45	-0.24
Economy +0.22		+0.44	-0.16	+0.20

* (1) All period, (2) rate hike period, (3) rate cut period, (4) zero interest rate period (2009/1ý2015/10, 2020/4ý2022/1)

*In the phase where the interest rate is raised every two meetings, the rate hike period also includes the meeting where the interest rate is kept unchanged.

*In Welch's t-test with the alternative hypothesis of ``(2) the average value in the rate hike phase > (3) the average value in the rate cut phase", it was determined that there was a significant difference at the 1% significance level for inflation, employment, and the economy.

4.4 Minutes from February 1971 to December 1992 analysis using

Before December 1992, there were no minutes in the current format, but there was a document similar to minutes called Minutes of Action. of Action as the data source, we compared each phase of policy stance in the same way as in 4.3. Comparing the averages over the entire period, Table 2 shows a higher score for perception of inflation, a lower score for perception of employment, and a match for perception of the economy. (Core PCE) and the average unemployment rate. This is the same as Table 1, which suggests that over the long-term period of the past 50 years, the Fed's reaction stance to inflation and employment data has been consistent and likely to be quantified using the model in this study. gender was suggested.

Table 2: Period Averages of Scores Using FOMC Minutes

	ÿ	ÿ	ÿ
Inflation - 0.05 Employment		+0.17	-0.34
	-0.31	-0.13	-0.63
	+0.22	+0.30	+0.08

Economy * (1) All period, (2) Rate hike period, (3) Rate cut period

*In Welch's t-test with the alternative hypothesis of ``(2) the mean value of the rate hike phase > (3) the mean value of the rate cut phase", it was determined that there was a significant difference at the significance level of 1% for inflation, employment, and the economy.

5 Summary

In this study, we used an entailment decision model based on zero-shot text classification to try to understand changes in Fed communication. Using data from December 2018 to March 2022, we attempted to compare each document published by the Fed. In the score, a

somewhat digital-like movement

While the score changed in 2010, a more gradual change in stance was confirmed for minutes and transcripts of interviews. Yes, and it is moving around a neutral level.

Next, we conducted a phase analysis using FOMC minutes over the long term from February 1971 to March 2022, and found that the Fed's communications on inflation, employment, and economic perceptions were consistent with actual policy changes. This result can be interpreted using the model in this study.

Acknowledgments

This research was inspired by the 2022 first half report of the Data Science school at the University of Tokyo. I would also like to thank all the TAs and supervisors who have guided me.

Considerations

The content and views of this paper belong to the individual author It is not the official position of the company.

References

[1] Ryo Ito, Shintaro Suda, Kiyoshi Izumi: Evaluation of FOMC Minutes by Topic-specific Polarity 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, VH Larsen, G Meggiorini: Fed Sentiment and Expectations: Evidence from Speeches by FOMC Members, (2022)

[4] S Cannon, Sentiment of the FOMC: Unscripted, economics 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 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)