Issues in Information Systems
Volume 25, Issue 3 pp. 252-266, 2024

DOI: https://doi.org/10.48009/3_iis_2024_120

Examining text consistency amongst FOMC statements and minutes subtext using deep learning

Adam Patterson, *Nichols College, adam.patterson@nichols.edu*

Abstract

This study explores a novel approach to evaluating the homogeneity of literary relationships between pairwise Federal Open Market Committee (FOMC) meeting minutes and statements using Natural Language Processing techniques (NLP). This is the first research to summarize minutes for analyzing consistency of committee communication. Bidirectional Encoder Representations from Transformers (BERT), a pre-trained deep learning model, is employed due to its computational efficiency, trailblazing performance, and alignment with text under analysis. First, BERT is utilized to create an abstractive summary of each meeting minute using the first and last 512 tokens. Second, BERT is deployed to generate vector representations for calculation of text similarity metrics between the minutes summary and corresponding statement published by the committee. Text similarity results from 46 meetings during the Powell chairmanship, 2019 through 2023, indicate two statistical outliers proxying a potential lack of transparency. Findings exhibit the March 2019 outlier may be attributed to (i) surprising economic data and (ii) the first meeting under Jerome Powell. The Russian invasion of Ukraine is a potential driver of diminished transparency for the March 2022 meeting.

Keywords: natural language processing, deep learning, monetary policy, federal open market committee, transparency

Introduction

The FOMC, or the committee, is the policy-making body within the Federal Reserve, a quasi-independent institution tasked with enacting monetary policy in the United States of America. The primary purpose of the FOMC is to promote the dual mandate of price stability and maximum employment mandated by Congress (Thornton, 2012). The FOMC primarily controls monetary policy by influencing the federal funds rate. Open market operations are a crucial tool the committee uses to adjust interest rates throughout the economy. Typical actions in pursuit of open market operations affect the money supply by acquiring and disposing of short-term government securities on banks' balance sheets, which influence the federal funds rate, a key interest rate governing overnight lending and borrowing between banks (Kopchak, 2011). Other levers of monetary policy influence include changing the discount rate and banks' reserve requirement. The FOMC establishes a specified range within which it aims to maintain the federal funds rate and deploys open market operations to achieve this target. According to Taylor (2001), the Trading Desk of the Federal Reserve Bank of New York may not always need to buy or sell securities as the market arbitrages the announced target range, causing the rate to adjust accordingly. Thus, communication between FOMC and the public is of utmost importance.

The committee conducts eight regularly scheduled meetings throughout the year, with statements issued on the date of policy decision and minutes disclosed three weeks after, to determine the appropriate monetary policy given the global macroeconomic environment and domestic business cycle—other than the

committee Chair, membership changes at the first meeting of each year. The statements provide a concise summary of the meeting, while the minutes detail discussion and decision-making at the meeting. According to the Federal Reserve (n.d.), the minutes "*describe the views expressed by policymakers and explain the reasons for the Committee's decisions. The minutes can help the public interpret economic and financial developments and understand the Committee's decisions.*". The committee's goal is to communicate current policy decisions and provide foreword guidance on the outlook of monetary policy. Although they are careful in the rhetoric that they communicate, the committee has a strong ability to move markets with publicly provided documents. Research notes that markets react differently to updates on the target range than policy communication itself (Farka & Fleissig, 2012; Gürkaynak et al., 2004; Lucca & Trebbi, 2009) while Handlan (2020) discovers that FOMC wording accounts for four times the variation in federal funds futures than the actual policy announcement. Thus, it is essential to determine if the communicated statements are consistent with underlying discussions in the minutes.

Since they are released many weeks apart, exploring potential heterogeneity in language between minutes and statements allows for consistent evaluation of clarity from the committee. Investigating text differences between pairwise documents allows insights into FOMC transparency, as according to Hansen et al. (2018), a vital feature of the central bank is transparency. Transparency is vital as committee members contain privileged information about salient economic issues and future monetary policy decisions. Potential information asymmetry could be a source of corruption, conspiracy, or misinformation from national leaders. Analyzing text similarity between these documents helps to ensure consistency, transparency, and precision in the Fed's messaging. The analysis performed in this paper attempts to confirm the unbiased nature of the FOMC and, subsequently, the quasi-independent Federal Reserve, given their privileged position and foreword knowledge.

Text similarity analysis may also assist in tracking potential deviations in the communication strategy over the Jerome Powell reign. Shifts in the underlying mechanism of minutes creation can be evaluated by comparing the summarized minutes of FOMC meetings with subsequent policy statements. This information is of value to the public as even subtle discrepancies in communication can influence investment decisions, interest rates, and market stability. Investors and policymakers would be interested to know if the FOMC has changed its minutes function during global exogenous shocks, such as the Coronavirus pandemic 2020.

The FOMC minutes provide a detailed account of the discussions and debates during each meeting, offering insight into the thought processes behind monetary policy decisions. When summarized and compared with the statements, this analysis may assist in bridging the gap between the FOMC's internal deliberations and the external messaging provided to the public in a more digestible form. Text alignment, measured with text similarity, fosters transparency, ensuring that the minutes and statements convey consistent messages. Homogeneity of messaging across both documents reduces the likelihood of market confusion and misinterpretation and will thus assist in the formation and development of efficiently allocated markets. Furthermore, text similarity analysis contributes to a better understanding of the Federal Reserve's credibility and accountability. FOMC's credibility may be bolstered if its communication strategy is clear and consistent throughout different economic cycles.

The scope of this report includes documents published with the Honorable Jerome Powell as Chairman of the committee. Limiting the range of analysis to documents only communicated by the Powell leadership attempts to control communication techniques amongst different regimes. The approximate eight yearly meetings from 2018, when he took over the Chairmanship, until 2023 yields 45 pairwise documents for analysis.

Additionally, the FOMC met for unscheduled sessions during the coronavirus pandemic, but only statements were released from some of these meetings. No minutes were released for the unscheduled sessions. For this reason, all unscheduled sessions were omitted from this analysis, as no pairwise observations exist to examine text similarity. The researcher acknowledges that this is a limitation of this study as salient information communicated around the time of extreme economic uncertainty must be omitted. The committee denied the opportunity for transparency to be evaluated during this time of interest. The researcher assumes the committee omitted the minutes to preserve a free exchange of views about economic conditions and alternative policy approaches. FOMC transcripts are the most detailed record of the meetings and are released five years after the policy date to protect this free exchange.

The contribution of this paper to current literature is to provide institutions with a more immediate, although potentially less accurate, framework to evaluate the transparency of the committee. According to research by Poole and Rasche (2003), which conducts a historical analysis of FOMC transparency, clarity is of interest to institutions as insights into the consistency of committee language can be valuable for economists, investors, and policymakers alike. Transparency can help lead to efficient markets (Acosta, 202) and provide more institutions with symmetric information to compete and correctly arbitrage the policy announcement through an information technology framework (Ramaswamy, 2022).

Literature Review

Extensive research has been conducted regarding the use of NLP techniques in FOMC documents. A large amount of this literature revolves around performing sentiment analysis on committee text (Gardner et al., 2022; Huang & Kuan, 2021; Shapiro & Wilson, 2021; Tadle, 2022) or examining the movements of financial assets from committee communications (Doh et al., 2021; Gorodnichenko et al., 2023; Gürkaynak et al., 2004; Nakamura & Steinsson, 2018; Swanson, 202l).

A small subset of the literature explores the clarity of monetary policy communication through text similarity analysis of committee compositions. Rha (2018) studies the transparency of the committee by proxy of similarity metrics between FOMC statements and mass media news articles reporting on the statements. Acosta (2023), comparable to this study, investigates the transparency of pairwise minutes and transcripts. The study indicates that transparency plays a crucial role in efficient monetary policy, as a more significant effect on interest rates is found when transparency is higher.

Fischer et al. (2023) is also analogous to this research in that pairwise committee documents are analyzed for insights into transparency. Like Acosta (2023), their research uses minutes and transcripts to evaluate policy transparency over the last five years. Results found that publicly providing the transcripts in real time could substantially enhance forecasting accuracy, suggesting potential room for increased transparency from the committee. Although, according to Apel et al. (2022), transcripts are more informative than minutes, the FOMC minutes appear 257 weeks earlier and offer a more immediate transparency metric.

Methodology

Literature from the FOMC was web scrapped directly from the Federal Reserve website (Federal Reserve, n.d.) using the python libraries *urllib3*, *BeautifulSoup* imported from *bs4*, and *requests*. Minutes and Statements from March 2018 (Powell's first meeting as chairman) through November 2023 were downloaded by their HTML and cleaned for analysis. Regular expression (regex) patterns were used to

match homogenous opening and closing headers for downloading a subset of desired text from the body of literature.

In the full-text minutes, the regular expression searches a pattern match in the text that starts with "Minutes of the Federal Open Market Committee" and ends with "Back to Top". All FOMC minutes start with this header title and end with the closure. The starting point for the FOMC statements is "Federal Reserve issues FOMC statement," and the ending is "Last Update:". These starting and ending points are the only differences in data preparation between the FOMC minutes and statements. The text between pattern matches is stored into a variable for analysis after conversion to lowercase form.

The resulting body of text was harvested as one long string and thus needed to be tokenized for imputation into pre-trained models. Tokenization is required for linguistic analysis as this is the step in which text is broken down into individual words, or tokens, for model input. This procedure is essential for algorithms to analyze and predict the sequence of words. The text similarity model utilized by Google (detailed in the next section) requires tokenizing text documents before input into the model. After tokenization, the first 512 tokens and the last 512 tokens from each document are saved separately for input into an abstractive summary, as the pre-trained BERT model has a maximum input length of 512 tokens. This step may be redundant as input must be passed through a BERT-specific tokenizer for input into the BERT model; however, tokenization must be performed to subset the text into 512 token lengths.

Deep Learning Model Used

A pre-trained deep learning transformer was used as the model of choice for summarizing FOMC minutes text and calculating text similarity metrics. Bidirectional Encoder Representations from Transformers (BERT), developed by artificial intelligence engineers at Google, was utilized to gather subtext information embedded in the minutes (Hugging Face, 2020). BERT is a deep learning model that performs well in various NLP tasks, including abstractive summaries, extractive summaries, sentiment analysis, and text similarity evaluation. BERT achieves state-of-the-art performance by analyzing semantic relationships between words (Liu & Lapata, 2019). Extractive summaries highlight essential information within the text, whereas abstractive summarization comprehends the entire text and generates a summary (Bhatia & Jaiswal, 2015). The abstractive approach gathers insights and relationships from the original minutes, even if the information was not explicitly stated.

The secret lies within the BERT name. Bidirectionality allows for mapping literary context both before and after the word through transformer architecture called encoders, similar to the human brain, hence the deep learning title. Important phrases or series of words are examined for underlying trends and patterns to determine the meaning of all literature used throughout the text. BERT looks at all words before evaluating the underlying meaning associated with the surrounding parameters of the given sample.

One significant benefit of using BERT is that the model is pre-trained. A pre-trained model has already been trained on a large corpus to understand lingual patterns, relationships, and trends. This allows the researcher to save time and effort as a model does not need to be initially trained. It is also computationally efficient, given a smaller resource constraint. BERT is pre-trained on two training corpora: (a) the BookCorpus dataset, containing over 800 million tokens from books, and (b) the English Wikipedia corpus, containing over 2.5 billion tokens of text (Mao et al., 2021).

During pre-training, BERT is given two separate prediction objectives to calculate efficient token parameters detailing the relationship within the corpus. The first objective, showcased in Figure 1, is to predict the missing word in a sentence, also called masked language modeling. BERT is given a sentence

with a token, or tokens, masked out. BERT will then predict the missing text and evaluate accuracy metrics, determining if the model's understanding of context is trending positively or negatively.

Figure 1: Masked Language Modeling

In the sentence above, BERT predicts a word to complete the sentence. Then, evaluation metrics are determined through embeddings to arbitrage the most probable word by predicting trial and error. BERT may predict the word guess in the above-masked language example. It may also predict the words describe, write, predict, spell, etc. There are many possible values for a bland singular sentence.

An additional technique is needed to supplement masked language modeling in achieving high-performance results. The second goal of BERT is to predict sentence ordering given a sequence of sentences, known as next-sentence prediction. This task assists BERT in gaining knowledge of the relationships between sentences. BERT is given two masked sentences and tasked with predicting if the sentences are next to each other in the corpus (Devlin et al., 2018).

Sentence prediction allows BERT to learn different ways in which sentences can be associated with each other. Masked language and sentence forecasting techniques allow the final pre-trained BERT model to understand the context of tokens and the relationships between sentences in summarizing FOMC minutes. Thus, BERT learns patterns in the corpus by combining masked language modeling and next-sentence prediction to provide a deep understanding of linguistics, hence the deep learning. BERT utilizes this literary relational knowledge by identifying important information while filtering out potential disinformation. As BERT develops a deep understanding of language through predictions associated with groups, or clusters, of words, the model can coherently concatenate sequences of tokens to deliver a summarized text of salient information.

BERT is also employed to generate vector representations of text that can be used downstream for text similarity metrics. The model utilizes its pre-training to represent subtle differences within the literature. Each text input into the model receives a vector representation, encoding semantic meaning into a high-dimensional space. Traditional mathematic techniques, such as cosine similarity, use the vector representations of the two texts to determine how much the literature resembles each other.

Implementation of the BERT Model

The pre-trained model is ready for sequence-to-sequence tasks once the summarizer model and tokenizer are downloaded, T5ForConditionalGeneration and T5Tokenizer, from the *transformers* package (Hugging Face, 2020). Variables were created for both the tokenizer and summarizer models. T5, or Text-to-Text Transfer Transformer, is a text-to-text learning model trained for various tasks. These tasks include state-of-the-art performance metrics in text summarization, machine translation, question answering, and natural language inference (Koroteev, 2021). BERT offers three different T5 model sizes (small, base, and large: offering 60M, 220M, and 770M parameters, respectively). The larger models are more accurate but slower and require more computational resources. This investigation utilizes the "t5-small" model for all analysis to reserve resources as the "t5-large" model was run on the researcher's computer, but the loop was not complete within a week and thus was aborted.

The next step is to input the tokenizer variables into a pipeline to create a text summarizer. A pipeline is a sequence of tasks performed in a specific order to achieve a desired result. A pipeline gives the machine learning model a series of steps to deploy. The summarizer pipeline created and required for running the BERT algorithm will perform the steps in order, as exhibited below in Figure 2.

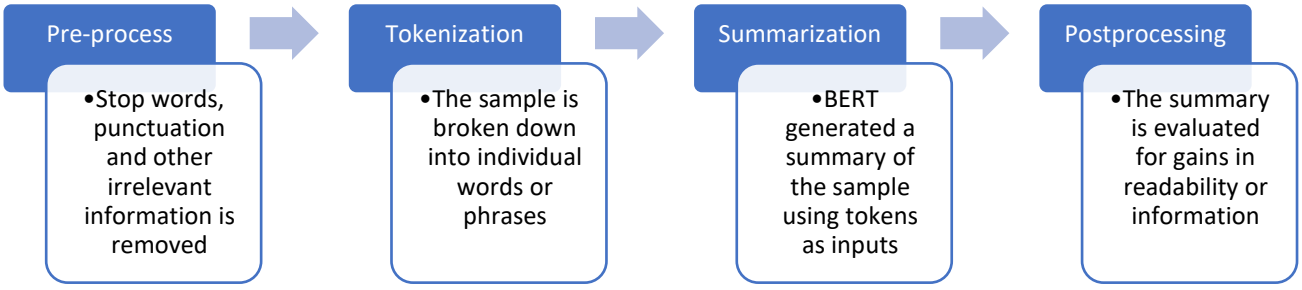

Figure 2: Summarizer Pipeline Steps

The data, the first and last 512 tokens of the minutes, is passed through the pipeline using the argument do_sample=FALSE to avoid using sampling techniques in summary creation. Two additional arguments, both a minimum and maximum length, are input to set the bounds of the abstractive summary. A 50-token minimum and 600-token maximum are used to create the abstractive range. Each created summary output is stored in a dictionary for analysis. The dictionary key corresponds to a number representing the sequence of FOMC minutes (With the first minutes equaling 0, the second minutes equaling 1, etc.), while the dictionary values are the text summaries. Two index values are randomly selected to provide an example of BERT summarized output from the first and last 512 tokens. Ten and thirty-seven were selected, indicating the June 2019 and November 22 meetings. Summarized minute endings, using the first and last 512 tokens, are showcased below, in Figure 3 and Figure 4, in original punctuation from BERT.

June 2019 BERT first 512-token summary	November 2022 BERT first 512-token summary
'the deputy manager of the system open market account provided an overview of developments in U.S. and global financial markets . market participants pointed to a number of factors as contributing to the heightened volatility and sustained declines in risk asset prices and interest rates over recent months . some market reports suggest investors perceived the FOMC as insufficiently flexible in its approach to adjusting the path for the federal funds rate or the process for balance sheet normalization .'	'the manager pro tem turned first to a discussion of financial market developments in the united states . over the period, financial conditions had tightened amid elevated volatility across financial markets . the market-implied path of the policy rate rose, with median federal funds rate values being viewed as indicating a commitment to sustaining a restrictive monetary policy stance . however, contacts were increasingly focused on when the Committee might slow the pace of future increases .'

Figure 3: BERT first 512-token summarized minute examples

June 2019 BERT last 512-token summary	November 2022 BERT last 512-token summary
'inflation and inflation for items other than food and energy remain near 2 percent . in support of these goals, the committee decided to maintain the target range for the federal funds rate . the committee will assess realized and expected economic conditions relative to its maximum employment objective and its symmetric 2 percent inflation objective .'	'the committee seeks to achieve maximum employment and inflation at the rate of 2 percent over the longer run . the committee anticipates that ongoing increases in the target range will be appropriate in order to attain a stance of monetary policy that is sufficiently restrictive to return inflation to 2 percent.'

Figure 4: BERT last 512-token summarized minute examples

BERT is also deployed for the second step of the analysis, obtaining text similarity metrics for pairwise values of summarized minutes and statements. The pre-trained model for text similarity is "bert-base-uncased", the base BERT model with separate handling of upper and lowercase letters. The sample has already been converted to lowercase in data preprocessing using regular expressions. The uncased option is the most accurate and efficient model for the state of the sample (Liu & Lapata, 2019).

BERTTokenizer is employed, instead of the previously used T5Tokenizer, as it performs tokenization for all BERT models while the T5 pre-processes for T5 models. The main difference between the two tokenizers is that BERT obtains word-piece tokens and can split words into sub-words, whereas T5 develops byte-pair encoding tokens by splitting words into pairs of bytes. The former is better suited for outliers, unrecognizable tokens, or general performance, while the latter is more proficient with global tokens (Bostrom & Durrett, 2020). The BERTTokenizer performs three ordered steps depicted in Figure 5.

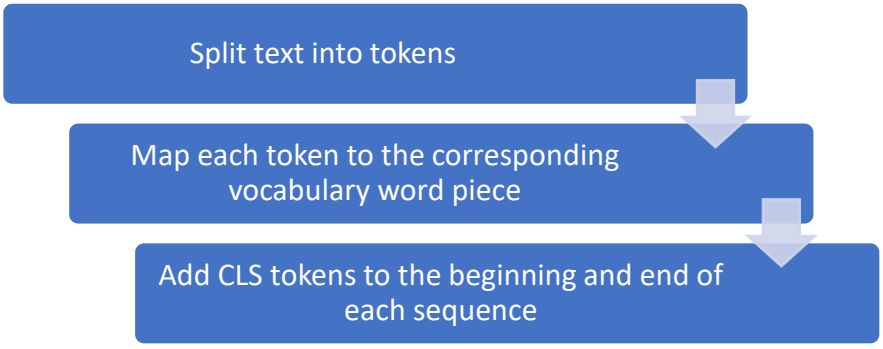

Figure 5: BERTTokenizer Steps

If the token is not found in the vocabulary word piece, it is split into subwords using a greedy algorithm. A greedy parameter matches the longest subword of a word to the vocabulary dictionary. The tokenizer then repeats the middle step for each subword until the word has been completed. The CLS tokens act similarly to HTML tags, identifying the start and beginning of a specific input.

The tokens are converted to tensors after the minutes and statements are tokenized. Tensors are multidimensional arrays used in machine-learning applications. While matrices are restricted to two dimensions, tensors can store data in more than two dimensions. An example is an image with height, width, and depth stored as information in a 3D tensor. Each token is mapped to its corresponding BERT vocabulary index and stored within the tensor.

The tensors are then passed through the BERT model to generate embeddings. Embeddings are vectorized representations of the token. They are generated by machine learning relationships between and within token structures, which allows for analysis between objects. Embeddings are generated through the PyTorch package in Python. Figure 6 denotes the steps performed to create embeddings.

Figure 6: Embedding creation process

For NLP purposes, embeddings allow relatedness quantification of concepts amongst given input. Embeddings assist models in assessing complex text through how data is stored. These token embeddings are then generalized as representations of the texts. Encoding representations of linguistics allows for many different NLP tasks to be performed. This process is the workhorse of the model as BERT stands for Bidirectional Encoder Representations from Transformers.

The final step in determining text similarity is calculating cosine similarity between the embeddings. Cosine similarity measures the closeness of two vectors and is bounded by negative one and one. A score of one denotes identical vectors, a negative one indicates two opposite vectors, whereas a zero result specifies an orthogonal relationship. In finding the cosine similarity, the dot product of the two embeddings (one being the transpose) is calculated while dividing the result by the product of embedding norms. Np.linalg.norm() is then used to compute embedding norms to achieve desirable output.

Output is saved into a dictionary for final analysis. Results are visualized by taking stored dictionary keys and values to produce a scatter plot within Python. The resulting image, labeled Figure 7, can be found in the Results/Findings section below. The 0 index indicates the March 2018 meeting, whereas the 45 index denotes November 2023. A complete list of meeting dates and their corresponding index can be found in Table 1 below.

Table 1: Index Value and Corresponding Meeting Dates

Index	Meeting Date	Index	Meeting Date
0	March 20-21 2018	23	January 26-27 2021
1	May 1-2 2018	24	March 16-17 2021
2	June 12-13 2018	25	April 27-28 2021
3	July/Aug 31-1 2018	26	June 15-16 2021
4	September 25-26 2018	27	July 27-28 2021
5	November 7-8 2018	28	September 21-22 2021
6	December 18-19 2018	29	November 2-3 2021
7	January 29-30 2019	30	December 14-15 2021
8	March 19-20 2019	31	January 25-26 2022
9	April/May 30-1 2019	32	March 15-16 2022
10	June 18-19 2019	33	May 3-4 2022
11	July 30-31 2019	34	June 14-15 2022
12	Septmber 17-18 2019	35	July 26-27 2022
13	October 29-30 2019	36	September 20-21 2022
14	December 10-11 2019	37	November 1-2 2022
15	January 28-29 2020	38	December 13-14 2022
16	March 17-18 2020	39	Jan/Feb 31-1 2023
17	April 28-29 2020	40	March 21-22 2023
18	June 9-10 2020	41	May 2-3 2023
19	July 28-29 2020	42	June 13-14 2023
20	Septmber 15-16 2020	43	July 25-26 2023
21	November 4-5 2020	44	September 19-20 2023
22	December 15-16 2020	45	Oct/Nov 31-1 2023

Results/Findings

Analysis of all text similarity scores indicates two statistical outliers. March 2018 and March 2022 are the meeting dates for these events. According to Doh et al. (2021), the first paragraph is critical in generating semantic differences from FOMC documents. The first paragraph of each outlier is examined and shown below. Surprise economic data can be interpreted as a reason for driving heterogeneity of the 2018 FOMC meeting documents. This was also the first meeting in which Jerome Powell acted as chairman. For March 2022, the Russian Invasion of Ukraine is considered the driving factor of dissimilarity between texts. After manually reading the full text of the minutes, these results are found for potential sentences that may be flagged as unique communication conveyed by the committee. Figure 7 presents noteworthy sentences in their corresponding minutes' opening paragraph.

March 2018 first paragraph	March 2022 first paragraph
"incoming economic data released in early February--particularly data on average hourly earnings--as raising concerns about the prospects for higher inflation and higher interest rates. These concerns reportedly contributed to a steep decline in equity prices and an associated rise in measures of volatility"	*"The manager turned first to a review of global financial market developments. Following the Russian invasion of Ukraine and the subsequent imposition of an array of sanctions, the ruble depreciated roughly 40 percent against the dollar. Prices of dollar-denominated Russian bonds plunged 80 to 90 percent"*

Figure 7: Sentences from FOMC minutes

The vector of text similarity scores between each minute and statement pair remained relatively constant. As shown in Table 1, the resulting vector comprised 45 observations and exhibited a mean of 0.8053 and a standard deviation of 0.0556. The median value is 0.8109. Comparatively, there was a large difference in text similarity between the two meetings, as denoted above. Text similarity metrics for the March 2022 meeting indicate a result of 0.644, whereas the March 2018 value is 0.631. With z scores of absolute value greater than three, the data points are labeled significant outliers as they are not within 99.73% of the sample. Identifying these events can drive insights into potential mechanisms driving a lack of transparency. Figure 8 exhibits the text-similarity of each meeting through visualization.

Table 1: Descriptive Statistics of Text Similarity Vector

N	Mean	Std. Dev	Median
46	0.8053	0.0556	0.8109

Discussion of Findings

These results are of interest to a wide variety of organizations as they can start to realize where, when, and how asymmetric information may occur regarding monetary policy. Findings indicate that larger macroeconomic shock events, such as the Russian invasion of Ukraine, may provide deviations in consistency amongst FOMC transparency. Institutions can understand that statements may not be the expected proxy of what is communicated in the minutes. The underlying mechanism for these deviations is not well understood. There could be many reasons, but the driving factor may be to keep human agents in the economic system calm and rational. If more damaging information is being communicated, financial volatility may occur, or people may act out of fear. This is the opposite of what the FOMC intends when they release policy adjustments or meeting documents.

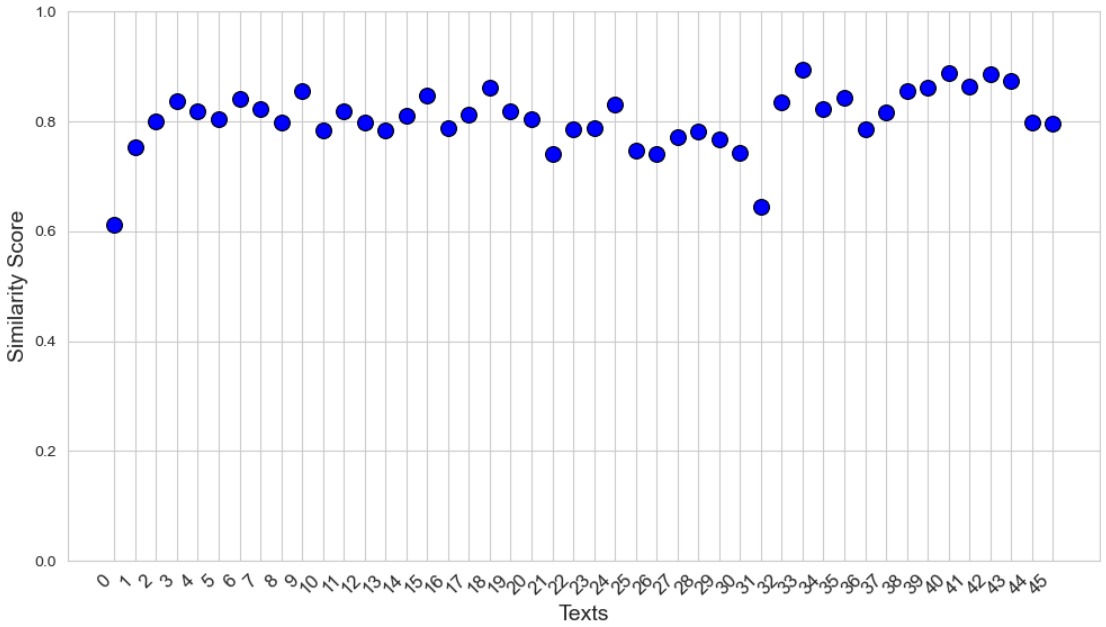

Figure 8: Text similarity from each meeting

Results were somewhat surprising as meetings surrounding the coronavirus pandemic of 2020 indicate no decreases in transparency. This may be due to the committee not releasing their standard pairwise documents for these special meetings. Text similarity results for meetings during the pandemic are comparable to the average, as showcased in Figure 1. Perhaps more dark and somber information, with the potential for significant volatility in the world financial system, will be somewhat masked from the statements. This will instill a sense of calmness in society that is needed to avoid downward economic spirals. It is also possible that the "heavier" discussions occurred during the special meetings in which no minutes were provided.

Of note is that the Israel-Hamas conflict, starting October 7, 2023, is included in the scope of this analysis as the November 2023 meeting is included. Interestingly, this military conflict did not impact consistency or transparency in the same way that the Russian invasion of Ukraine did. This result provides insights into the differing levels of macroeconomic shocks. The Russian/Ukraine event may have had more impact on FOMC discussion as Russia is a large global oil producer while Ukraine is a large international agricultural exporter. Significant economic sanctions imposed on Russia from the Western world may also drive these differences.

Conclusion

The methodology described in this paper is original for understanding text consistency amongst FOMC documents. Deep learning methods, such as implementing the BERT model, have not been utilized to abstractively summarize committee minutes and compare the resulting information to corresponding meeting statements employing text similarity measures. BERT was selected given its computational efficiency and high-performance measures with a broad range of NLP tasks, including abstractive summaries and vector representations for text similarity. The selection of BERT is a practical compromise between the need to increase time and computing power with respectable accuracy metrics. The limitations of BERT, stated in the following subsection, guide the summary approach.

The research results indicate that the FOMC is mainly consistent in its messaging to public markets. Out of a 46-meeting sample, only the March 2019 and March 2022 meetings were statistical outliers. The five unscheduled coronavirus special meetings during March 2020 were not included in the sample as the minutes were not released to the public. Of the two outlier meetings found, there is potential for seasonality to drive the result, as both meetings took place in March. Using previous research from Doh et al. (2021), stating that semantic differences are primarily attributed to information in the opening paragraphs, it is hypothesized that unforeseen economic volatility drove the March 2019 outlier, while the Russian invasion of Ukraine contributed to producing the March 2022 outlier meeting.

This paper makes the following contributions:

1. It provides a more immediate measure of FOMC transparency than previous studies by using minutes instead of transcripts (~257 weeks sooner).
2. It details the impact of unexpected economic data on the consistency of FOMC communication.
3. It highlights the potential for macroeconomic "shock" events to impact committee transparency, as evidenced by the Russian invasion of Ukraine.

Limitations

One major limitation of using BERT is that the maximum token length input is 512 due to computational constraints and memory limitations (Devlin et al., 2018). Thus, the entirety of the meeting minutes was not able to be input into the model for a single abstractive summary. Text truncation was utilized in response to this limitation. The first 512 tokens were kept for the first abstractive summarization, as previous research has stated the importance of the opening paragraph (Doh et al., 2021). The last 512 tokens were also included for a second abstractive summary as they generally contain salient information regarding monetary policy actions. For example, the committee's final decision regarding the federal funds rate is always toward the latter, or last paragraph, of the minutes. All output for the minutes and corresponding summary was verified in a Jupyter notebook.

In dealing with this limitation, the original text could be sliced into many separate pieces of 512 token length. The summaries could then be aggregated. The researcher did not feel this was appropriate as the summarized output will not summarize the total text. It would not gather the most salient information from the whole text but rather grab subtext from individual subsets of text. The resulting abstractive summary would be rather long and may contain important information in a few paragraphs but not the entire text. This is not a valid proxy of minutes summary and may add little value to organizations. While the chosen methodology for this paper, allowing tokens to be cut, is not ideal, it is deemed more favorable than the immediate alternatives, with many suggestions for future research.

Future research

Research acknowledging the 512-token limitation still utilizes BERT-based models as a baseline (Yang et al., 2020) to evaluate new NLP algorithms that allow longer token sequences, such as Big Bird (Zaheer et al., 2020). In this sense, the BERT analysis in this paper acts identically to a linear regression baseline in numeric-based modeling. Although the researcher rarely assumes a linear data-generating process, more appropriate algorithms have a reference to give its results context. It is suggested that future research be performed utilizing algorithms that allow for longer token sequences and include the entirety of the FOMC minutes.

Further, future research may implement techniques to achieve more insightful BERT baseline performance. One such method for exploration is to consider various data pre-processing methods to account for the 512-token limit (Hou et al., 2022) or manually train a summarizer model from an input training corpus (Ding et al., 2020). The latter is costly and potentially requires access to a particular type of machine for proper training (Limsopatham, 2021).

Society would benefit from future research determining if the text similarity metrics using all minutes' language correlate with transparency metrics found in Fischer et al.'s (2023) research examining statements with transcripts. Institutions could gather further insights into forecasting what may occur in the transcripts or adjustments in the federal funds rate. This information can give more clarity to institutions. Public markets could further benefit from additional research studying the transparency of minutes.

Studies hereafter could also perform diagnostic analytics to understand better why the Russian invasion of Ukraine was such a heterogeneous event in terms of information conveyed by the committee, especially compared to the Israel conflict of October 2023. Examining the financial and current accounts between nations could help extract insight. Sanctions and the Russian economy de-coupling from SWIFT may be driving mechanisms in interest rate and trade goods deviation and ultimately be of priority for monetary policy influencers. It could be interesting to relate these sanctions and trade movements with information communicated through text. Similar insights can be developed in the March 2019 outlier meeting. The surprising economic data from 2019 may be compared to forecasted market values.

References

Acosta, M. (2023). A new measure of central bank transparency and implications for the effectiveness of monetary policy. *International Journal of Central Banking*, *19*(3), 49-97.

Apel, M., Blix Grimaldi, M., & Hull, I. (2022). How much information do monetary policy committees disclose? Evidence from the FOMC's minutes and transcripts. *Journal of Money, Credit and Banking*, *54*(5), 1459-1490.

Bhatia, N., & Jaiswal, A. (2015). Trends in extractive and abstractive techniques in text summarization. *International Journal of Computer Applications*, *117*(6).

Bostrom, K., & Durrett, G. (2020). Byte pair encoding is suboptimal for language model pretraining. *arXiv preprint arXiv:2004.03720*.

Devlin, J., Chang, M. W., Lee, K., & Toutanova, K. (2018). Bert: Pre-training of deep bidirectional transformers for language understanding. *arXiv preprint arXiv:1810.04805*.

Ding, M., Zhou, C., Yang, H., & Tang, J. (2020). Cogltx: Applying bert to long texts. *Advances in Neural Information Processing Systems*, *33*, 12792-12804.

Doh, T., Kim, S., & Yang, S.-K. (2021). How You Say It Matters: Text Analysis of FOMC Statements Using Natural Language Processing. *Federal Reserve Bank of Kansas City Economic Review*, *106*(1), 25–40.

Farka, M., & Fleissig, A. R. (2012). The effect of FOMC statements on asset prices. *International Review of Applied Economics*, *26*(3), 387-416.

Federal Reserve. (n.d.). Meeting calendars and information [Webpage]. Retrieved from https://www.federalreserve.gov/monetarypolicy/fomccalendars.htm

Federal Reserve. (n.d.). Minutes of the Federal Open Market Committee (FOMC) [Webpage]. Retrieved from https://www.federalreserve.gov/faqs/minutes-federal-open-market-committee-fomc.htm

Fischer, E., McCaughrin, R., Prazad, S., & Vandergon, M. (2023). Fed Transparency and Policy Expectation Errors: A Text Analysis Approach. *FRB of New York Staff Report*, (1081).

Gardner, B., Scotti, C., & Vega, C. (2022). Words speak as loudly as actions: Central bank communication and the response of equity prices to macroeconomic announcements. *Journal of Econometrics*, *231*(2), 387-409.

Gorodnichenko, Y., Pham, T., & Talavera, O. (2023). The voice of monetary policy. *American Economic Review*, *113*(2), 548-584.

Gürkaynak, R. S., Sack, B. P., & Swanson, E. T. (2004). Do actions speak louder than words? The response of asset prices to monetary policy actions and statements. *The Response of Asset Prices to Monetary Policy Actions and Statements*.

Handlan, A. (2020). Text shocks and monetary surprises: Text analysis of fomc statements with machine learning. *Published Manuscript*.

Hansen, S., McMahon, M., & Prat, A. (2018). Transparency and deliberation within the FOMC: A computational linguistics approach. *The Quarterly Journal of Economics*, *133*(2), 801-870.

Hou, L., Pang, R. Y., Zhou, T., Wu, Y., Song, X., Song, X., & Zhou, D. (2022). Token dropping for efficient BERT pretraining. *arXiv preprint arXiv:2203.13240*.

Huang, Y. L., & Kuan, C. M. (2021). Economic prediction with the FOMC minutes: An application of text mining. *International Review of Economics & Finance*, *71*, 751-761.

Hugging Face. (2020). *BERT.* Retrieved from https://huggingface.co/transformers/v2.9.1/model_doc/bert.html

Kopchak, S. J. (2011). The liquidity effect for open market operations. *Journal of Banking & Finance*, *35*(12), 3292-3299.

Koroteev, M. V. (2021). BERT: a review of applications in natural language processing and understanding. *arXiv preprint arXiv:2103.1194*.

Limsopatham, N. (2021). Effectively leveraging BERT for legal document classification. *Proceedings of the Natural Legal Language Processing Workshop 2021* (pp. 210-216).

Liu, Y., & Lapata, M. (2019). Text summarization with pretrained encoders. *arXiv preprint arXiv:1908.08345*.

Lucca, D. O., & Trebbi, F. (2009). Measuring central bank communication: an automated approach with application to FOMC statements (No. w15367). *National Bureau of Economic Research*.

Mao, R., Lin, C., & Guerin, F. (2021). Combining pre-trained word embeddings and linguistic features for sequential metaphor identification. *arXiv preprint arXiv:2104.03285*.

Nakamura, E., & Steinsson, J. (2018). High-frequency identification of monetary non-neutrality: the information effect. *The Quarterly Journal of Economics*, *133*(3), 1283-1330.

Poole, W., & Rasche, R. H. (2003). The impact of changes in FOMC disclosure practices on the transparency of monetary policy: are markets and the FOMC better" synched"?. *Review-Federal Reserve Bank of Saint Louis*, *85*(1), 1-10.

Ramaswamy, M. (2022). Towards a framework for better decision-making in small businesses. *Issues in Information Systems*, *23*(3).

Rha, K. (2018). A Measure of Monetary Policy Transparency: How Newsworthy are the Actions of the FOMC?

Shapiro, A. H., & Wilson, D. J. (2021). Taking the Fed at its word: A new approach to estimating central bank objectives using text analysis. *The Review of Economic Studies*, *89*(5), 2768–2805.

Swanson, E. T. (2021). Measuring the effects of federal reserve forward guidance and asset purchases on financial markets. *Journal of Monetary Economics*, *118*, 32-53.

Tadle, R. C. (2022). FOMC minutes sentiments and their impact on financial markets. *Journal of economics and business*, *118*, 106021.

Taylor, J. B. (2001). Expectations, open market operations, and changes in the federal funds rate. *Review-Federal Reserve Bank of Saint Louis*, *83*(4), 33-48.

Thornton, D. L. (2012). The Dual Mandate: Has the Fed Changed Its Objective?. *Federal Reserve Bank of St. Louis Review*, *94*(2), 117-133

Yang, L., Zhang, M., Li, C., Bendersky, M., & Najork, M. (2020, October). Beyond 512 tokens: Siamese multi-depth transformer-based hierarchical encoder for long-form document matching. In *Proceedings of the 29th ACM International Conference on Information & Knowledge Management* (pp. 1725-1734).

Zaheer, M., Guruganesh, G., Dubey, K. A., Ainslie, J., Alberti, C., Ontanon, S., ... & Ahmed, A. (2020). Big bird: Transformers for longer sequences. *Advances in neural information processing systems*, *33*, 17283-17297.