Economic Signals in Deliberations of the Federal Open Market Committee: A Topic Model Approach Data Science in Action

Christian Langerfeld

Department of professional and intercultural communication Norwegian School of Economics Bergen, Norway

21 November, 2024

NHH



Bio: Christian Langerfeld

- Associate professor in *Digital text analysis* at the Norwegian School of Economics
- Background in corpus linguistics
- Collaborations with researchers from Finance and Business Management
- Research interests: analyzing financial narratives for various purposes

Plan for the lecture

- 1. Introduction: Textual analysis
- 2. The Federal Open Market Committee
- 3. The data some preprocessing
- 4. The set-up of the analysis
- 5. Linguistic signals of the pending financial crisis
- 6. Topic models introduction
- 7. Topic models estimating a model
- 8. Results

Textual Analysis and financial narratives

- Textual analysis involves applying natural language processing (NLP) and other computational techniques to extract meaningful insights from text data.
- Financial narratives, such as annual reports, earnings calls, and analyst notes, are rich in qualitative information that can influence investor behavior.
- These narratives often contain cues about sentiment, risk, future performance, and strategic priorities.

Examples:

- Sentiment analysis
- Document classification
- Keyword extraction
- Text summarization
- Topic models
- etc.

Textual analysis in finance



BAN432: Applied Textual Data Analysis for Business and Finance

- Course at the master level in the Business Analytics program at NHH
- Collaboration with Max Rohrer from the Department of Finance/NHH → interdisciplinary, linguistics and finance
- Textual analysis usually involves 3 steps:
 - 1. Data collection: APIs, web scraping
 - 2. Pre-processing: Giving structure to unstructured textual data and preparing it for the analysis
 - 3. Analysis: Gaining insights through e.g., sentiment analysis, document clustering, keyword extraction, topic models
- The topic of today's talk is something our students have worked on some years ago.

About the FOMC

The Federal Open Market Committee (FOMC)

The FOMC is a branch of the Federal Reserve System responsible for setting U.S. monetary policy.

It was established to achieve key economic goals, namely:

- Maximum employment
- Stable prices (inflation control)
- Moderate long-term interest rates
- The FOMC holds eight scheduled meetings annually, though it may meet more frequently if necessary.
- After each meeting a transcript is prepared.

Structure of the FOMC

- The FOMC comprises twelve voting members:
 - Seven members of the Federal Reserve Board of Governors
 - The president of the Federal Reserve Bank of New York
 - Four of the remaining eleven Reserve Bank presidents, who serve on a rotating basis
- The chair of the FOMC is also the chair of the Federal Reserve Board.

Economic impact of FOMC decisions

Market reactions: Financial markets respond to FOMC decisions and policy statements, often anticipating economic trends. Rate hikes can lead to stock market declines, while rate cuts generally support market growth.

Impact on consumers and businesses:

- For consumers:
 - Higher interest rates mean more expensive loans and credit (e.g., mortgages, auto loans), discouraging spending.
 - Lower rates reduce borrowing costs, encouraging spending and investment.
- For businesses:
 - Lower rates reduce the cost of financing expansions, boosting hiring and investments.
 - Higher rates may lead to reduced business growth and expansion.



Figure 1: FOMC meeting in 2016. Souce: Wikipedia

Data: The FOMC transcripts

- FOMC transcripts are detailed, word-for-word records of the discussions and deliberations during Federal Open Market Committee meetings.
- They provide insights into the economic analyses, policy debates, and individual viewpoints of committee members
- They reveal the rationale behind interest rate decisions and other monetary policies.
- These transcripts are released with a five-year delay to protect sensitive information and maintain market stability.

Jump to the FOMC webpage.

Workflow:

- 1. Define a time range, e.g. c(2005:2008)
- 2. Read the page for each year line-by-line or parse it.
- Find and extract the links to the pdf documents (transcripts); they end with meeting.pdf.
- 4. Download the pdf documents using the links extracted in 3.

Once the transcripts are downloaded they have to be pre-processed. Workflow:

- 1. Convert the pdf to text
- 2. Identify where the actual transcript starts.
- 3. Remove page headers, page numbers, etc.
- 4. Generate an XML file.

From pdf to xml

January 28-29, 2003

3 of 195

Transcript of the Federal Open Market Committee Meeting on January 28-29, 2003

January 28—Afternoon Session

CHAIRMAN GREENSPAN. Good afternoon, everybody. I have a welcome for Sandy

Pianalto. Since she will not officially be a Reserve Bank President for a number of days, we will

put your welcome in escrow, and you may withdraw it at the appropriate time.

MS. PIANALTO. Thank you.

<doc meeting_date="2003-01-29" chair="Greenspan">CHAIRMAN
GREENSPAN. <participant name="CHAIRMAN GREENSPAN.">Good
afternoon, everybody. I have a welcome for Sandy Pianalto. Since
she will not officially be a Reserve Bank President for a number
of days, we will put your welcome in escrow, and you may
withdraw it at the appropriate time. </participant>MS. PIANALTO.
<participant name="MS. PIANALTO.">Thank you. </participant>

The research question

Did the FED anticipate the financial crisis of 2007?

The set-up

When did the financial crisis start?

Define:

- a crisis corpus
- a pre-crisis corpus
- a control corpus (when no one possibly knew about the upcoming crisis)

The financial crisis



Figure 2: The TED spread. Source: Wikipedia

Pre-crisis, crisis, and control corpus

- ▶ We set the start of the financial crisis to 2007-08-01.
- Crisis corpus: 9 meetings after that date
- Pre-crisis corpus: 9 meetings before first "crisis meeting"
- Control corpus: 9 meetings before pre-crisis corpus starts



Defining the FOMC meeting sample splits

Examining the data – negative words

- Complile a word list of words with a negative tone
- Count how many negative words occure in each transcript
- Scale the count by the total number of words in a document

Sentiment in the three samples



Examing the data – Keywords

- Extract the 50 keywords that are strongest associated with the financial crisis
- Use bi-grams
- ► For each bi-gram compute: relative freq crisis sample relative frequency control sample
- Limit to 50 with the highest score
- Count how many times they occure in the 3 samples

Keywords associated with the crisis corpus



Crisis keywords - trend

Trend analysis - relative frequency of identified keywords



Topic modeling

Topic modeling – introduction

Text 1:

Waldhof Mannheim, officially SV Waldhof 07 Mannheim, is a historic German football club based in Mannheim. Baden-Württemberg. Founded in 1907, the club is known for its passionate fan base and long-standing traditions. Waldhof enjoyed its peak during the 1980s, competing in the **Bundesliga** and earning respect as a tough opponent. In recent years, the team has primarily competed in lower leagues, with their home games played at the Carl-Benz-Stadion. Despite challenges, Waldhof Mannheim remains a symbol of local pride, combining a rich history with ambitions for future success.

Text 2

The 2024 U.S. elections were marked by intense competition, with former President Donald Trump representing the Republican Party and Vice President Kamala Harris standing as his Democratic opponent. Trump's campaign focused on economic revitalization and immigration reforms, while Harris emphasized social justice and climate policies. The elections also saw record voter turnout, reflecting the polarized political landscape. Ultimately, the Republican candidate prevailed, with significant victories in key swing states.

Typical words for the topic *football*:

football, club, league, Waldhof Mannheim, Carl-Benz-Stadion, Bundesliga Typical words for the topic *US*-elections:

elections, Donald Trump, Republican, Kamala Harris, Democratic, voter, swing states, etc.

Basic assumption in topic modeling:

- All topics occur in all documents of a corpus.
- There are different probabilities for each topic to occur in a given document.
- All words occur in all topics
- There are different probabilities for each word to occur in a given topic.

• Estimating a topic model aims at answering two questions:

- 1. What is the probability of a topic to occur in a given document?
- 2. What is the probability of a word to belong to a given topic?

Title	Bundesliga	football	and	in	elections	voter
"History of Waldhof Mannheim"	1	1	2	5	0	0
"US elections: The day after"	0	0	3	2	2	1
"Scholz attends Bundesliga match"	3	2	5	3	2	1

Distribution of 6 terms in 3 texts

Title	Topic Football	Topic Elections	\sum
"History of Waldhof Mannheim"	1	0	1
"US elections: The day after"	0	1	1
"Scholz attends Bundesliga match"	0.5	0.5	1

1. What is the probability of a topic to occur in a given document?

Title	Bundesliga	football	and	in	elections	voter	\sum
Football Topic	0.45	0.45	0.05	0.04	0	0.01	1
Election Topic	0.01	0	0.03	0.06	0.6	0.3	1

2. What is the probability of a word to belong to a given topic?

Topic modeling – Analysing the data

Terms in Topic 6

uncertainty paper financial commercial productivity potential growth predominant developments r confidence corporate assessment e secono -mortgage marke demand **Credit** subprime investment increased adjustment statement moderate underlying implications

Topic 6 – trend

Posterior values for topic 6



Topic 7

easing exhibit dollar balance spreads points basis large capital realrisk market economic banksmarkets bank assets o many oar institutions committee commercial lending lending

Topic 7 – trend

Posterior values for topic 7



Results

RQ: Did the FED anticipate the financial crisis of 2007?

- Some evidence that the FED started to talk about issues regarding the financial crisis before the market was aware of it.
- The frequency of negative words went up in the pre-crisis sample
- The frequency of keywords related to the crisis went also up in the pre-crisis sample
- Using topic modeling, we identified topics that are related to the crises; they became more prominent in the pre-crisis sample

You find the data and the R-code on OSF.

OSF.IO/UV4XY

christian.langerfeld@nhh.no