Analysis of FOMC Statements Using Topic Modeling and Sentiment Analysis

Fatemeh Abbasian-Abyaneh

August 15, 2024

Abstract

This report presents an analysis of Federal Open Market Committee (FOMC) statements using text mining techniques, specifically Topic Modeling through Latent Dirichlet Allocation (LDA) and Sentiment Analysis. The goal is to derive insights into the policy stance and sentiment reflected in the FOMC statements over time.

Contents

1 Introduction

The period from 2015 onward has been marked by significant fluctuations in the U.S. economy. Economic indicators such as unemployment, inflation, and GDP growth have shown varying trends, reflecting the impact of global and domestic events, including the COVID-19 pandemic. The Federal Reserve, as the central banking system of the United States, plays a pivotal role in managing these economic conditions through its monetary policy. The Federal Open Market Committee (FOMC), a key body within the Federal Reserve, is responsible for setting the direction of this policy.

The FOMC holds regular meetings, typically eight times a year, during which it assesses economic conditions and makes decisions regarding interest rates and other monetary policies. These meetings result in the release of statements that summarize the Committee's economic outlook and the measures it intends to take. These statements are crucial as they not only signal the Fed's intentions to the markets but also influence household and business decisions across the country. The tone of these statements typically begins with a description of the overall economy, followed by the specific policy actions being implemented, whether they are expansionary or contractionary.

Expansionary policies generally involve lowering interest rates and purchasing securities to inject liquidity into the economy, aimed at stimulating growth and reducing unemployment. On the other hand, contractionary policies typically involve raising interest rates and selling securities to curb inflation and prevent an overheated economy. Understanding the nature and implications of these policies is essential for stakeholders across the economy.

Given the importance of FOMC statements, this study seeks to develop a text-based measure that can systematically capture the sentiment and policy stance conveyed in these communications. By analyzing the segments of these statements during significant economic and political events, this study aims to provide insights into the strictness and nature of the Federal Reserve's policies over time. The results could offer a valuable tool for forecasting economic conditions and understanding the Federal Reserve's response to various challenges.

The remainder of this report is organized as follows: Section 2 provides an overview of the U.S. economy from 2015 to 2024, highlighting key economic conditions and the Federal Reserve's responses. Section 3 outlines the methods used for data collection, preprocessing, and analysis, including sentiment analysis and topic modeling. The results of the analysis are presented in Section 4, followed by a discussion of the findings in Section 5. Finally, Section 6 concludes the report, summarizing the insights gained and their implications for understanding the Federal Reserve's policy stance over time.

2 Overview of the US Economy (2015-2024)

2.1 Economic Conditions and Federal Reserve Responses

2.1.1 2015 to 2016: Post-Great Recession Recovery

The US economy was still in a recovery phase following the Great Recession. The Federal Reserve's policies were primarily focused on supporting economic growth and stabilizing inflation, which remained below target. Interest rates were kept low to encourage borrowing and investment.

2.1.2 Late 2016 to Early 2018: Post-Election Optimism

Following the 2016 presidential election, there was a surge of optimism in the markets, driven by expectations of deregulation and tax cuts. The Federal Reserve began to cautiously raise interest rates to prevent the economy from overheating, reflecting confidence in the strength of the recovery.

2.1.3 2019: Trade Tensions and Economic Concerns

In 2019, the US economy faced uncertainties due to trade tensions, particularly with China. The Federal Reserve adjusted its policy stance by pausing rate hikes and eventually lowering rates to support economic stability amid these tensions.

2.1.4 2020: COVID-19 Pandemic

The COVID-19 pandemic triggered a severe economic downturn, leading the Federal Reserve to implement unprecedented measures, including cutting interest rates to near zero and launching extensive asset purchase programs. These measures aimed to provide liquidity to the financial system and support economic activity during the crisis.

2.1.5 2021 to 2022: Recovery and Inflation Concerns

As the economy began to recover, inflation emerged as a significant concern. The Federal Reserve maintained a supportive policy stance but started signaling potential rate hikes to curb rising prices. The labor market showed signs of improvement, but supply chain disruptions contributed to inflationary pressures.

2.1.6 2023 to 2024: Tighter Monetary Policy

With inflation remaining elevated, the Federal Reserve adopted a tighter monetary policy stance, characterized by successive rate hikes. This period marked the most aggressive tightening cycle in decades, aimed at bringing inflation back to target levels.

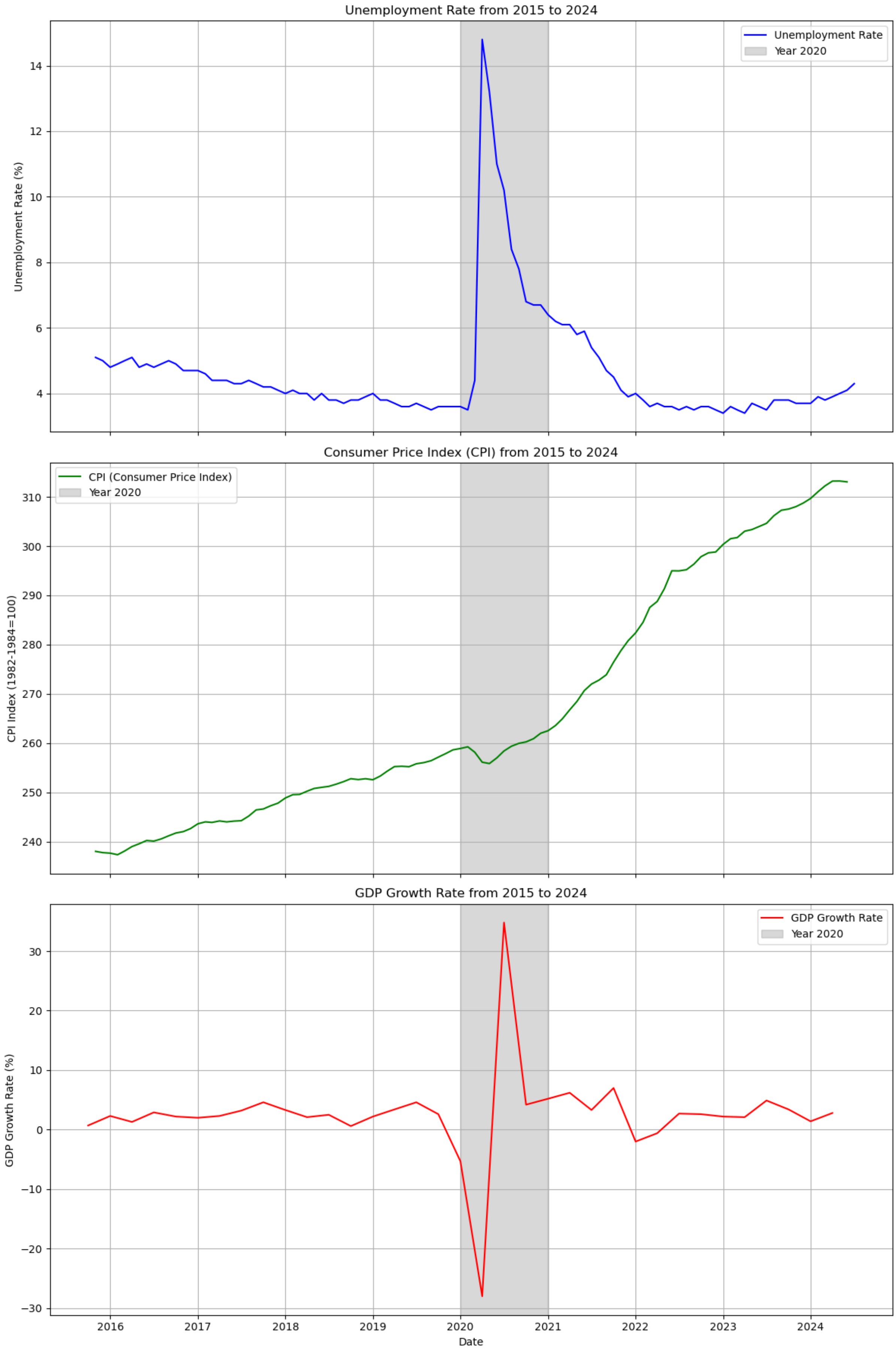


Figure 1: Economic Indicators of the US from 2015 to 2024: Unemployment, Inflation, and GDP Growth Rate. Source: FRED

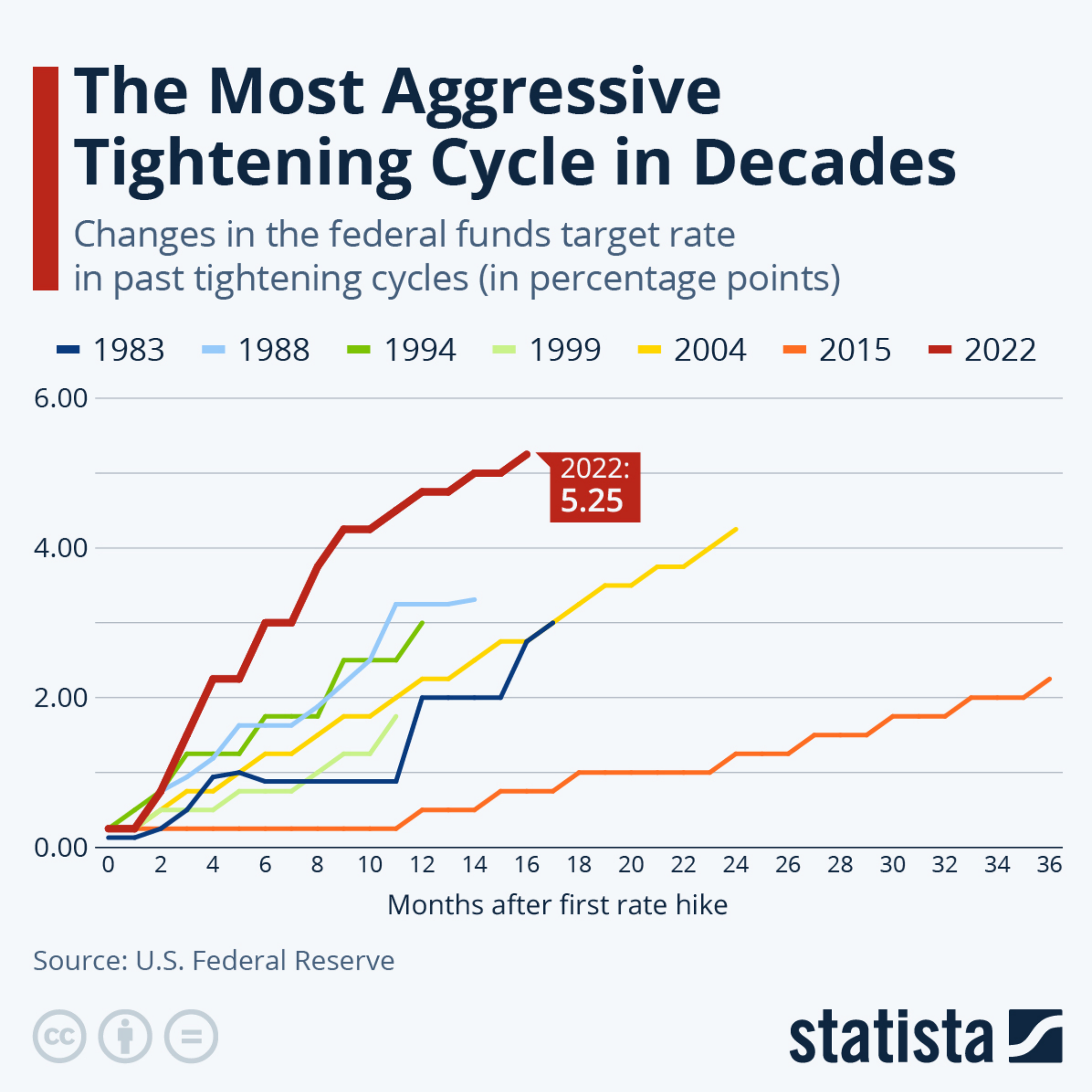


Figure 2: The Most Aggressive Tightening Cycle in Decades: Changes in the Federal Funds Target Rate (2022 vs. Past Cycles). Source: Statista

3 Methods

3.1 Data Collection and Preprocessing

In this study, FOMC statements from December 2015 to July 2024 were collected from the Federal Reserve's official website, which provides a comprehensive archive of these documents. The statements are generally released in January, March, April, May, September, November, and December each year. After downloading, the statements were saved with filenames in the format 'year_order of the statement in the year', and the corresponding file names and their months were recorded in an Excel sheet named *data* to facilitate easy retrieval during analysis.

FEDERAL RESERVE press release

For release at 2 p.m. EDT March 16, 2016

Information received since the Federal Open Market Committee met in January suggests that economic activity has been expanding at a moderate pace despite the global economic and financial developments of recent months. Household spending has been increasing at a moderate rate, and the housing sector has improved further; however, business fixed investment and net exports have been soft. A range of recent indicators, including strong job gains, points to additional strengthening of the labor market. Inflation picked up in recent months; however, it continued to run below the Committee's 2 percent longer-run objective, partly reflecting declines in energy prices and in prices of non-energy imports. Market-based measures of inflation compensation remain low; survey-based measures of longer-term inflation expectations are little changed, on balance, in recent months.

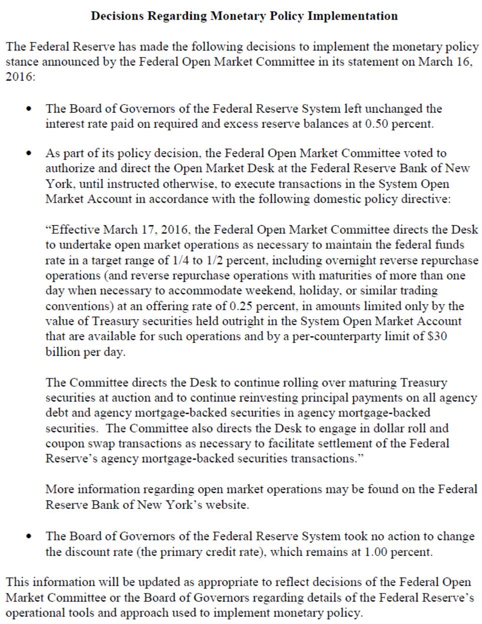
Consistent with its statutory mandate, the Committee seeks to foster maximum employment and price stability. The Committee currently expects that, with gradual adjustments in the stance of monetary policy, economic activity will expand at a moderate pace and labor market indicators will continue to strengthen. However, global economic and financial developments continue to pose risks. Inflation is expected to remain low in the near term, in part because of earlier declines in energy prices, but to rise to 2 percent over the medium term as the transitory effects of declines in energy and import prices dissipate and the labor market strengthens further. The Committee continues to monitor inflation developments closely.

Against this backdrop, the Committee decided to maintain the target range for the federal funds rate at 1/4 to 1/2 percent. The stance of monetary policy remains accommodative, thereby supporting further improvement in labor market conditions and a return to 2 percent inflation.

In determining the timing and size of future adjustments to the target range for the federal funds rate, the Committee will assess realized and expected economic conditions relative to its

Figure 3: Sample FOMC Statement - March 16, 2016

The structure of the FOMC statements typically involves an initial discussion on the economic sentiment, where the Committee provides an overview of the current economic conditions, including details on inflation, employment, and other key indicators. This is followed by a section titled *Decisions Regarding Monetary Policy Implementation*, where the specific policy measures decided during the meeting are detailed.



Decisions Regarding Monetary Policy Implementation

The Federal Reserve has made the following decisions to implement the monetary policy stance announced by the Federal Open Market Committee in its statement on March 16, 2016:

- The Board of Governors of the Federal Reserve System left unchanged the interest rate paid on required and excess reserve balances at 0.50 percent.

- As part of its policy decision, the Federal Open Market Committee voted to authorize and direct the Open Market Desk at the Federal Reserve Bank of New York, until instructed otherwise, to execute transactions in the System Open Market Account in accordance with the following domestic policy directive:

  "Effective March 17, 2016, the Federal Open Market Committee directs the Desk to undertake open market operations as necessary to maintain the federal funds rate in a target range of 1/4 to 1/2 percent, including overnight reverse repurchase operations (and reverse repurchase operations with maturities of more than one day when necessary to accommodate weekend, holiday, or similar trading conventions) at an offering rate of 0.25 percent, in amounts limited only by the value of Treasury securities held outright in the System Open Market Account that are available for such operations and by a per-counterparty limit of $30 billion per day.

  The Committee directs the Desk to continue rolling over maturing Treasury securities at auction and to continue reinvesting principal payments on all agency debt and agency mortgage-backed securities in agency mortgage-backed securities. The Committee also directs the Desk to engage in dollar roll and coupon swap transactions as necessary to facilitate settlement of the Federal Reserve's agency mortgage-backed securities transactions."

  More information regarding open market operations may be found on the Federal Reserve Bank of New York's website.

- The Board of Governors of the Federal Reserve System took no action to change the discount rate (the primary credit rate), which remains at 1.00 percent.

This information will be updated as appropriate to reflect decisions of the Federal Open Market Committee or the Board of Governors regarding details of the Federal Reserve's operational tools and approach used to implement monetary policy.

Figure 4: Decisions Regarding Monetary Policy Implementation - March 16, 2016

Given this structure, the text of each statement was split into two corpora: one containing the sentiment-related content and the other focused on policy decisions. These were saved as 'filename_semantic' and 'filename_policy', respectively. The sentiment corpus was then analyzed to assess the tone of the statements, determining the extent to which the language implied a pessimistic or optimistic outlook.

The policy corpus, on the other hand, was used to define a measure of how expansionary or contractionary the policy measures were. This involved detailed text preprocessing steps such as lowercasing, removing punctuation, tokenization, and lemmatization to improve the quality of the analysis. Additionally, for the policy analysis, the most frequent and least frequent words, along with month names, were removed to focus on the most informative terms.

The text preprocessing ensured that the subsequent sentiment and topic modeling analyses were conducted on high-quality, standardized text data, providing more reliable and interpretable results. Specifically, both the sentiment-related and policy-related corpora underwent standard preprocessing steps, including converting all text to lowercase, removing punctuation, tokenization (splitting text into individual words or tokens), and lemmatization (reducing words to their base or root form). These preprocessing steps are crucial for minimizing noise in the data and ensuring that the analyses focus on the most meaningful linguistic patterns within the FOMC statements.

3.2 Sentiment Analysis

3.2.1 Sentiment Scoring by Rule-Based Approach

Overview of Rule-Based Sentiment Analysis Rule-based sentiment analysis is a method used to quantify the sentiment expressed in a given text by assigning a numerical score. This score typically ranges from -1 to 1, where -1 indicates a highly negative sentiment, 0 represents neutrality, and 1 indicates a highly positive sentiment. The approach works by analyzing the words in the text and assigning scores based on predefined sentiment lexicons.

Initial Sentiment Analysis with TextBlob For the initial analysis, I utilized **TextBlob**, a Python library for processing textual data. TextBlob performs sentiment analysis by assigning polarity scores to words in the text using general-purpose lexicons. The scores are averaged to produce an overall sentiment score for each document. However, due to the semi-neutral and conservative tone often found in FOMC statements, TextBlob was not fully effective in capturing the pessimistic outlooks expressed in these texts.

Enhanced Sentiment Analysis with a Financial-Specific Lexicon Recognizing the limitations of general-purpose sentiment analysis tools for this specific context, I implemented a rule-based sentiment analysis approach using a financial-specific lexicon. The **Loughran-McDonald Sentiment Word Lists** was employed for this purpose. This lexicon is specifically designed for analyzing financial documents and categorizes words into several sentiment-related classes, such as:

- **Positive**: Words that convey optimism (e.g., "growth," "strong," "profit").
- **Negative**: Words that indicate pessimism or concern (e.g., "loss," "decline," "risk").
- **Litigious**: Words associated with legal issues (e.g., "lawsuit," "regulation").
- **Uncertainty**: Terms indicating ambiguity or unpredictability (e.g., "uncertain," "doubt").

- **Constraining**: Words implying limitations (e.g., "restrict," "limit").
- **Superfluous**: Words suggesting excess (e.g., "redundant," "overhead").

For the analysis, only the **Positive** and **Negative** categories were considered. The **Litigious**, **Uncertainty**, and **Constraining** categories were excluded because their frequent occurrence in such documents could reduce the significance of the overall sentiment analysis.

Calculation of Sentiment Scores The sentiment score for each FOMC statement was calculated as follows:

- **Text Preprocessing**: The text of each statement was preprocessed, which included:
  - **Lowercasing**: Converting all text to lowercase for consistency.
  - **Tokenization**: Splitting the text into individual words (tokens).
  - **Stopword Removal**: Removing common words that do not contribute to sentiment.
  - **Lemmatization**: Reducing words to their root form.
- **Lexicon Matching**: After preprocessing, the text was analyzed using the Loughran-McDonald lexicon. Words matching the positive and negative categories were identified, and their counts were summed to produce scores.
- **Sentiment Scoring**: The final sentiment score for each document was calculated by subtracting the negative score from the positive score. This net score was then normalized to a range between -1 and 1, providing a clear measure of the overall sentiment expressed in the statement.

Conclusion The use of the Loughran-McDonald financial-specific lexicon allowed for a more accurate and contextually relevant sentiment analysis of the FOMC statements, effectively capturing the nuanced tones that are typical of such financial documents. By focusing on positive and negative sentiments, the analysis provides a clearer picture of the Federal Reserve's outlook as conveyed in these critical communications.

Reference: Loughran-McDonald Master Dictionary

3.3 Policy Stance Analysis

3.3.1 Overview of Topic Modeling and LDA

Topic modeling is a powerful technique in text mining that helps identify abstract themes or "topics" across a collection of documents. The Latent Dirichlet Allocation (LDA) is one of the most widely used topic modeling algorithms. LDA is particularly useful in understanding the underlying themes within documents, which, in the context of Federal Open Market Committee (FOMC) statements, can be interpreted as expansionary or contractionary policy directions.

How LDA Works LDA assumes that each document in a corpus is a mixture of several topics, and each topic is represented by a distribution over words. Here's a brief overview of how LDA operates:

- **Document-Topic Distribution**: LDA posits that each document is composed of a mix of topics. For instance, an FOMC statement may cover topics related to inflation, employment, and monetary policy.
- **Topic-Word Distribution**: Each topic is characterized by a distribution of words. For example, a topic on inflation might frequently include words like ”inflation,” ”prices,” and ”costs.”
- **Inference Process**: LDA infers the topics and their distribution across documents by analyzing word co-occurrences. It assigns probabilities to the likelihood that a specific word in a document is associated with a particular topic.
- **Topic Assignment**: Once the model is trained, LDA assigns a topic distribution to each document, showing the proportion of each topic within the document.
- **Interpretation**: The identified topics are then interpreted by analyzing the words most closely associated with each topic. For instance, a topic with words like ”stimulus,” ”support,” and ”growth” may be indicative of an expansionary policy stance, while a topic with words like ”hike,” ”restrict,” and ”control” may suggest a contractionary stance.

Application in Policy Stance Scoring By categorizing topics within FOMC statements as either expansionary or contractionary, LDA enables the quantification of the policy stance for each document. Over time, this allows for tracking the evolution of the Federal Reserve's policy direction.

3.3.2 Methodology

Refinement of Word Frequencies in LDA To ensure that the LDA model's results were meaningful and unbiased, the word frequencies in the text were refined by removing high-frequency and low-frequency words. This step was essential for improving the distinction between topics.

- **Removal of High-Frequency Words (Above 78th Percentile)**: Common terms that appear frequently across all documents were removed to prevent them from skewing the topic distributions. These words often include non-informative terms that could dilute the significance of more meaningful content-specific words.
- **Removal of Low-Frequency Words (Below 5th Percentile)**: Words that occur very infrequently were also removed as they might introduce noise and lead to overfitting. These rare words are often not representative of broader trends and could distort the topic modeling process.
- **Impact on LDA Model**: The refined word set allowed for a more balanced representation of topics, ensuring that the LDA model identified topics based on words that reflect their true importance in the context of FOMC statements.

Normalization of LDA Scores Between -1 and 1 To make the policy stance scores comparable across different documents, the raw LDA scores were normalized to a scale between -1 and 1:

- **Calculation of Raw Scores**: For each document, the score was calculated based on the proportion of words associated with expansionary and contractionary topics.
- **Normalization Formula**: The raw score was normalized using the following formula:

$$\text{Normalized Score} = \frac{2 \times (\text{Raw Score} - \text{Min Score})}{\text{Max Score} - \text{Min Score}} - 1 \tag{1}$$

3.3.3 Identified Topics

After tuning the LDA model, the following three distinct topics were identified:

Table 1: Top 30 Bigrams for Each Identified Topic

Topic 1	Topic 2	Topic 3
excess	purchase	monthly
maturity	functioning	aggregate
trading	smooth	minimum
holiday	needed	bid
accommodate	cmb	reinvest
similar	chair	standing
available	increased	needed
convention	temporarily	stated
value	discretion	modest
weekend	release	reason
held	sustain	deviation
outright	support	maturing
limited	authorize	received
authorize	establishment	extent
rolling	existing	bill
maturing	level	le
reinvesting	reinvest	redeem
including	excess	discretion
received	commercial	temporarily
small	funding	increased
acceptable	additional	chair
reason	deviation	approved
deviation	reason	establish
release	modest	raise
level	stated	director
point	edt	submitted
approved	term	request
existing	est	percentage
establishment	pace	taking
taking	current	point

Interpretation of Topics Topic 1 and Topic 3 are associated with contractionary measures, focusing on aspects such as tightening liquidity and controlling inflation. Topic 2 is linked to expansionary measures, reflecting actions aimed at stimulating economic growth.

Scoring Mechanism The policy stance score for each document was calculated as the weighted probability of topics in each document, multiplied by their assigned sign (where contractionary topics are weighted as -1 and expansionary topics as +1). The final score indicates the overall policy stance, with positive values suggesting expansionary measures and negative values indicating contractionary measures.

Note: Custom stop words (e.g., "part," "related," "allow," etc.) and months' names were removed to enhance the distinction between topics.

4 Results

4.1 Sentiment Analysis Results

The results of the sentiment analysis using a financial-domain-specific lexicon have been saved in the `sentiment_score` Excel sheet. Below are the visualizations of the average sentiment scores over time, as calculated using both general sentiment analysis and the financial-specific lexicon.

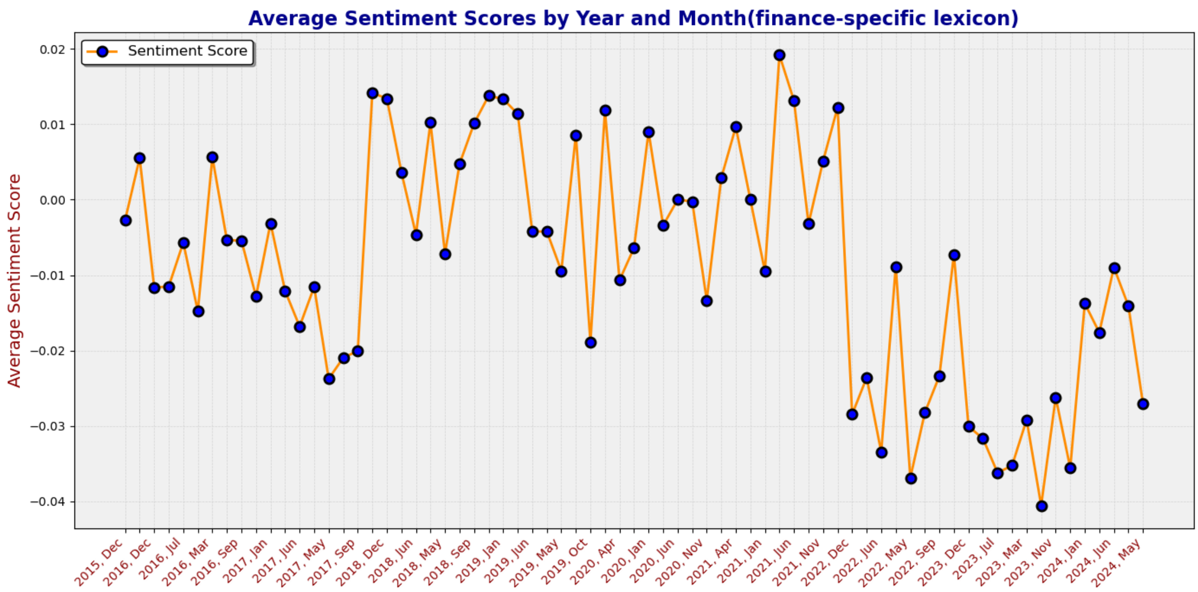


Figure 5: Average Sentiment Scores by Year and Month (financial-specific lexicon)

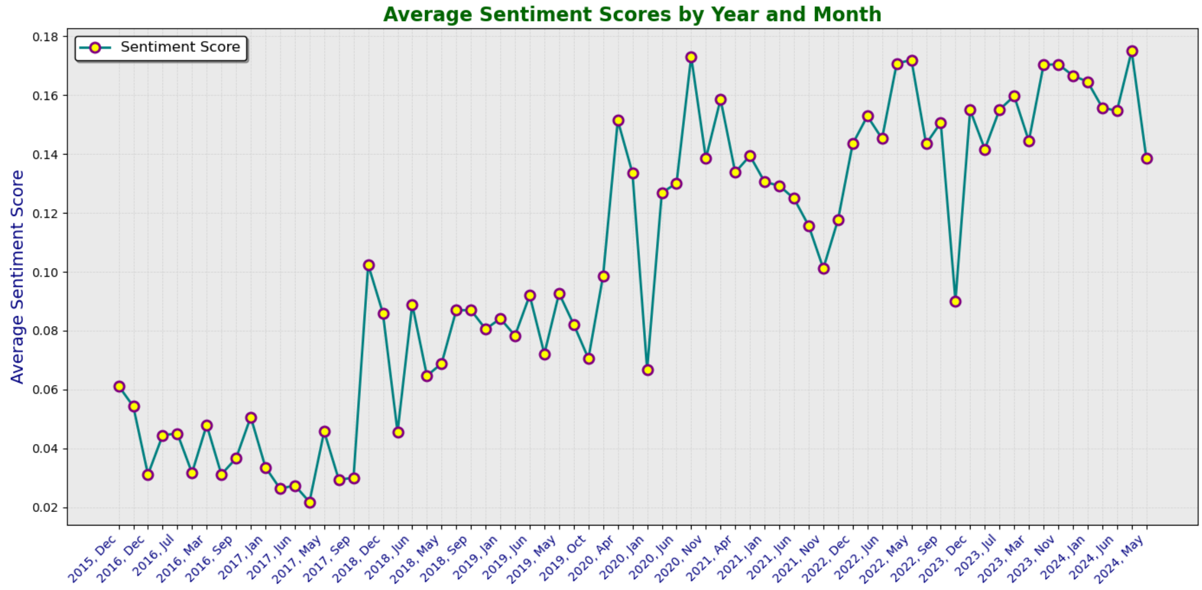


Figure 6: Average Sentiment Scores by Year and Month (general sentiment analysis)

4.1.1 Event-Based Comparison and Interpretation of the Methods

2015 to 2016: Post-Great Recession Recovery

- **General Sentiment Analysis:** The sentiment during this period shows relatively stable but slightly positive trends, reflecting a neutral to optimistic interpretation of the economic conditions.
- **Financial-Specific Lexicon:** Captures more fluctuations and occasional negative sentiment, reflecting the FOMC's cautious outlook due to global uncertainties like China's slowdown and Brexit.

Late 2016 to Early 2018: Post-Election Optimism

- **General Sentiment Analysis:** The sentiment scores during this period reflect a strong upward trend, indicating broad optimism following the 2016 U.S. Presidential Election.
- **Financial-Specific Lexicon:** While also showing optimism, this method exhibits more variability, highlighting the FOMC's balance between post-election optimism and ongoing economic risks.

2019: Trade Tensions and Economic Concerns

- **General Sentiment Analysis:** The sentiment scores show some fluctuations but remain generally positive, potentially underestimating the impact of trade tensions.
- **Financial-Specific Lexicon:** More pronounced dips in sentiment, indicating the FOMC's concern about the potential negative impact of U.S.-China trade tensions on the global economy.

2020: COVID-19 Pandemic

- **General Sentiment Analysis:** A sharp decline in sentiment is observed during the pandemic, reflecting the economic shock.
- **Financial-Specific Lexicon:** Captures a more severe drop in sentiment, accurately reflecting the deep concerns and unprecedented economic disruption caused by COVID-19.

2021 to 2022: Recovery and Inflation Concerns

- **General Sentiment Analysis:** High sentiment scores during this period indicate strong optimism as the economy began to recover.
- **Financial-Specific Lexicon:** Shows mixed sentiment with noticeable dips, reflecting both the economic recovery and emerging inflation concerns, providing a more nuanced interpretation.

2023 to 2024: Tighter Monetary Policy

- **General Sentiment Analysis:** A slight decrease in sentiment is observed, indicating some caution but overall positive trends.
- **Financial-Specific Lexicon:** More pronounced negative sentiment is observed, reflecting the FOMC's increasing caution as it tightens monetary policy to combat inflation, signaling concern over potential economic slowdowns.

4.1.2 In Summary

The financial-specific lexicon provides a more accurate and detailed reflection of the FOMC's sentiment, particularly in capturing caution and negative outlooks during periods of economic uncertainty, such as trade tensions and the COVID-19 pandemic. In contrast, the general sentiment analysis tends to lean towards a more positive or neutral interpretation, potentially missing critical nuances. The financial-specific approach better aligns with actual economic events, making it a more reliable tool for understanding the FOMC's outlook.

4.1.3 Limitations and Areas for Improvement

While the rule-based approach using a financial-specific lexicon is effective in capturing nuanced sentiment in FOMC statements, it has certain limitations:

- **Static Lexicon:** The lexicon is static and may not evolve to capture new terminologies or changes in the financial language over time.
- **Context Ignorance:** The approach may not fully capture the context in which certain words are used, leading to potential misclassification of sentiment.
- **Granularity:** The sentiment is analyzed at the word level rather than at the phrase or sentence level, which might overlook the overall sentiment conveyed by the combination of words.

Future improvements could involve incorporating machine learning-based sentiment analysis techniques that can adapt to new language trends and better understand the context in which words are used.

4.2 Policy Stance Result

The results of the policy stance analysis, based on Latent Dirichlet Allocation (LDA) topic modeling, have been saved in the `policy_score` Excel sheet. Given that policy stance tends to be more lethargic to change compared to sentiment, the scores have been averaged yearly. This provides a more significant measure of the policy stance over time. The graph below illustrates the average yearly policy stance scores, normalized between -1 and 1, where -1 indicates a strongly contractionary stance and 1 indicates a strongly expansionary stance.

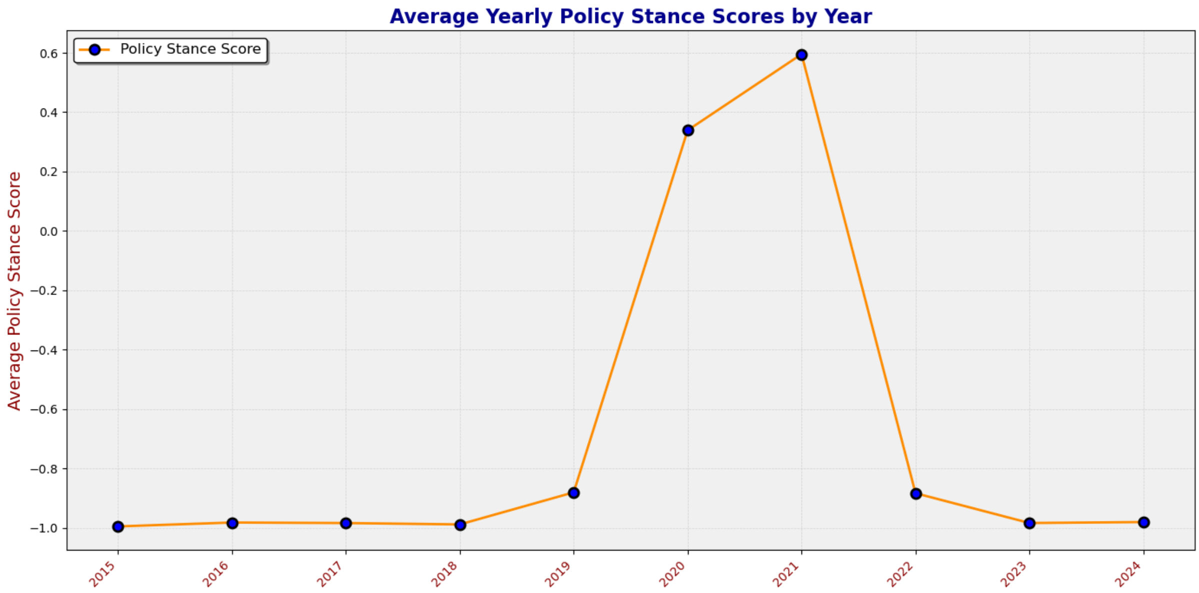


Figure 7: Average Yearly Policy Stance Scores by Year

4.2.1 Analysis of Policy Stance in the USA Based on LDA Model and Historical Context

The graph above, generated from Latent Dirichlet Allocation (LDA) topic modeling of Federal Open Market Committee (FOMC) statements, illustrates the evolution of the U.S. monetary policy stance from 2015 to 2024. Each point on the graph represents the average policy stance score for a given year.

Periods of Policy Stance and Their Historical Context

- **2015-2019: Gradual Tightening (Contractionary Policy)**
  - **Duration:** This period marks the aftermath of the Great Recession, during which the Federal Reserve gradually transitioned from an expansionary stance to a more contractionary policy. The Fed began to raise interest rates from historically low levels, reflecting improved economic conditions and concerns over potential inflation.

- **Initiating Events:** The recovery of the U.S. economy, falling unemployment rates, and the need to normalize monetary policy after years of quantitative easing.
- **Policy Strictness:** This period was characterized by cautious tightening, with the Fed gradually increasing interest rates to avoid derailing the economic recovery.

- **2020-2021: Emergency Expansion (Expansionary Policy)**
  - **Duration:** The most dramatic shift in the policy stance occurred in 2020-2021, as the COVID-19 pandemic struck the global economy. The Fed responded with aggressive expansionary policies, including slashing interest rates to near zero and implementing large-scale asset purchases.
  - **Initiating Events:** The global pandemic led to a sharp economic downturn, necessitating emergency measures to stabilize financial markets and support economic activity.
  - **Policy Strictness:** The expansionary measures during this period were unprecedented in scale, aimed at preventing a deep and prolonged recession.

- **2022-2024: Aggressive Tightening (Contractionary Policy)**
  - **Duration:** Starting in late 2021 and continuing through 2024, the Fed shifted back to a contractionary stance in response to surging inflation, which reached multi-decade highs. This period is marked by rapid interest rate hikes and the reduction of the Fed's balance sheet through quantitative tightening.
  - **Initiating Events:** Persistent inflation, driven by supply chain disruptions, labor shortages, and high consumer demand post-pandemic, triggered this aggressive tightening.
  - **Policy Strictness:** The contractionary policies during this period were notably harsher compared to the 2015-2019 period, as the Fed acted decisively to control inflation.

Comparison of Policy Strictness Across Periods

- **2015-2019 vs. 2022-2024:** Both periods were contractionary, but the strictness of the policy stance is markedly different. The gradual tightening from 2015 to 2019 was moderate and cautious, designed to avoid jeopardizing the economic recovery. In contrast, the post-2021 tightening was much more aggressive, with rapid interest rate hikes reflecting the urgency of addressing high inflation.

- **2020-2021 Expansion vs. Other Periods:** The expansionary stance during 2020-2021 was the most aggressive in recent history, driven by the unique challenges of the pandemic. The scale and speed of the measures taken were far greater than any expansionary actions during previous economic downturns.

Compatibility and Sensibility of the LDA Model Results The LDA model, by categorizing the language of FOMC statements into distinct topics, successfully captures the overall shifts in the Federal Reserve's policy stance over the years. The model's results, as depicted in the graph, align well with the historical narrative:

- The model accurately reflects the gradual tightening from 2015-2019, followed by a sharp expansionary phase in 2020-2021, and then an aggressive contractionary phase from 2022 onwards.
- The clear peaks and troughs in the graph correspond to the major policy shifts in response to economic conditions, demonstrating the model's ability to identify significant changes in monetary policy language.

Areas for Improvement in the Model While the LDA model provides a valuable overview of policy stance changes, there are some limitations to consider:

- **Nuanced Detection of Policy Strictness:** The model may not fully capture the varying degrees of strictness within contractionary or expansionary periods. For example, the harsher contractionary stance from 2022 onwards compared to 2015-2019 may not be as distinctly represented due to the similar language used across both periods. To address this, a more granular model could incorporate additional context or data, such as economic indicators, to better differentiate the intensity of policy measures.
- **Impact of Seemingly Neutral Language:** FOMC statements often contain neutral or cautious language that may not fully convey the strictness of policy actions. This could lead to some periods appearing less differentiated in the model than they are in practice. Incorporating sentiment analysis or advanced linguistic features might help in capturing subtle shifts in tone that correlate with policy strictness.
- **Data Availability:** The model's effectiveness is partly constrained by the data it analyzes. If FOMC statements lack explicit language reflecting the severity of policy changes, the model may underrepresent these shifts. Expanding the dataset to include minutes of FOMC meetings or speeches by Federal Reserve officials could provide a richer source of information for analysis.

Conclusion The LDA model provides a coherent and sensible framework for analyzing the evolution of U.S. monetary policy stance over time. While the model successfully captures major policy shifts, further refinements could enhance its ability to detect nuanced differences in policy strictness, especially during periods of aggressive tightening or expansion. Incorporating additional data sources and advanced linguistic techniques could improve the model's accuracy and depth, making it a more powerful tool for understanding the Federal Reserve's approach to monetary policy.

5 Discussion

Discussion

This section provides a critical analysis of the results obtained from both sentiment and policy stance analyses. It examines the alignment of these results with historical economic events, highlights the strengths and limitations of the methodologies employed, and suggests potential areas for improvement in future research.

5.1 Alignment with Historical Events

The results from both sentiment and policy stance analyses demonstrate a clear alignment with significant economic events that have shaped the U.S. economy from 2015 to 2024. For instance, the sharp decline in sentiment scores during the COVID-19 pandemic reflects the severe economic downturn and uncertainty during that period. Similarly, the aggressive contractionary stance observed in the policy stance analysis post-2021 aligns with the Federal Reserve's response to surging inflation.

5.2 Strengths of the Analyses

The methodologies employed in this study offer several advantages:

- **Sentiment Analysis:** The use of a financial-specific lexicon allowed for a more nuanced interpretation of the sentiment expressed in FOMC statements, particularly during periods of economic uncertainty. This approach provided a more accurate reflection of the cautious and often neutral tone typical of such official communications.
- **Policy Stance Analysis:** The application of LDA topic modeling effectively captured the shifts in monetary policy over time, allowing for a systematic and objective categorization of FOMC statements into expansionary or contractionary stances. The normalization of scores between -1 and 1 provided a clear and interpretable measure of policy direction.

5.3 Limitations of the Methodologies

Despite the strengths, there are limitations to consider:

- **Rule-Based Sentiment Analysis:** While the financial-specific lexicon improved the accuracy of sentiment detection, the rule-based approach may still miss subtle nuances in language, particularly when the statements employ neutral or mixed tones. The conservative nature of FOMC language can lead to underestimating the sentiment intensity.
- **LDA Topic Modeling:** Although LDA was successful in identifying key topics, it may not fully capture the varying degrees of strictness within contractionary or expansionary policies. This is particularly evident in the post-2021 period, where the aggressive tightening measures might not be distinctly represented due to the similarities in language across different periods.
- **Data Granularity:** The analysis was constrained by the granularity of the data available. FOMC statements alone may not fully encapsulate the Federal Reserve's policy stance, as they are often supplemented by meeting minutes, speeches, and other communications. Including these additional sources could provide a more comprehensive view of the policy stance.

5.4 Suggestions for Improvement

To enhance the robustness and depth of the analyses, the following improvements are suggested:

- **Advanced Sentiment Analysis:** Incorporating more sophisticated sentiment analysis techniques, such as machine learning models trained on financial text, could capture more nuanced sentiment shifts. These models could also account for the contextual meaning of words in different economic environments.
- **Enhanced Topic Modeling:** Refining the LDA model to account for varying degrees of policy strictness could improve the differentiation between contractionary and expansionary periods. Incorporating economic indicators or using a hierarchical topic model might provide additional insights.
- **Inclusion of Additional Data Sources:** Expanding the dataset to include FOMC meeting minutes, speeches by Federal Reserve officials, and other related documents could offer a richer context for the analysis. This would likely improve the accuracy and interpretability of both sentiment and policy stance scores.

6 Conclusion

This study applied advanced text mining techniques to analyze Federal Open Market Committee (FOMC) statements, providing insights into the sentiment and policy stance over the past decade. The results align well with historical economic events, demonstrating the effectiveness of using financial-specific sentiment analysis and LDA topic modeling in understanding monetary policy direction. While the methodologies employed have proven useful, there are opportunities for further refinement, particularly in capturing the nuances of policy strictness. Overall, this analysis offers a valuable framework for assessing the Federal Reserve's communications and their impact on economic expectations.

7 References

- Federal Reserve Economic Data (FRED), Federal Reserve Bank of St. Louis. Available at: `https://fred.stlouisfed.org/`
- Board of Governors of the Federal Reserve System. Federal Open Market Committee (FOMC) Calendars, Statements, and Minutes. Available at: `https://www.federalreserve.gov/monetarypolicy/fomccalendars.htm`
- Statista, "The Most Aggressive Tightening Cycle in Decades." Available at: `https://www.statista.com/chart/28437/interest-rate-hikes-in-past-tightening-cycles/`
- Loughran-McDonald Sentiment Word Lists, University of Notre Dame. Available at: `https://sraf.nd.edu/loughranmcdonald-master-dictionary/`