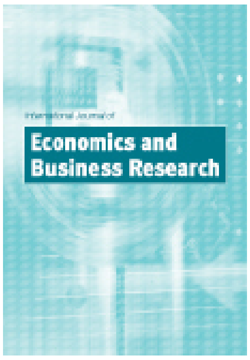


International Journal of Economics and Business Research

ISSN online: 1756-9869 - ISSN print: 1756-9850
https://www.inderscience.com/ijebr

Extracting the Federal Reserve's priorities and transparency from FOMC statements using textual data mining tools

Reyadh Faras, Faleh Alshameri, Ahmad Bash

DOI: 10.1504/IJEBR.2025.10072392

Article History:

Received:	20 April 2025
Last revised:	09 June 2025
Accepted:	16 June 2025
Published online:	20 July 2025

Extracting the Federal Reserve's priorities and transparency from FOMC statements using textual data mining tools

Reyadh Faras*

College of Business Administration,
Kuwait University,
P.O. Box 5486, Safat 13055, Kuwait
Email: reyadh.faras@ku.edu.kw
*Corresponding author

Faleh Alshameri

School of Business,
University of Maryland Global Campus,
Adelphi, Maryland, USA
Email: Faleh.alshameri@faculty.umgc.edu

Ahmad Bash

College of Business Studies,
The Public Authority for Applied Education and Training (PAAET),
P.O. Box 23167, Safat 13092, Kuwait
Email: ay.bash@paaet.edu.kw

Abstract: In this paper, we aim to investigate the priorities of the Federal Reserve (Fed) and transparency under three different chairpersons. Specifically, we investigated the chairperson's influence on monetary policy over time and what issues were considered when setting new monetary targets under each chairperson. Data came from the statements of Federal Open Market Committee (FOMC) meetings from March 2006 to May 2024. The sample period covered the chairmanships of Ben Bernanke, Janet Yellen, and current chairman Jerome Powell. Textual data mining is a tool used to extract terms that correspond to several economic issues that the Fed may consider when setting monetary policy during FOMC meetings. We tracked the frequency of mentioning certain terms related to economic issues throughout 146 FOMC statements. The results show that Fed Chairman Bernanke has completely different ranking priorities from the other two. Yellen and Powell have identical priorities rankings (as revealed in the FOMC statements). Overall, inflation and the financial market were the highest priorities under all three chairmanships. With respect to transparency, Bernanke's term has the lowest level, and Powell had the highest level, followed by Yellen.

Keywords: Federal Reserve; FOMC; inflation targeting; monetary policy; text mining; clustering; information extraction.

Reference to this paper should be made as follows: Faras, R., Alshameri, F. and Bash, A. (2025) 'Extracting the Federal Reserve's priorities and transparency from FOMC statements using textual data mining tools', *Int. J. Economics and Business Research*, Vol. 29, No. 15, pp.23–46.

Biographical notes: Reyadh Faras is an Associate Professor of Economics at Kuwait University. Before that he worked at different public and international institutions inside and outside of Kuwait. His areas of concentration include monetary economics, public finance, international trade, political economy, and competition. He has many published papers in refereed scientific journals. He holds a PhD in Economics from West Virginia University, two Masters in Economics from Florida State University and University of Colorado, and a Bachelor's in Economics from Kuwait University.

Faleh Alshameri is a Professor of Computer Information Systems. He received his PhD from the Volgenau School of Engineering and Information Technology, George Mason University. His research interests include text mining, image mining, data science and big data analytics. He has presented and published number of research papers in refereed journals and conferences.

Ahmad Bash is an Associate Professor of Finance at the Public Authority for Applied Education and Training (PAAET), College of Business Studies, Insurance and Banking Department. He has received his BA in Finance and Financial Institutions from Kuwait University; Master of Economics from Kuwait University; Master of Risk Management from Southampton University; Master of Commercial Law from La Trobe University; and PhD in Finance from RMIT University.

1 Introduction

Central banks' communications with the public have emerged in recent decades as a new tool of monetary policy used to manage economic expectations. This new tool has significantly impacted financial markets and contributed to lowering market volatility, exchange rate volatility, output variation, and inflation volatility (Weber, 2019). It has also recently garnered significant attention from scholars (Andrade and Ferroni, 2021; Del Negro et al., 2023; Swanson, 2021). In the past, central banks implemented their economic policy with great secrecy, believing that monetary policy is more effective when it surprises the markets (Dotsey, 1987). Currently, in developed and developing countries, many central banks regularly use several channels to communicate with the markets. In addition, the central banks' communications have been playing an important role in influencing economic outcomes and shaping market expectations (Blinder et al., 2008). For example, Rafferty and Tomljanovich (2002) found that high levels of policy transparency by central banks is associated with a marked enhancement of forecasting that will eventually lead to improvement in market efficiency. Such a finding contradicts the perspective that claims that transparency lowers market efficiency.

Being the most important (and influential) monetary authority in the world, the actions of the Federal Reserve (Fed) have a powerful impact on the world economy, financial markets, and global monetary policy. For decades, financial experts and academics tried to predict the monetary policy that would be implemented by the Fed, which was not explicitly revealing its future intentions. This changed when the Fed started issuing statements following its periodic meetings in which it revealed its future monetary policy, and most importantly, the targeted policy rate, or federal funds rate. This change has led to enormous growth in research analysing the content of Federal Open Market Committee (FOMC) statements.

Moreover, the Fed communicates qualitative information to the public and to financial markets through statements, minutes, speeches by officials, and published reports. Thus, the increased commitment to transparency, especially since the mid-1990s, has directed the focus as much on the Fed's words as on its actions as indicators of changes in future policy (Boukus and Rosenberg, 2006). Bernanke et al. (1999) described central bank transparency from the perspective of inflation-targeting in nine countries. Blinder et al. (2001) also provided a detailed analysis of the transparency of five central banks.

It is well documented in the literature that financial markets are highly responsive to monetary policy changes, most importantly to changes in the interest rate (i.e., federal funds rate). The response used to be substantial when the Fed was not issuing the FOMC statements, especially when policy changes were not anticipated. Bernanke and Kuttner (2005) found that an unexpected 25-basis-point cut in the federal funds rate would lead to an increase in stock prices of one percent.

However, it is believed that the timely and detailed statements issued on the FOMC meetings have helped reduce the ambiguity of monetary policy future direction, which minimised the volatility in financial markets responses to policy changes. Nevertheless, markets continued to be responsive to the statements issued but in a more predictable way. Many studies have focused on how Fed communication affects the expectations of market participants and asset prices. For example, Gürkaynak et al. (2005) examined how monetary policy affects asset prices, finding that FOMC statements significantly affect stock prices and bond yields. Adding that the size of the effect depends, in part, on forward guidance of changes in futures prices, which is inferred from the FOMC statements. This confirms the fact that markets seriously take the statements as signals of the direction of future monetary policy (Campbell et al., 2012; Moessner, 2013; Swanson and Williams, 2014).

This paper contributes to the literature by providing new insights from analysing FOMC statements in two directions. First, we are interested in identifying the priorities that the Fed emphasises in its statements. For each chairperson's tenure, we investigated which priority they included the most in their FOMC statements. Second, we assess the level of Fed transparency regarding disclosing the future directions of monetary policy by considering the amount of information included in the FOMC statements.

Our results have several implications. First, they provide valuable understanding of how the Fed communicates to the financial markets and the public in terms of its monetary policy and the role that the Fed plays in minimising uncertainty. For example, our results will help investors be more capable of anticipating the market path. Second, our results will relatively straightforwardly help businesses and consumers predict inflation and growth directions in the future.

We analysed the content of statements from the FOMC's meetings, which occur eight times a year, from March 2006 to May 2024. Using text mining tools, we found that Yellen and Powell had quite different economic priorities from Bernanke. However, Yellen and Powell's priority rankings were the same as each other's (as seen in the

FOMC statements). Under all three chairmanships, the financial sector and inflation were given top priority.

The remainder of the paper is structured in the following manner: Section 2 presents a comprehensive overview of the existing literature about central bank communication and the use of text mining. Section 3 outlines the data and the research methodology we employed. Section 4 presents the results and analysis. Finally, Section 5 concludes the paper.

2 Literature review

2.1 FOMC statements

Central banks, through their issued statements, provide additional information that enhances the common approach of solely counting on movements in the policy rates to reflect changes in the stance of monetary policy (Siklos, 2020). Information plays a key role in asset-pricing theory because it affects investors' expectations. According to the efficient-market hypothesis by Eugene Fama, prices should reflect all available information. Evans and Lyons (2005) found that the arrival of news affects end-user participants such as hedge funds, mutual funds, and non-financial corporations. Kryvtsov and Petersen (2021) found that central bank communication effects and stabilises individual and aggregate outcomes.

The FOMC started in the early 1980s, intending to hold eight regularly scheduled meetings per year during which its members would discuss the economic outlook and formulate monetary policy. Whatever policy change is decided at the meeting is instantly implemented through open-market operations. However, before 1994, there was no public announcement of these policy changes, and market participants could only know about them by the size and direction of the open market operations on the next day (Jegadeesh and Wu, 2017). Things changed in January 1994 when policy changes were made public in a short statement released immediately after the meeting. Moreover, during each meeting, detailed records of the discussions are kept, then summarised in the form of meeting minutes (i.e., FOMC statements), which are released to the public three weeks after the policy decision date (Tumala and Omotosho, 2019).

FOMC statements contain information about many aspects of the economy, including interest rates, the money supply, inflation, unemployment, and economic growth (Rohlfs et al., 2016). In studying the European Central Bank communications during policy decisions, Andrade and Ferroni (2021) found that two elements from high-frequency monetary surprises have a significant effect: information about upcoming macroeconomic conditions (Delphic shocks) and information about upcoming monetary policy shocks (Odyssean shocks). The FOMC is recognised as an important market mover, and many researchers have found that equity and interest rate markets frequently respond to the release of FOMC communications (Farka and Fleissig, 2013; Gürkaynak et al., 2005; Mueller et al., 2017; Rosa, 2011).

Analysing the content of documents (i.e., content analysis) is defined as "the method of studying the communication process and quantifying the content of documents in terms of intensity and direction of meaning" [Mazis and Tsekrekos, (2017), p.180], has been applied in much research in fields such as the political sciences, social sciences, finance, and economics. Content analysis is a relatively new field of research, around three decades old. Initially, researchers used manual techniques; then with the advancement of text mining (as will be shown in the next subsection), new computer-based techniques were introduced to identify hidden messages and measure sentiment. Content analysis can be used to link certain actions of central banks or predict the consequences of these actions (Mazis and Tsekrekos, 2017).

Early attempts to analyse the content of the Fed statements (or announcements) used a manual approach. In one of the first papers, Cook and Hahn (1988) analysed the impact of the information content of discount rate announcements on market interest rates. They found that the Fed systematically used certain types of discount rate announcements to signal changes in its policy instrument, the federal funds rate. Romer and Romer (1989) used the FOMC statements to discern the Fed's intentions by looking for both a clear statement of a belief that the current level of inflation needed to be lowered and an indication of the consequences on output that would bring inflation to the desired level.

2.2 Textual data mining

The manual approach had major drawbacks: it was subjective, labour intensive, and affected by reader bias (Boukus and Rosenberg, 2006). Criticism of the manual approach to content analysis led to a new automated approach known as textual data mining. It includes various computational tools and statistical techniques commonly used to quantify text (Bholat et al., 2015).

Textual data mining identifies patterns in natural language by labelling documents with keywords and analysing their co-occurrence frequency, effectively extending numeric data-mining techniques to handle unstructured or partially structured text (Al-Hassan et al., 2013).

Information and communication technology advancements, combined with artificial intelligence, have enabled researchers to develop innovative methods for identifying and extracting valuable insights from the vast amounts of data generated by increased web usage. This has facilitated information navigation, summarisation, and organisation to uncover intriguing relationships within large and semi-structured datasets through various computational techniques (Alshameri and Green, 2020). The currently used text mining techniques provide researchers with valuable tools to extract information from text documents. There are multiple uses of text mining, including information retrieval, information extraction, document classification, and document comparison (Hendry and Madeley, 2010). Given its usefulness, it has been widely used in economics and finance as a tool to uncover important information in published statements and social media (Chahadah et al., 2018; Gupta et al., 2020; Cicekyurt and Bakal, 2025).

Over the last decades, different automated machine learning algorithms have been developed. One of the most used approaches is the latent Dirichlet allocation (LDA) (Blei et al., 2003). The LDA model effectively identifies topics and extracts relevant features from extensive textual datasets through a three-layer Bayesian network comprising a document layer, a topic layer, and a word layer. Each text reflects a mixture of topics, and each topic comprises a probability distribution of words (Cheng et al., 2019).

Edison and Carcel (2021) employed LDA to examine and categorise the transcripts of the FOMC statements from 2003 to 2012, including a total of 45,346 passages across eight distinct topics. The findings indicated that conversations regarding economic modelling were particularly prevalent during the global financial crisis (GFC), with a notable rise in discussions pertaining to the banking system in the years following the GFC.

Other approaches include latent semantic analysis (LSA) (Landauer et al., 1998), valence aware dictionary and sentiment reasoner (VADER) (Hutto and Gilbert, 2014), and financial sentiment analysis with bidirectional encoder representations from transformers (FinBERT) (Devlin, 2018).

Clustering algorithms offer a relational framework that enhances analytical processes by revealing characteristics that impart significance and identify the most critical features (Al-Hassan et al., 2013). The development of fast and high-quality document-clustering algorithms can significantly improve the efficiency of extracting relevant and useful information from vast amounts of data (Alshameri et al., 2012). Compact and significant clusters are structured to facilitate effective navigation and browsing experiences. These clusters significantly enhance the retrieval process through techniques such as cluster-driven dimensionality reduction, term-weighting, and query expansion. They serve to elucidate the characteristics of data distribution while simultaneously illustrating the interrelationships among datasets. This dual function not only enhances data analysis and extraction but also enriches the contextual understanding of the work involved (Alshameri and Green, 2020).

Data clustering is a significant methodology for categorising analogous data items. By structuring extensive textual information into coherent clusters, individuals can gain a comprehensive overview of the data. Clustering techniques aim to uncover natural groupings within textual documents, ensuring that the clusters demonstrate a high degree of similarity within themselves while maintaining a low degree of similarity among different clusters (Al-Hassan et al., 2013). Text clustering is an automated technique that organises and categorises text segments into distinct groups. The primary objective of this clustering process is to identify and classify the similarities among various physical or abstract entities. As a standard example of unsupervised learning, text clustering encompasses several methodologies, including partition-based, hierarchical, density-based, grid-based, and model-based approaches.

A considerable number of studies have focused on the relationship between central bank communication and macroeconomic outcomes. Shapiro and Wilson (2019) utilised text mining techniques to scrutinise the transcripts of FOMC statements, minutes, and speeches delivered by FOMC members between 1986 and 2013. They derived a measure of monetary policy sentiment from the analysed documents, revealing that negative sentiments expressed by the FOMC were inversely correlated with the growth of output in the USA.

Siklos (2020) employed two distinct content analytic methodologies, namely Wordscores and DICTION, to examine six decades of FOMC statements. Wordscores assesses the content in relation to a predetermined benchmark whereas DICTION utilises a specialised algorithm to identify various aspects of sentiment and tone within the text. The analysis reveals a significant correlation between the content of FOMC statements and economic indicators, particularly real GDP growth and fluctuations in the federal funds rate. Notably, this relationship evolved post-1993, coinciding with the public release of the statements with a delay.

Huang and Kuan (2021) undertook a sentiment analysis of the FOMC statements, utilising text mining techniques to derive sentiment indicators. They applied an adaptive Bayesian methodology to construct sentiment indicators corresponding to each of the Federal Reserve's mandates. The empirical findings reveal that the indicators tailored to specific mandates display unique patterns, effectively highlighting the FOMC's policy priorities across various timeframes. Furthermore, the analysis demonstrates that these indicators possess predictive capabilities for economic variables, yielding enhanced out-of-sample forecasting performance.

An automated text mining procedure was established utilising the Bank of International Settlements repository of speeches delivered by senior executives of central banks. This methodology aimed to analyse and compare the objectives and strategies of various central banks, including the Federal Reserve, the European Central Bank, the Bank of England, and the Reserve Bank of Australia, spanning from the late 1990s to 2016. A series of indicators were introduced to assess the intensity of speeches across five distinct domains: monetary conditions, financial stability, external competitiveness, labour and social conditions, and economic activity (Ernst and Merola, 2018).

Tumala and Omotosho (2019) employed text mining methodologies to investigate the monetary policy communication strategy of the Central Bank of Nigeria from 2004 to 2019. Various text mining techniques have been employed in contexts pertinent to monetary policy. These applications include analysing central bank communications during banking reform initiatives, predicting financial instability through sentiment derived from textual data, evaluating sentiment scores in reports on financial stability, and exploring the relationship between financial stability and stock market returns, among other relevant topics.

In general, text mining has been used in studies to analyse the content of the FOMC statements in three main areas. The first area is to infer the Fed sentiment (or tone) with respect to future monetary policy, whether it is hawkish, dovish, or neutral (Cannon, 2015; Hubert and Fabien, 2017; Kahveci and Odabas, 2016; Mazis and Tsekrekos, 2017; Shapiro and Wilson, 2019; Tadle, 2022; Arismendi-Zambrano et al., 2024; Chau et al., 2025). The second use is to measure the response of financial markets to the FOMC statements (Boukus and Rosenberg, 2006; Farka and Fleissig, 2013; Gidofalvi and Elkan, 2001; Gu et al., 2022; Lucca and Moench, 2015; Lucca and Trebbi, 2009; Osowska and Wójcik, 2024; Schumaker et al., 2012). The third and final area, which is also the focus of this study, analyses the topics of interest that the Fed highlighted in the FOMC statements (Boukus and Rosenberg, 2006; Edison and Carcel, 2021; Jegadeesh and Wu, 2017; Kim et al., 2023; Siklos et al., 2018). Of these areas, the last one has received the least attention from researchers. Therefore, we are motivated to delve further into investigating the topics of interest of the Fed because of the lack of coverage by other researchers.

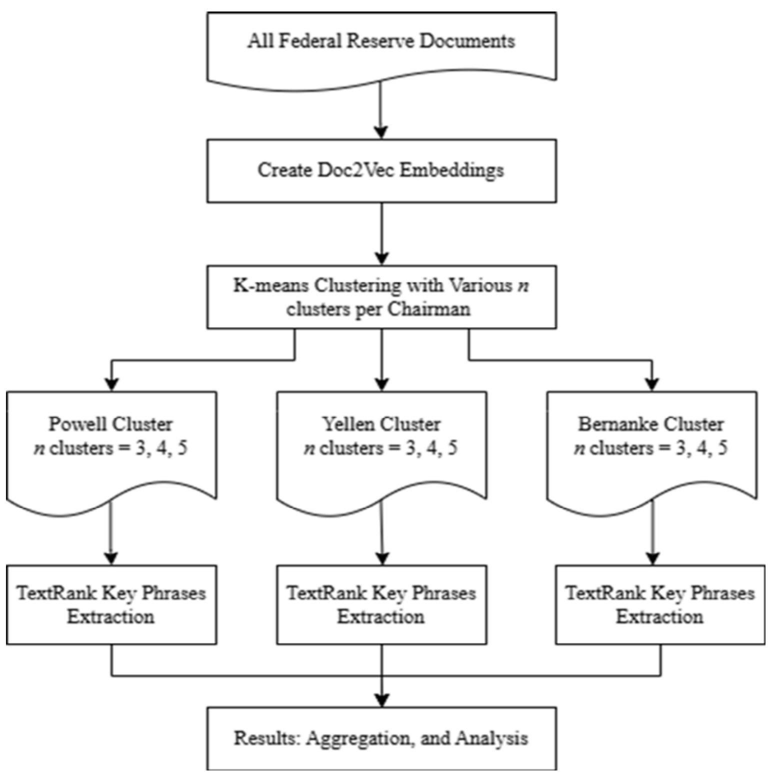
3 Data and research methodology

In this study we are interested in extracting the key phrases of FOMC statements from March 2006 to May 2024. The sample period covers the full chairmanship of Ben Bernanke (March 2006–January 2014), Janet Yellen (March 2014–January 2018), and current chairman Jerome Powell (March 2018–present). We analysed a total of 146 statements spanning the three chairmanships, 64, 32 and 50, respectively. The dataset publicly available from the Fed's official website. The code and additional materials used in this analysis are available from the authors upon reasonable request to support replication and future research.

In analysing the content of the FOMC statements, we used the textual data mining technique, which extracts certain words and phrases and categorises them in groups that are called clusters. Then we analysed these clusters to obtain useful information, which we used to draw conclusions about the priorities of the Fed under the three chairmanships that we are covering in this study, Bernanke, Yellen, and Powell.

The 146 statements were partitioned based on the term of each chairperson into various clusters using the K-means clustering method. The more sophisticated analysis involves three steps, creating document embeddings, clustering these document embeddings for each chairperson, and performing key phrase extractions for each document in each cluster for each set of clustering results. Figure 1, describes the research methodology framework, which provides a summary of the steps used in the research methodology, starting from the initial phase where we collected the dataset (146 statements), then the K-means clustering algorithm we used to cluster the dataset into different clusters to extract the useful information and the key phrases from these clusters. The last phase of the research methodology is the stage where we analysed and aggregated the results to show the main topics discussed in the FOMC' statements and Fed's priorities. The following research methodology subsections provide more detail information about the research methodology steps used in this study.

Figure 1 Research methodology framework



3.1 Text embeddings and clustering

Before we started the clustering process, we needed to create a form of text embedding such as TF-IDF term frequency-inverse document frequency, a commonly used simple document representation. Text embedding models are designed to translate sequences of text into embedding vectors. These embedding vectors are valuable because they represent a concept of semantic similarity; inputs that share similar meanings should have embeddings that are situated close together in vector space (Morris et al., 2023). Embeddings are frequently utilised for a variety of tasks, including search, clustering, and classification (Muennighoff et al., 2022).

Our exploratory analysis shows that these documents are in various but protracted lengths, high in repetitive words, and rich in semantics. Therefore, we chose to use a different embedding that could produce a dense and ideally fixed-length vector for each document that was not sensitive to word frequency and had notions of semantics. The Doc2Vec algorithm was the ideal approach for this analysis. Doc2Vec represents a neural network-driven methodology that facilitates the acquisition of distributed representations for documents. This technique operates under an unsupervised learning paradigm, assigning each document a fixed-length vector within a high-dimensional space. The learning process ensures that documents with similar content are positioned close to one another in this vector space. This allows for comparing documents through their vector representations, enabling various applications such as document classification, clustering, and similarity assessment.

Upon creating the document embeddings, we started clustering these documents using the K-means algorithm. The K-means algorithm is a commonly used unsupervised learning technique designed to sort observations into n clusters in which similarity is maximised within each cluster. Because we were interested in examining external context in this study, we first divided the dataset by chairperson. Then for each partition, we performed K-means clustering with n clusters, for examples 3, 4, and 5 clusters.

3.2 Key phrases extraction

After the documents clustering process for each chairperson, we were interested in summarising and describing the clusters with key phrases from the original documents. To achieve that, we employed the TextRank algorithm to extract key phrases from each paragraph in each document, per set of clustering results, for example, key phrases for Chairwoman Yellen from clustering results (n = 3). TextRank is a graph-based ranking algorithm that originated from PageRank, a ranking algorithm for web pages used by search engines. Instead of web pages, TextRank algorithms process word embedding and/or sentence embedding and rank these embeddings with a similarity matrix. Some typical steps in this TextRank key phrases extraction involve creating the phrase embeddings for paragraphs in each document, calculating a normalised cosine similarity matrix M for those embeddings, and then using the PageRank algorithms to compute the rank of each phrase. For each paragraph, we selected the top five ranked phrases as the candidate phrases, and then for each document, we selected the top 50 phrases from the candidate phrases from the paragraphs.

3.3 Data cleaning process

In the field of text mining, it is standard procedure to eliminate the introductory statements from each document and all chart annotations, thereby retaining only the core text. The FOMC statements exhibited considerable variation in length, spanning from 4,028 to 19,157 words and comprising between 42 and 366 paragraphs. Following the exclusion of common high-frequency and stop words – like ‘a’, ‘the’, ‘of’, ‘for’, and ‘with’, which are considered irrelevant for clustering – the dataset’s total word count was reduced to 915,362. Before this cleaning process, the total word count reached 1,598,365. The number of stop words varied in each statement within the dataset.

To prevent inaccuracies in counting and to mitigate semantic discrepancies, we recognised compound words and the specialised terminology frequently employed in these statements. It is essential to recognise compound words to comprehend the meaning of a text effectively. For example, the term ‘rate’ was found more than 2,000 times within these documents; however, without additional context, its specific reference remained ambiguous. By contrast, the phrases ‘interest rate’, ‘unemployment rate’, ‘federal fund rate’, and ‘inflation rate’ possess distinctly different interpretations, rendering their respective frequency significantly more informative. Likewise, the term ‘market’ occurred more than 1,000 times whereas the compound expressions ‘labour market’, ‘financial market’, ‘open market committee’, and ‘foreign exchange market’ carry specific connotations. These occurrences underscore the importance of systematic organisation and compilation to index compound phrases and other essential terminology accurately.

4 Results and analysis

4.1 Most common terms

As a first step in analysing the Fed’s priorities during the March 2006–May 2024 period under the leadership of the three chairpersons (Bernanke, Yellen, and Powell), we identified the most common words mentioned in the FOMC statements. We tried four combinations of term words and n-grams (i.e., unigram, bi-gram, tri-gram, and four-gram). We found that the two-word terms (bi-grams) were the best, given that most economic terms comprise two words (e.g., economic growth, financial market, monetary policy, foreign markets, exchange rates, unemployment rate).

The results in Figures 2, 3, and 4 show the ten most common bi-gram terms appearing in the FOMC statements under the three Fed chairpersons. As expected, ‘federal funds/ funds rate’ was the most common bi-gram and ranked first in all chairmanships. It appeared 960, 869, and 617 times in the Powell, Yellen, and Bernanke statements, respectively. This is not a surprise because the federal funds rate is the policy rate used by the Fed to implement monetary policy. The second most common bi-gram was ‘labour market’, which ranked second during Bernanke and Yellen’s terms. The other most common terms were ‘economic activity’, ‘monetary policy’, ‘inflation’, and ‘financial market’.

Figure 5, features a word cloud that encapsulates the most frequently mentioned words in our dataset. This graphical representation serves to accentuate significant terms by varying their font sizes based on their relative frequency (Tumala and Omotosho, 2019). Upon reviewing Figure 5, it is evident that terms like 'inflation', 'rate', 'market', 'federal', 'economic', 'policy', 'monetary', and 'labour' stand out prominently. This finding supports our earlier insights regarding the Fed Monetary Policy.

Figure 2 Most common (bi-gram) terms (Bernanke: 2006–2014) (see online version for colours)

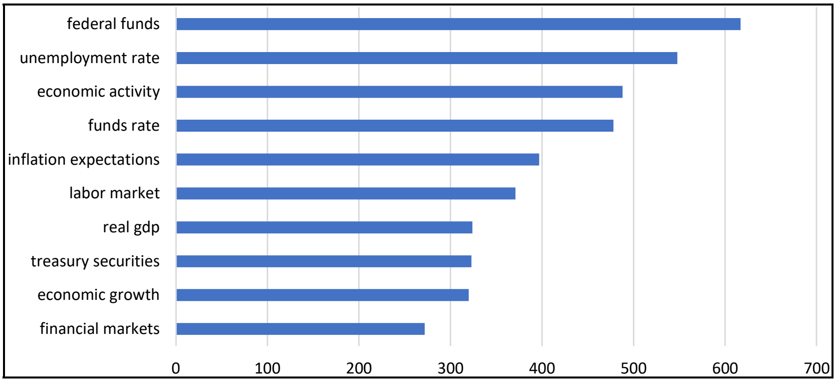


Figure 3 Most common 'bi-gram' terms (Yellen: 2014–2018) (see online version for colours)

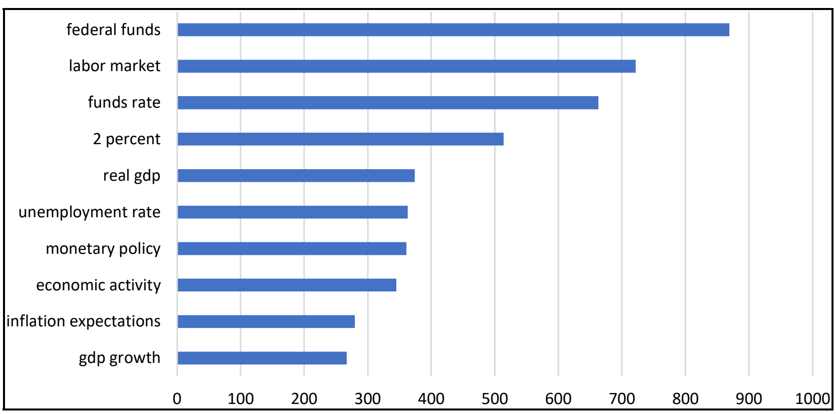


Figure 4 Most common 'bi-gram terms' (Powell: 2018–2024) (see online version for colours)

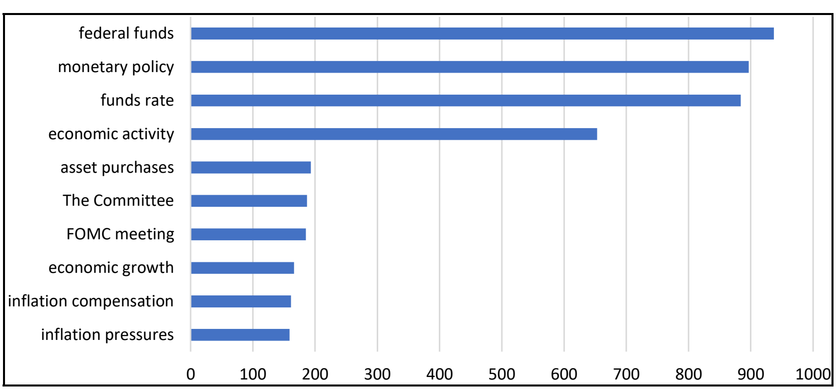
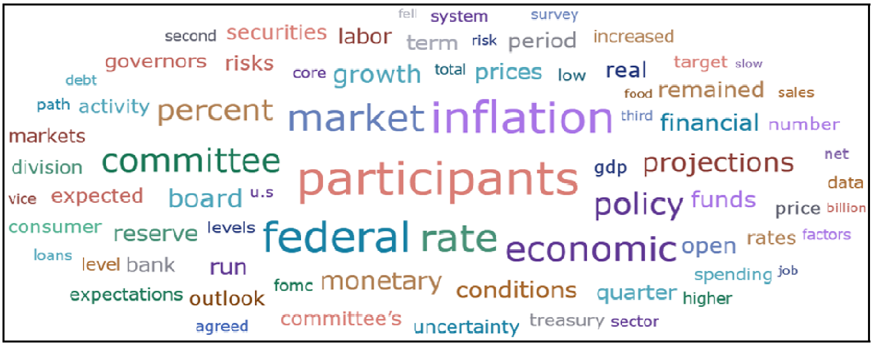


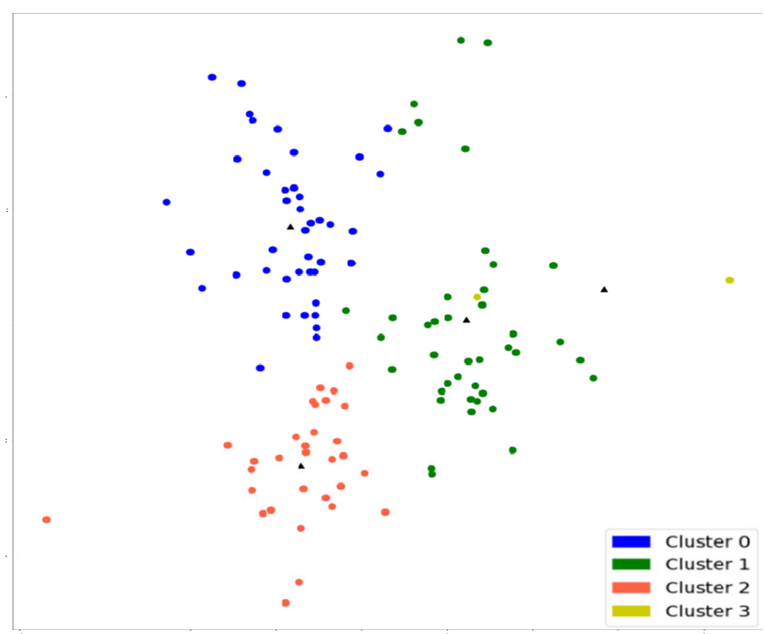
Figure 5 Word Cloud of the statements March 2006–May 2024 (see online version for colours)



4.2 Dominant clusters

To effectively partition the dataset comprising 146 statements into meaningful subsets, we aimed to identify the most significant characteristics of the data. This involved analysing the similarities within each subset (or cluster), simultaneously elucidating the distinctions between different clusters. We experimented with various configurations regarding the number of clusters and ultimately decided to focus on four primary clusters. This approach maximised intra-cluster similarity while minimising inter-cluster similarity. Figure 6, shows the distribution of the FOMC statements into four clusters based on the similarities between the statements. Each cluster has several statements. The total number of statements in the four clusters is 146 which is equal to the total number of statements in the dataset, this means that each statement belongs to only one cluster.

Figure 6 The FOMC clustering distribution (see online version for colours)



The second step in analysing the Fed's priorities was to dive deeply into the terms used in the statements' content related to the six topics (or issues) that the statements included. We categorised the common bi-gram terms into different clusters. Then each cluster was labelled based on the common terms in each cluster (e.g., macroeconomy, inflation, financial market). Figures 7, 8, and 9 show the dominant cluster of each FOMC statement. As shown in the figures, most of the time the dominant cluster changes between a statement and the one following it or after two statements.

Figure 7 FOMC statements: dominant clusters (Bernanke: 2006–2014) (see online version for colours)

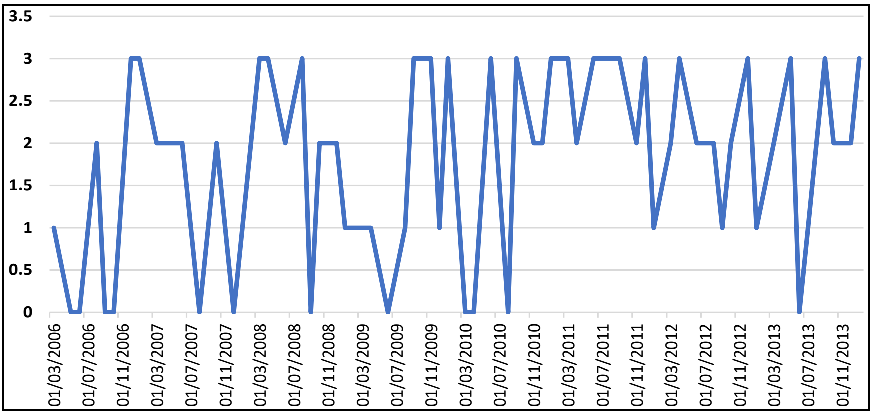
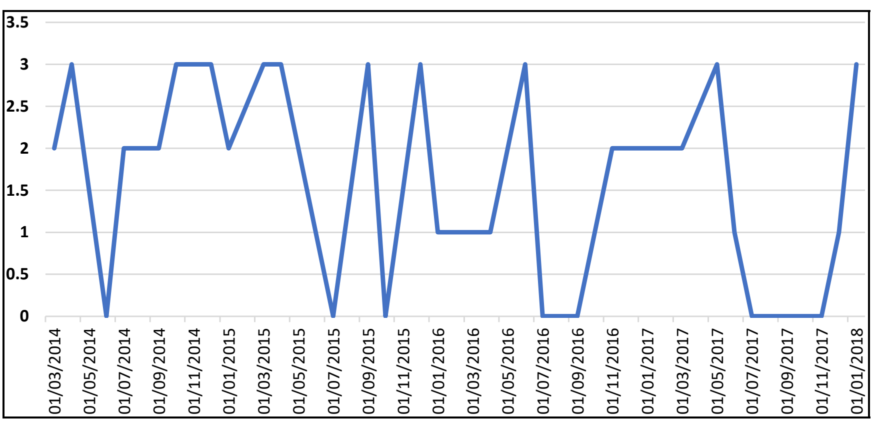


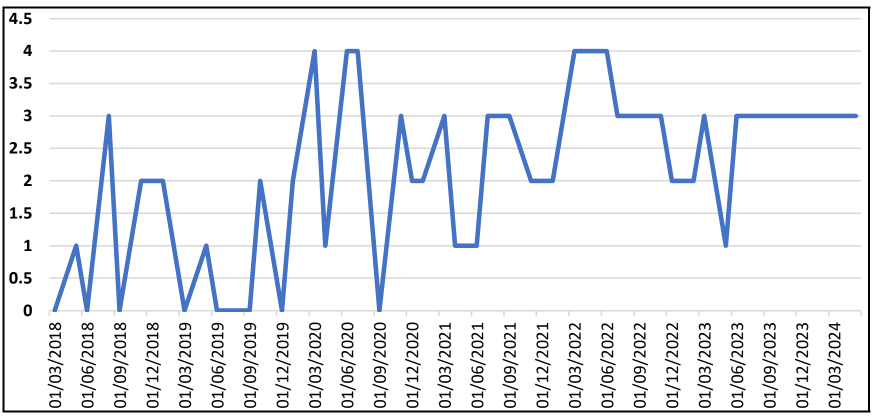
Figure 8 FOMC statements: dominant clusters (Yellen: 2014–2018) (see online version for colours)



However, on fewer occasions, a given cluster dominated more than two consecutive statements. It happened only twice during Bernanke's term when cluster 1 dominated three consecutive statements (January–April 2009) and cluster 3 did also (June–September 2011). During Yellen's term, it happened three times when cluster 1 dominated three statements (January–April 2016), cluster 2 dominated four statements (November 2016–March 2017), and cluster 0 dominated three statements (July–November 2017).

The Powell term was distinguished by persistent domination of the clusters from the FOMC statements. Nine times a given cluster dominated between three and six consecutive statements. A given cluster dominated three statements five times: cluster 0 (June 2019–September 2019), cluster 3 (October 2019–January 2020 and November 2021–January 2022), cluster 2 (September–December 2022), and cluster 1 (January–May 2024). Additionally, a given cluster dominated four consecutive statements twice: cluster 3 (November 2018–March 2019) and cluster 1 (March–July 2020). Moreover, one time a given cluster dominated five consecutive statements: cluster 1 (March–September 2021). Last, on one occasion only, cluster 2 dominated six consecutive statements (May–December 2023).

Figure 9 FOMC statements: dominant clusters (Powell: 2018–2024) (see online version for colours)



Overall, the above analysis shows that the Bernanke term was the least stable (or persistent) term with respect to continuity of clusters among the three chairmanships. In contrast, the Powell term was significantly the most stable (or persistent) term. This shows that there are significant differences among the three chairpersons, as evident in the dominance of clusters over the FOMC statements.

Table 1a The distribution of FOMC statements (Bernanke: 2006–2014)

Cluster	*Statements*
0	May 2006, June 2006, September 2006, October 2006, August 2007, December 2007, September 2008, June 2009, March 2010, April 2010, August 2010, and June 2013
1	March 2006, September 2007, January 2008, January 2009, March 2009, April 2009, August 2009, December 2009, January 2012, September 2012, January 2013, and July 2013
2	August 2006, March 2007, May 2007, June 2007, October 2007, June 2008, October 2008, December 2008, November 2010, December 2010, April 2011, November 2011, March 2012, June 2012, August 2012, October 2012, March 2013, October 2013, and December 2013
3	December 2006, January 2007, March 2008, April 2008, August 2008, September 2009, November 2009, January 2010, June 2010, September 2010, January 2011, March 2011, June 2011, August 2011, September 2011, December 2011, April 2012, December 2012, May 2013, September 2013, and January 2014

Table 1b The distribution of FOMC statements (Yellen: 2014–2018)

Cluster	*Statements*
0	June 2014, July 2015, October 2015, July 2016, September 2016, July 2017, September 2017, and November 2017
1	June 2015, January 2016, March 2016, April 2016, June 2017, and December 2017
2	March 2014, July 2014, September 2014, January 2015, November 2016, December 2016, February 2017, and March 2017
3	April 2014, October 2014, December 2014, March 2015, April 2015, September 2015, December 2015, June 2016, May 2017, and January 2018

Table 1c The distribution of FOMC statements (Powell: 2018–2024)

Cluster	*Statements*
0	March 2018, June 2018, September 2018, June 2019, July 2019, September 2019, and September 2020
1	Aug 2018, March 2020, April 2020, June 2020, July 2020, March 2021, April 2021, June 2021, July 2021, September 2021, March 2022, May 2022, July 2022, March 2023, January 2024, March 2024, and May 2024
2	May 2018, May 2019, November 2020, June 2022, September 2022, November 2022, December 2022, May 2023, June 2023, July 2023, September 2023, November 2023, and December 2023
3	November 2018, December 2018, January 2019, March 2019, Oct 2019, December 2019, January 2020, December 2020, January 2021, November 2021, December 2021, January 2022, and February 2023

To summarise the FOMC statements for each chairperson, Tables 1a, 1b and 1c show the distribution of the FOMC statements into four clusters.

4.3 Dominant topics

The third step in the analysis was to investigate each cluster to identify the dominant topics in each cluster. To do so, we categorised the terms in each cluster under the six topics that we had previously identified. Table 2 shows the six topics and the most commonly occurring terms for each topic. The selected topics served as proxies to achieve the twin goals of this paper, transparency and priorities. The first topic was monetary policy, which reveals the amount of information the Fed discloses on the directions of future monetary policy. The other five topics highlighted the main issues the Fed is concerned with in setting its policy goals. For example, the topics of inflation and labour market are related to the Fed's dual mandate of price stability and maximum employment. Macroeconomy is yet another important issue that corresponds to economic performance and achieving sustainable economic growth. The financial market topic relates to preserving the stability of the financial market and avoiding financial crises that may destabilise the economy. Finally, we were interested in knowing whether the Fed considers the state of the global economy in setting its policy.

The word cloud shown in Figure 10 presents the most frequent words in each of the general topics discussed in Table 2.

Figure 10 Word Cloud of the general topics and the most common terms in each topic, (a) monetary policy (b) inflation (c) labour market (d) macroeconomy (e) financial market (f) global economy (see online version for colours)

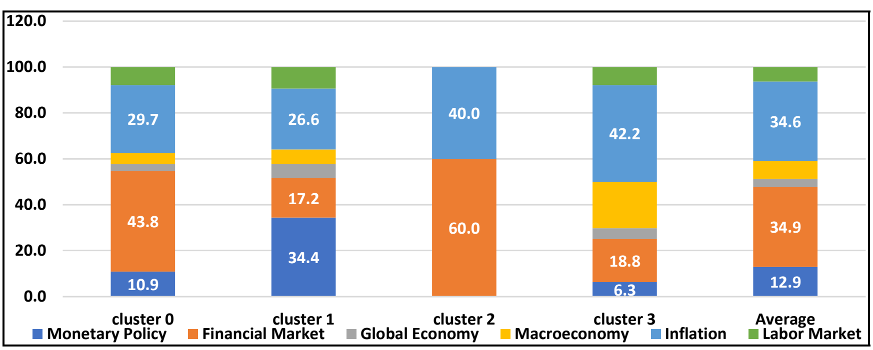
Figures 11, 12, and 13 show the breakdown of each cluster into six topics for every term of the three chairmanships. In addition, and to make it easier for comparison, we added a column showing the average share of each topic across the four clusters.

During the Bernanke term, the financial market was the dominant issue in two clusters, and, on average, was the most dominant issue, accounting for 34.9% of clusters' content. Notably, the financial market was the dominant issue only during the Bernanke term. This is partially because of the great turbulence in the financial market during the GFC, which took place during his term. The second most important issue was inflation, which dominated cluster 3, and was, on average, the second most dominant issue, accounting for 34.6%. This was also expected given the importance of price stability as one component of the Fed's dual mandate and because the Fed started implementing inflation-targeting policy in 2012.

Table 2 The general topics and the most common terms in each topic

Topic	*Terms*
Monetary policy	Federal funds, funds rate, policy rate, FOMC participants, monetary policy, policy target, dual mandate, target range, and open market
Inflation	Core inflation, price inflation, energy price, PCE inflation, inflation target, core, PCE, and oil price
Labour market	Civilian employment, civilian unemployment, labour market, wage rate, jobless claims, non-farm payroll, labour force, hourly earnings, and maximum employment
Macroeconomy	Economic growth, GDP growth, industrial production, personal consumption, household spending, business investment, fixed investment, government purchases, economic outlook, government spending, and manufacturing output
Financial market	Bond yield, financial system, banking system, financial markets, asset price, stock market, commercial banks, bank bankruptcy, credit conditions, default rate, equity price, lending activities, market liquidity, bank reserves, financial stability, treasury bills, market volatility, financial condition, treasury securities, and interest rate
Global economy	Foreign currencies, foreign economies, global economy, global markets, European union, emerging economies, geopolitical, international markets, and international trade

Figure 11 FOMC statements: topics per cluster (Bernanke: 2006–2014) (see online version for colours)



The third issue was monetary policy, which dominated cluster 1 and accounted, on average, for 12.9% of the clusters' content. This shows how open the Fed has been in recent years in talking about their stance and the direction of monetary policy.

The ranking of issues during the Yellen and Powell terms differs from Bernanke's with respect to issues dominating the clusters. Both Yellen and Powell have similar rankings for the dominant issues, monetary policy, inflation, and the financial market. This shows, first, that the financial market moved to the third position in importance after the financial crisis. Second, monetary policy jumped to the first position, which may signal that the Fed was talking more about monetary policy indicators such as the federal funds rate. Last, inflation affirmed its importance by maintaining its position as the second dominant issue during both Yellen and Powell's terms, accounting for 30.1% and 26.8%, respectively. This shows that the Fed has been continuing its inflation targeting policy in recent years.

Figure 12 FOMC statements: topics per clusters (Yellen: 2014–2018) (see online version for colours)

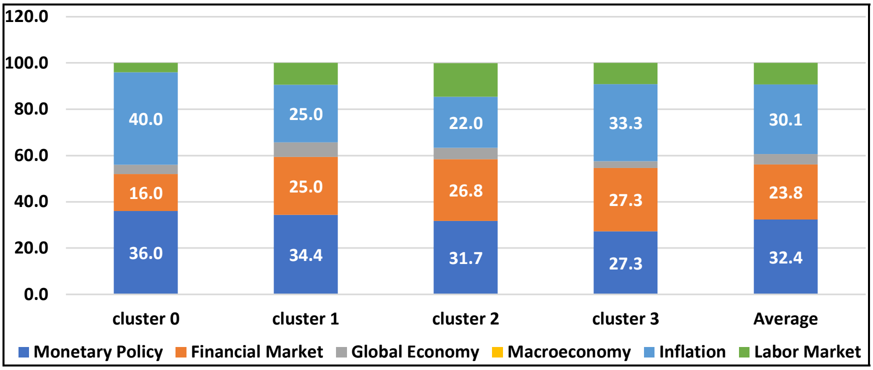
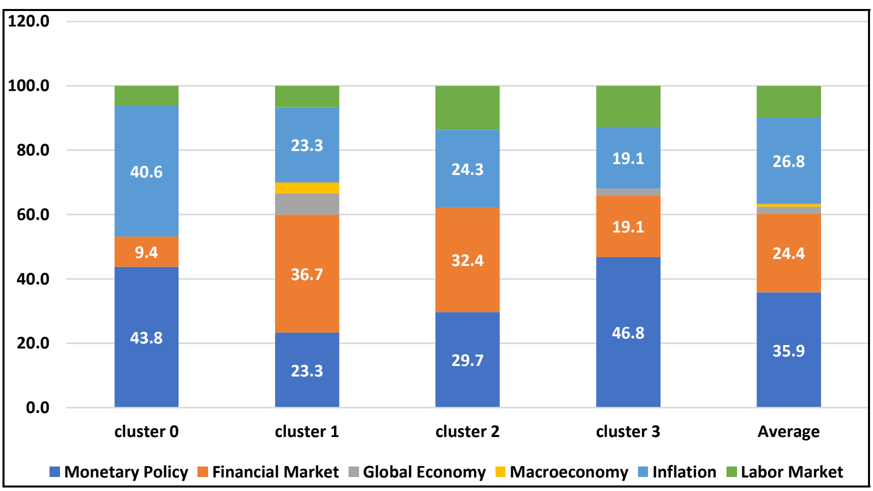


Figure 13 FOMC statements: topics per cluster (Powell: 2018–2024) (see online version for colours)



4.4 *Fed priorities*

The next step in the analysis was to rank the five priorities that the Fed may be concerned with while implementing monetary policy: inflation, macroeconomy, labour market, financial market, and global economy. The ranking was based on the average weights of these priorities in the clusters, which were calculated in Figures 11, 12, and 13, under the three chairmanships. Table 3 shows the Fed priorities during the terms of Bernanke, Yellen, and Powell.

Table 3 Fed priorities during the terms of Bernanke, Yellen, and Powell

Rank	*Bernanke*	*Yellen*	*Powell*
1	Financial market	Inflation	Inflation
2	Inflation	Financial market	Financial market
3	Macroeconomy	Labour market	Labour market
4	Labour market	Global economy	Global economy
5	Global economy	Macroeconomy	Macroeconomy

Bernanke had a totally different ranking of priorities than Yellen and Powell, and the latter two had identical rankings. On the one hand, we see that during Bernanke's term, financial market stability ranked as the top priority, heavily influenced by the turbulence in the market during the GFC. On the other hand, the financial market moved to the second rank during the Yellen and Powell terms, suggesting that the Fed continued to monitor financial market stability even after the financial crisis.

Inflation was ranked second during Bernanke's term, reflecting the Fed's adherence to its mandate of price stability and implementation of inflation-targeting policy since 2009. Chairpersons Yellen and Powell continued to closely monitor inflation, which moved it to the first rank. This is consistent with other major central banks in the world that implement inflation targeting in their policies by setting a 2% inflation rate target.

However, the macroeconomy (or economic growth) received varying degrees of attention from the three chairpersons. Again, Bernanke gave it more attention, as it was ranked third, and Yellen and Powell gave it the least attention; that is, it ranked last. The great recession during Bernanke's term played a role in keeping the stability of the macroeconomy within the Fed's priorities. During Yellen and Powell's terms, the macroeconomy had a decent and stable performance, which explains why macroeconomy was ranked last.

The labor market received some attention from the Fed. As reflected in the FOMC statements, it was ranked third in Yellen and Powell's terms and fourth in Bernanke's. Employment is one of the Fed's dual mandates and is one of the variables closely monitored by the Fed in setting the direction of monetary policy.

The priority that received the least attention from the Fed during the Bernanke term was the global economy, ranked fifth; it ranked fourth during the Yellen and Powell terms. Because the USA is the largest global economy and given the dominant role of the US dollar in international trade and investment, the Fed seems to pay less attention to the condition of the global economy in setting its policy interest rates. In fact, other central banks usually follow the Fed in setting the direction of their monetary policy.

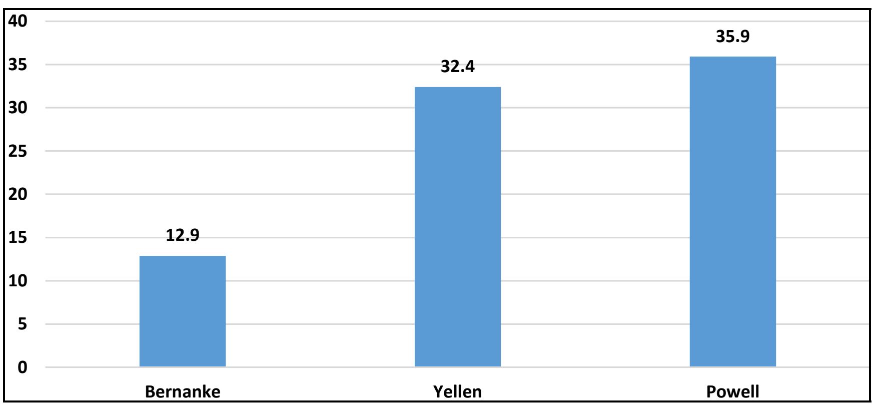
Overall, the analysis of the Fed's economic priorities reveals that the Bernanke term stands in contrast with the other two terms in the ranking of priorities in setting US monetary policy. This may indicate that Bernanke's priorities were different, so we can conclude that the Fed chairperson may have some influence on the Fed's priorities. The Yellen and Powell terms had identical ranking of priorities but were different from Bernanke's. This evidence is supported by our analysis of the FOMC statements using the text mining technique to extract the most important terms mentioned in these statements.

4.5 Fed transparency

The final step in the analysis is to rank the terms of the three chairpersons with respect to transparency. We measured the level of transparency by the relative weight of monetary policy in the clusters among the three chairpersons. A higher weight for monetary policy reflects the level of transparency during each chairmanship, as measured by how much information the FOMC statements provided regarding monetary policy.

Figure 14 shows the relative weight under the three chairmanships. It is clear that there is a significant difference in the level of transparency among the three chairpersons. Bernanke's term had the lowest level of transparency with a relative weight of 12.9%. The level of transparency was significantly higher during the Yellen term at 32.4% and even higher during Powell's term to reach 35.9%. This result is consistent with the fact that the Fed has been more open in recent years in talking about the future directions of monetary policy, especially after implementing the forward guidance since the GFC.

Figure 14 Transparency of monetary policy (see online version for colours)



In general, the priorities and transparency trends derived from FOMC statements offer significant insights for multiple stakeholders. For example, policymakers, being considerate of the changes in the Fed's focus, can aid in coordination efforts and assist in predicting its actions in varying economic contexts. While for financial analysts, these insights enhance forecasting models and investment strategies by underscoring the economic indicators that the Fed is currently prioritising. As for the public, heightened transparency in the way the Fed communicates its future policy direction can alleviate uncertainty and enable consumers and firms to make more informed financial decisions.

The findings of the study indicated that inflation and financial market stability consistently ranked as primary priorities, while the level of the Fed's transparency has significantly increased from Bernanke to Powell, reflecting a shift towards more open communication regarding future policy directions. These insights enhance predictability and mitigate market volatility, ultimately contributing to a heightened stability of the economic environment.

5 Conclusions

In this study, we investigated the economic priorities of the Federal Reserve from March 2006 to May 2024, during which there were three different chairpersons. We aimed at investigating the influence of the chairperson on the course of monetary policy over time and what issues were considered when establishing new monetary targets under each chairperson, as reflected in the statements of FOMC meetings. We used textual data mining tools to extract terms that corresponded to a number of economic issues that the Fed considers in setting monetary policy during the FOMC meetings. In doing so, we tracked the frequency of mentioning certain terms relating to a number of economic issues throughout the 146 FOMC statements.

The results show that the Fed chairman Bernanke had completely different economic priorities than the other two. Yellen and Powell had identical ranking of priorities (as revealed in the FOMC statements). Overall, inflation and the financial market received the highest priorities under the three chairmanships.

We acknowledge the importance of the GFC and recommend this for future research by categorising the statements into three different categories to investigate the main topics or issues before, during, and after the GFC. In addition, more common economic and financial words could be investigated, and the same methodology can be applied to other central banks worldwide.

While the advancement in text mining tools (i.e., software) has helped researchers better extract meaningful information from central banks' statements, it still has several limitations. First, the produced output from the software depends on the program code used by the researcher, which may produce different results for the same statement. Second, published statements are drafted in a more general way, which may not reflect actual discussions in the meetings that could guide future policy direction. Third, text mining, by itself, cannot precisely conclude whether the statements' language is hawkish or dovish. Fourth, contextual meaning and the evolution of language may not be adequately captured by automated processes. Finally, variations in linguistic style over time or among different chairpersons can influence the consistency of interpretations. Therefore, researchers mostly would rely on their own judgment in interpreting the intended messages included in the statements, making the process more subjective than envisioned.

In this research we faced these challenges as well. For example, we had to use our own judgment and economic intuition in determining the appropriate number of clusters and number of words per phrase that best capture our objective of determining the Fed priorities and level of transparency.

Declarations

All authors declare that they have no conflicts of interest.

References

Al-Hassan, A., Alshameri, F. and Sibley, E. (2013) 'A research case study: difficulties and recommendations when using a textual data mining tool', *Information & Management*, Vol. 50, No. 7, pp.540–552.

Alshameri, F. and Green, N. (2020) 'Analyzing the strength between mission and vision statements and industry via machine earning', *The Journal of Applied Business Research*, Vol. 36, No. 3, pp.121–128.

Alshameri, F., Greene, G. and Srivastava, M. (2012) 'Categorizing top fortune company mission and vision statements via text mining', *International Journal of Management & Information Systems*, Vol. 16, No. 3, pp.227–238.

Andrade, P. and Ferroni, F. (2021) 'Delphic and Odyssean monetary policy shocks: evidence from the euro area', *Journal of Monetary Economics*, Vol. 117, pp.816–832.

Arismendi-Zambrano, J., Kypraios, E. and Paccagnini, A. (2024) 'Informativeness of the federal reserve chair communication's sentiment on the monetary policy uncertainty', *Annals of Operations Research*, pp.1–45.

Bernanke, B., Laubach, T., Mishkin, F. and Posen, A. (1999) *Inflation Targeting: Lessons from the International Experience*, Princeton University Press, New Jersey.

Bernanke, B.S. and Kuttner, K.N. (2005) 'What explains the stock market's reaction to Federal Reserve policy?', *The Journal of Finance*, Vol. 60, No. 3, pp.1221–1257.

Bholat, D., Hansen, S., Santos, P., and Schonhardt-Bailey, C. (2015) *Text Mining for Central Banks: Handbook*, Vol. 33, pp.1–19, Centre for Central Banking Studies, Bank of England, London.

Blei, D.M., Ng, A.Y. and Jordan, M.I. (2003) 'Latent Dirichlet allocation', *Journal of Machine Learning Research*, January, Vol. 3, pp.993–1022.

Blinder, A., Goodhart, C., Hildebrand, P., Lipton, D. and Wyplosz, C. (2001) *How do Central Banks Talk?*, Geneva Reports on the World Economy, International Center for Monetary and Banking Studies (ICMB), Switzerland.

Blinder, A.S., Ehrmann, M., Fratzscher, M., De Haan, J. and Jansen, D.J. (2008) 'Central bank communication and monetary policy: a survey of theory and evidence', *Journal of Economic Literature*, Vol. 46, No. 4, pp.910–945.

Boukus, E. and Rosenberg, J.V. (2006) *The Information Content of FOMC Minutes*, SSRN: 922312.

Campbell, J.R., Evans, C.L., Fisher, J.D., Justiniano, A., Calomiris, C.W. and Woodford, M. (2012) 'Macroeconomic effects of Federal Reserve forward guidance [with comments and discussion]', *Brookings Papers on Economic Activity*, pp.1–80.

Cannon, S. (2015) 'Sentiment of the FOMC: unscripted', *Economic Review*, Vol. 100, No. 4, Federal Reserve Bank of Kansas City.

Chahadah, A.A., Refae, G.A.E. and Qasim, A. (2018) 'The use of data mining techniques in accounting and finance as a corporate strategic tool: an empirical investigation on banks operating in emerging economies', *International Journal of Economics and Business Research*, Vol. 15, No. 4, pp.442–452.

Chau, F., Deesomsak, R. and Shaikh, R. (2025) 'Does Fed communication affect uncertainty and risk aversion?', *Review of Quantitative Finance and Accounting*, Vol. 64, No. 2, pp.713–756.

Cheng, X., Huang, D., Chen, J., Meng, X. and Li, C. (2019) 'An investigation on factors affecting stock valuation using text mining for automated trading', *Sustainability*, Vol. 11, No. 7, p.1938.

Cicekyurt, E. and Bakal, G. (2025) 'Enhancing sentiment analysis in stock market tweets through BERT-based knowledge transfer', *Computational Economics*, pp.1–23.

Cook, T. and Hahn, T. (1988) 'The information content of discount rate announcements and their effect on market interest rates', *Journal of Money, Credit and Banking*, Vol. 20, No. 2, pp.167–180.

Del Negro, M., Giannoni, M.P. and Patterson, C. (2023) 'The forward guidance puzzle', *Journal of Political Economy Macroeconomics*, Vol. 1, No. 1, pp.43–79.

Devlin, J. (2018) *BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding*, arXiv preprint arXiv:1810.04805.

Dotsey, M. (1987) 'Monetary policy, secrecy, and federal funds rate behavior', *Journal of Monetary Economics*, Vol. 20, No. 3, pp.463–474.

Edison, H. and Carcel, H. (2021) 'Text data analysis using latent Dirichlet allocation: an application to FOMC transcripts', *Applied Economics Letters*, Vol. 28, No. 1, pp.38–42.

Ernst, E. and Merola, R. (2018) *Central Bank Communication: A Quantitative Assessment*, International Labour Office Research Department Working Paper, No. 33.

Evans, M.D. and Lyons, R.K. (2005) 'Do currency markets absorb news quickly?', *Journal of International Money and Finance*, Vol. 24, No. 2, pp.197–217.

Farka, M. and Fleissig, A.R. (2013) 'The impact of FOMC statements on the volatility of asset prices', *Applied Economics*, Vol. 45, No. 10, pp.1287–1301.

Gidofalvi, G. and Elkan, C. (2001) *Using News Articles to Predict Stock Price Movements*, No. 17, Department of Computer Science and Engineering, University of California, San Diego.

Gu, C., Chen, D., Stan, R. and Shen, A. (2022) 'It is not just what you say, but how you say it: why tonality matters in central bank communication', *Journal of Empirical Finance*, Vol. 68, pp.216–231.

Gupta, A., Dengre, V., Kheruwala, H.A. and Shah, M. (2020) 'Comprehensive review of text-mining applications in finance', *Financial Innovation*, Vol. 6, pp.1–25.

Gürkaynak, R.S., Sack, B.P. and Swanson, E.T. (2005) 'Do actions speak louder than words? The response of asset prices to monetary policy actions and statements', *International Journal of Central Banking*, Vol. 1, No. 1, pp.55–93.

Hendry, S. and Madeley, A. (2010) *Text Mining and the Information Content of Bank of Canada Communications*, https://ssrn.com/abstract=1722829.

Huang, Y.L. and Kuan, C.M. (2021) 'Economic prediction with the FOMC minutes: an application of text mining', *International Review of Economics & Finance*, Vol. 71, pp.751–761.

Hubert, P. and Fabien, L. (2017) *Central Bank Sentiment and Policy Expectations*, Bank of England Staff Working Paper, No. 648.

Hutto, C. and Gilbert, E. (2014) 'Vader: a parsimonious rule-based model for sentiment analysis of social media text', in *Proceedings of the International AAAI Conference on Web and Social Media*, May, Vol. 8, No. 1, pp.216–225.

Jegadeesh, N. and Wu, D. (2017) *Deciphering Fedspeak: The Information Content of FOMC Meetings*, SSRN: 2939937.

Kahveci, E. and Odabaş, A. (2016) 'Central banks' communication strategy and content analysis of monetary policy statements: the case of Fed, ECB and CBRT', *Procedia-Social and Behavioral Sciences*, Vol. 235, pp.618–629.

Kim, W., Spörer, J.F. and Handschuh, S. (2023) *Analyzing FOMC Minutes: Accuracy and Constraints of Language Models*, arXiv preprint arXiv:2304.10164.

Kryvtsov, O. and Petersen, L. (2021) 'Central bank communication that works: lessons from lab experiments', *Journal of Monetary Economics*, Vol. 117, pp.760–780.

Landauer, T., Foltz, P. and Laham, D. (1998) 'An introduction to latent semantic analysis', *Discourse Processes*, Vol. 25, Nos. 2&3, pp.259–284.

Lucca, D.O. and Moench, E. (2015) 'The pre-FOMC announcement drift', *The Journal of Finance*, Vol. 70, No. 1, pp.329–371.

Lucca, D.O. and Trebbi, F. (2009) *Measuring Central Bank Communication: An Automated Approach with Application to FOMC Statements*, No. w15367, National Bureau of Economic Research.

Mazis, P. and Tsekrekos, A. (2017) 'Latent semantic analysis of the FOMC statements', *Review of Accounting and Finance*, Vol. 16, No. 2, pp.179–217.

Moessner, R. (2013) 'Effects of explicit FOMC policy rate guidance on interest rate expectations', *Economics Letters*, Vol. 121, No. 2, pp.170–173.

Morris, J.X., Kuleshov, V., Shmatikov, V. and Rush, A.M. (2023) *Text Embeddings Reveal (Almost) as Much as Text*, arXiv preprint arXiv:2310.06816.

Mueller, P., Tahbaz-Salehi, A. and Vedolin, A. (2017) 'Exchange rates and monetary policy uncertainty', *The Journal of Finance*, Vol. 72, No. 3, pp.1213–1252.

Muennighoff, N., Tazi, N., Magne, L. and Reimers, N. (2022) *MTEB: Massive Text Embedding Benchmark*, arXiv preprint arXiv:2210.07316.

Osowska, E. and Wójcik, P. (2024) 'Predicting the reaction of financial markets to Federal Open Market Committee post-meeting statements', *Digital Finance*, Vol. 6, No. 1, pp.145–175.

Rafferty, M. and Tomljanovich, M. (2002) 'Central bank transparency and market efficiency: an econometric analysis', *Journal of Economics and Finance*, Vol. 26, No. 2, pp.150–161.

Rohlfs, C., Chakraborty, S. and Subramanian, L. (2016) 'The effects of the content of FOMC communications on US's treasury rates', in *Proceedings of the 2016 Conference on Empirical Methods in Natural Language Processing*, No. 2096–2102.

Romer, C.D. and Romer, D.H. (1989) 'Does monetary policy matter? A new test in the spirit of Friedman and Schwartz', *NBER Macroeconomics Annual*, Vol. 4, pp.121–170.

Rosa, C. (2011) 'Words that shake traders: the stock market's reaction to central bank communication in real time', *Journal of Empirical Finance*, Vol. 18, No. 5, pp.915–934.

Schumaker, R.P., Zhang, Y., Huang, C.N. and Chen, H. (2012) 'Evaluating sentiment in financial news articles', *Decision Support Systems*, Vol. 53, No. 3, pp.458–464.

Shapiro, A.H. and Wilson, D. (2019) *Taking the Fed at its Word: Direct Estimation of Central Bank Objectives using Text Analytics*, Federal Reserve Bank of San Francisco, February.

Siklos, P. (2020) 'U.S. monetary policy since the 1950s and the changing content of FOMC minutes', *Southern Economic Journal*, Vol. 86, No. 3, pp.1192–1213.

Siklos, P., Amand, S.S. and Wajda, J. (2018) *The Evolving Scope and Content of Central Bank Speeches*, CIGI Papers, No. 202.

Swanson, E.T. (2021) 'Measuring the effects of Federal Reserve forward guidance and asset purchases on financial markets', *Journal of Monetary Economics*, Vol. 118, pp.32–53.

Swanson, E.T. and Williams, J.C. (2014) 'Measuring the effect of the zero lower bound on medium- and longer-term interest rates', *American Economic Review*, Vol. 104, No. 10, pp.3154–3185.

Tadle, R.C. (2022) 'FOMC minutes sentiments and their impact on financial markets', *Journal of Economics and Business*, Vol. 118, p.106021.

Tumala, M. and Omotosho, B.S. (2019) 'A text mining analysis of central bank monetary policy communication in Nigeria', *CBN Journal of Applied Statistics*, Vol. 10, No. 2, pp.73–107.

Weber, C.S. (2019) 'The effect of central bank transparency on exchange rate volatility', *Journal of International Money and Finance*, Vol. 95, pp.165–181.