Economics Terminologies in FOMC Minutes
CSCI699 Introduction to Information Extraction: project final report

Akira Matsui
Student ID:5421284665
amatsui@usc.edu

Abstract

Professional knowledge is widely used to solve complex problems in our society. In social science, the vast amount of labor has been devoted to public policy issues. Many policymakers use their professional knowledge in social sciences to solve social problems. However, it is unknown whether these professionals use their knowledge explicitly in policy making situations. Here, we show that professional knowledge plays a pivotal role when policymakers explain their policies to the public. We find that the important terminology semantic change when central banks change their behavior. We select the technical terminology in the Federal Reserve policy goal and compute the word embedding of these terminologies in the Federal Open Market Committee minutes. The changes in the embedding vectors affect the federal funds rate and monetary base. Our results demonstrate that the Federal Reserve Bank (FRB) places considerable value on the concepts in its monetary policy goal and the technical terminologies convey good indicators of the FRB's behavior. We anticipate this study to be a starting point to explore how we should use professional knowledge when explaining our findings to the public.

1 Introduction

Many research findings are used to solve real-world problems. Economics is a typical example. Tons of economics research is devoted to solving a wide range of economic problems, from stable marriage problems to optimal monetary policy problems. Even though professionals in economics use their economics knowledge in their minds when they tackle problems in the economy, do they use this knowledge explicitly in discussions? If so how these are used?

We hypothesize that policymakers utilize concepts from economics when they explain difficult decisions such as policy rule changes. In other words, economics terminology plays a pivotal role to explain their policy. To study this hypothesis, we use the data from Federal Reserve Bank (FRB), which is the central bank in the United States.

We utilize the dataset made from Federal Open Market Committee (FOMC) minutes, and an information extraction technique, to extract relationships between economics terminologies and FRB behavior. We focus on the data from FRB data for the following reason. First, we can expect that FOMC members use principal economics concepts because all of the FOMC members are economic experts both from academia and practical business. Second, the FRB is required to keep their transparency by the federal law, which means that the FRB always have responsibility to explain their policy to the public. The minutes of the committee are open to the public and they have specific goals of monetary policy that enable us to understand what they value. Lastly, we can clearly observe the the FRB’s behavior by studying their policy tools to stimulate the economy such as Federal Reserve Banks.

After FOMC meeting, the monetary policy decision becomes open to the public. At this point, central banks need to announce their decision properly. This communication by the central banks is considered one of the important factors to implement a successful monetary policy [7, 32]. Therefore, a central bank deeply values their communication not only with the market participants but also with the general public. Beside, central banks have to remain transparent by reporting
how these decisions are being made because "better public understanding makes the policy more credible and effective." The importance of Transparency is well known in both theoretical and empirical economic studies [10].

Although the specific goals slightly vary among central banks, they aim at keeping economic conditions healthy and reducing uncertainty. The aim of monetary policy, more specifically a central bank’s goal is considered to promote maximum employment, stable prices, and economic growth. In recent years, monetary policy is usually a set of policy operations using several tools, such as purchasing government bonds, assets and setting reserve bank rate [6].

To test this hypothesis, we use FOMC minutes data from 1993Q1 to 2018Q1, and map the FRB mandate to the terminologies in economics dictionary to select the important terminologies. In addition, we use the historical data of Federal Funds Rate (FFR) and Monetary Base (MB) to see changes in the semantics of the terminologies reflect the policy changes. To extract the information of the semantic of the terminologies, we use Word2Vec to obtain the word embedding of the terminologies. Then, we study how the changes in the word embedding affect FRB policy changes.

2 Related Research

In this section, we survey the related research that uses the documents published by the central banks. Most works use NLP techniques to extract information from the minutes and study some important concepts in economics such as communication, uncertainty and transparency. Others are interested in extracting features from the minutes to forecast economic outcomes. As discussed in the Introduction, we are interested more generally in how professionals utilize the professional knowledge to explain their behavior in a policy making situation.

2.1 Communication

Most studies about central banks' communications use NLP to extract features from the documents by the central bank and find the relationships between these features and economic indicators such as interest rates or monetary policy goals.

Lucca and Trebbi (2009) use sentiment analysis with the FOMC minutes and show how the sentiments in the minutes affect Treasury rates. To find the effect of sentiment on Treasury rates, they use Vector Auto Regression (VAR) model, which is the mainstream econometric method in Macroeconomics. They find that communication has an impact on Treasury rates and also the minutes tell us the critical implication in monetary policy such as rule-based interest rate and the Taylor-rule. Tang (2017) studies whether the central bank communications affect interest rates. [30]. The study presents the methods with the Naive Bayes model and find a strong relationship the labor-related contents in the minutes and monetary policy response to labor news. Hendry and Madeley (2010) study more specific aspect of the relationship between a central bank communication and Treasury rates by using Latent Semantic Analysis (LSA) [12]. They study how the communication of the Bank of Canada (BOC) affects returns and volatility in short-term as well as long-term interest rate markets. They find that strong effects appear in interest rates.

There are studies on the communication differences in the central banks. Eyup and Odabas (2016) classify the communication strategies [13]. They classify the policy statements into several categories from the FRB, the European Central Bank (ECB) and Central Bank of the Republic of Turkey (CBRT). Keida and Takeda (2017) show the changes in monetary policy communication among the two different governors [17]. As an anecdotal evidence, Mr. Kuroda, the BOJ governor, uses the completely different style of speech and interviews than the former governor. Keida and Takeda (2017) use tf-idf and similarity evaluation using vector space model (SVM), and then compare the differences among two BOJ governors speeches.

Footnotes:

2To understand why uncertainty matters in monetary policy see Bekaert et al. (2013). [4]
3They argue that their semantic analysis extracts the changes in the monetary policy.
4To understand these rules, see 9.4.2 Monetary policy rules in Kuttner (2018) [15]
5In monetary policy, an attenuating volatility of a interest rate is considered as significant since high volatility means high uncertainty in future economic conditions. For example, see 9.4.1 Monetary policy implementation in Kuttner (2018) [15].
2.2 Uncertainty and Transparency

Some researchers use NLP to create uncertainty index. Andrés (2017) use LDA to compose economic policy uncertainty index from news text [3]. Saltzman and Yung (2018) construct uncertainty measurements from the FOMC documents covering the long period (from 1970 to 2018) [24]. They also use VAR to find the relation between their uncertainty measurements and economic condition.

Acosta (2015) use LSA to make a transparency measurement over 32 years period from FOMC minutes. By using their measurements, this study shows that the Act’s requirement increased the central bank’s transparency [1].

2.2.1 Attitude of committee members

Suda et al. (2018) extract disagreements among FOMC members and study its effect on asset price [28]. They classify the topics in FOMC members speeches and define disagreement as the dispersion of the quantified sentiments. The study shows that the disagreement among FOMC members dilutes the announcement of the future monetary policy direction.

Apel and Grimaldi (2012) use semantic analysis to study the attitudes of the committee members [2]. They use the minutes of the Swedish central bank and find the attitudes of the committee members are useful to predict monetary policy decisions.

2.3 Economic prediction

There is a lot of NLP studies that predict macroeconomic outcomes. Here, I focus on the studies related to monetary policy. Rohlfis et al. (2016) predict monetary policy targets and Effective Federal Funds Rate (EFFR) by extracting the topics in FOMC statements with Latent Dirichlet Allocation (LDA) [23].

Moniz and Jong (2014) predict interest rate expectations using machine learning. They use the Bank of England Monetary Policy Committee Minutes (MPCM) from 1997 to 2014 and ensemble methods to predict interest rate expectations in the financial market [20]. Ramachandran and DeRose (2018) classify FOMC meetings in 2017 [21]. Štajner et al. (2016) focus on the specific classification problem. They classify speculative and non-speculative speech in the transcripts of FOMC monetary policy meetings [25]. Miller and McCoy (2014) study the content changes in FOMC transcripts and detect the changes between the pre and the post-financial crisis [19]. They use LDA to extract topic from the transcripts and classify them. Tan and Lee (2018) study the effect of emphasis on the listeners’ reception in FOMC members’ speech [21]. To this aim, they examine rhetoric patterns, hedging, in the transcripts of all FOMC meetings from 1977 to 2008.

While the related researches use the concepts in a specific discipline or features such as sentiments, we are focusing on the minimum unit of language, words. Therefore, our study is more general than other related research. The implication of our study can be applied to outside of economics as well: professional terminologies convey important information to understand their behavior.

3 Data

In this section, we provide a brief overview of data collection and describe the data. We crawl the FOMC minutes from 1967 to 2018 and the economics terminology dictionary.

The FOMC minutes are the minutes of the FOMC meeting to announce the monetary policy decision to the public. As discussed in the Related Research section, the minutes by the central banks provide the important information about the central bank behavior. The FOMC minute plays the pivotal role in monetary policy. The FOMC minutes include not only the detail of policy decision making but also the perspectives and the outlook of the economic condition at that time.

To have the FOMC minute datasets, we use the existing datasets containing the minutes from 1967 to 2007. In addition, we write a crawler to obtain the minutes from 2008 to 2018. We also collect the economic terminologies from the website that gives brief explanations of basic economics concepts, The Economics Classroom. This website is run by the professor in economics.

---

Figure 1: Number of words in minutes in each quarter

After preprocessing the FOMC minute, we decided to focus on the minutes in 1993Q1-2018Q1 for two reasons. First, the minutes before 1993 are much shorter than those minutes after 1993. Second, some minutes before 1993 do not have the terminologies that we select in the Methods section. The Figure 1 shows the number of words in minutes over the entire periods. This plot describes the length of minute increases over time. The striking point is between 1992 and 1993 periods. The FRB changed the style of the minute and they started to publish digital version minutes in 1993. To prevent this fact from affecting our analysis, we study the minute from 1993Q1 to 2018Q1 in the rest of this paper.

To identify the relation between the economic terminology usages and the monetary policy changes, we use the historical data of the FFR and the MB. As discussed in the Related Research, FFR and MB are the important monetary policy operation tools and the FRB change them to conduct their policy packages. Therefore, we can assume that the changes in these policy tools can be interpreted as the changes in monetary policy. The Figure 2 shows log difference of FFR and MB. The most clear change of the monetary policy was made after the financial crisis in 2008.

4 Methods

In this section, we will explain the methods to identify the relationships between changes in the terminology usage by the central bank and in the central bank behavior. First, we compute the terminologies semantic among periods. We utilize Word2Vec to obtain the word embedding for terms and define the semantic change as the cosine distance between embedding vectors. Second, we use Vector Autoregression to identify the central bank behavioral response to the terminology meaning changes in the minutes.

4.1 Word2Vec

Word2Vec is a well-known word embedding method introduced by Mikolov et al. (2013) [18]. This methods use Skip-gram model to learn syntactic and semantic word representations. Our methods are highly inspired by Hamilton et al. (2016) [11]. They use word embedding methods to capture word semantic changes from historical corpora and they study the statistical law of the word semantic.

To identify the word semantic transitions across time periods, we construct word embedding for each time period. We construct word embedding from each FOMC minute quarterly corpora by gensim [22]. Then, we aligned the vectors into the same coordinate axes. As Hamilton et al. (2016) [11] discuss, this is because low dimension embedding vectors can yield arbitrary orthogonal transformations. To compare word vectors of the same word from different periods, we follow Hamilton et al. (2016) [11] and use orthogonal Procrustes to align the embedding vectors.

We also select the terminologies to be used in the analysis by the following procedure. First, we focus on the words in the economic dictionary. The words in the dictionary can be interpreted as important terminologies in economics. Second, we select the words on which the FRB place importance from the dictionary. To this aim, I map concepts in the Fed’s mandate [13] to the words in the

13https://www.federalreserve.gov/faqs
dictionary. The Fed’s mandate describes what the Federal Reserve seek to achieve through its monetary policy and this can be interpreted as what they place importance on. The Fed’s mandate says

The Federal Reserve works to promote a strong U.S. economy. The Congress has directed the Fed to conduct the nation’s 1) monetary policy to support three specific goals: maximum 2) sustainable employment, 3) stable prices, and moderate long-term interest rates.\(^\text{14}\)

We map the three terms in the Fed’s mandate to the economic terminologies: monetary policy, sustainable employment, and stable prices. The first term, "monetary policy", is straightforward because we have the same word in the dictionary. However, we need to translate the other two, "sustainable employment" and "stable prices" as they are not in the dictionary and high level concepts. We translate "sustainable employment" as "labor" because "sustainable employment" is about the labor conditions in the economy. What we can learn from "sustainable employment" is the FRB pays close attention to "labor." For "stable prices", We choose "inflation" because inflation is a typical indicator of price level and we also assume that the FRB pays close attention to "inflation".

We compute how the word embedding vectors of the selected words vary across the time periods. We use cosine distances between the word vectors for the same word, \(s\) at time period \(t\) as

\[
y_t^s = 1 - \frac{\|w_t^s \cdot w_{t+1}^s\|}{\|w_t^s\| \|w_{t+1}^s\|}
\]

where \(w_t^s\) is the word embedding for word \(s\) at time period \(t\). Note that \(y_t^s\) is non-negative, i.e.) \(0 \leq y_t^s \leq 2\) for any word \(s\) and \(0 \leq t\). By using the cosine distance, we can focus on how much word embeddings change between time \(t\) and \(t + 1\).

4.2 Vector Autoregression

To test our hypothesis that the FRB change their language when they change their policy behavior, we use a statistical model. Since the data used in this paper is timeseries data, we will use VAR, which is a generalized autoregressive(AR) model[26, 27]\(^\text{15}\). By using VAR, we can consider

\[
y_t = c + \Phi_1 y_{t-1} + \cdots + \Phi_p y_{t-p} + \epsilon_t, \epsilon_t \sim W.N.(\Sigma)
\]

where \(c\) is \(n \times 1\) interception vector , \(\Phi\) is \(n \times n\) coefficient matrix, and \(\Sigma\) is variance-covariance matrix. The vector, \(y_t\), has three endogenous variables at time \(t\). \(y_{t}^{ff}\) is a log difference of Federal Fund Rate, \(y_{t}^{mb}\) is a log difference of Monetary Base. In addition, we incorporate \(y_t^s\) into the model, which is a cosine difference of word embedding of terminology \(s\).

VAR allows us to identify the central bank behavior respond to a terminology semantic change in the minute by using impulse response. Impulse response function, \(IRF_{ij}(k)\), is the function of time \(k\), which shows the response of variable \(y_i\) to the changes in variable \(y_j\) at time \(k\).

\[
IRF_{ij}(k) = \frac{\partial y_{i,t+k}}{\partial e_{jt}}
\]

By using the impulse response function, \(IRF_{ij}(k)\), we can study how a change\(^{\text{16}}\) in a variable, \(y_j\), affects other \(y_i\) over time periods. In this analysis, we focus on how a change in \(y^s\) affects the other variables, \(y_{t}^{ff}, y_{t}^{mb}\). In other words, we focus on how the changes in the word embedding of the same words affects the FRB policy changes.

5 Results

5.1 Case Study

Before answering the research question, we conduct a case study to examine a typical example of when the FRB change both their behavioral and terminologies semantic. We study the data around the financial crisis in 2008, which caused serious economic problems. It is clear that FRB changed their policy during this time period [8, 31, 5]. Therefore, we give how the word embedding changes around these time periods to study the case of the FRB changing its behavior. Table 1 shows the top three similar words of "labor" by Word2Vec from annual corpora. We can

\(^{14}\)the numbers are assigned by the author

\(^{15}\)The lecture note by Zivot provides a good introduction.

\(^{16}\)In this paper, one standard deviation change
see, after the crisis, the most similar word changed from "ebullient" to "disorderly". Although there are some negative words before the crisis, we can see more negative words after the crisis such as "tip(s)", "fragile", "disorderly" and "foreclosed". This case study shows that looking into an important words provides a clue to understand the FRB’s perspectives and behavior at that time.

5.2 Impulse Response Analysis

I estimate the VAR model and compute the impulse responses with a 95% confident interval. Figures 3, 4 and 5 show the impulse response for each model. The most notable result is for the model with $s = \text{"labor"}$, Figure 3. The second row in Figure 3 shows a change in the word embedding of "labor" largely decreases FF and the third row shows that a change in the "labor" embedding increases MB. This means that when they change the meaning of "labor", they also change their policy behavior to stimulate the economy. Although the other two columns shows that the same results, the impulse responses are weak. The last column shows the changes in the word embedding of "monetary policy" has a weak effect on the policy behavior. These results show that each terminology affects the policy behavior to different degree.

The other interesting finding is that in all models (Figures 3, 4 and 5) a change in the terminology word embedding does not have a long effect on itself. The change in a terminology semantic not largely affect the semantic of that terminology in the following time periods. In the all first rows in Figure 3, 4 and 5, the impulse responses converge faster than other rows. This indicate that when they change the meaning of a term, they will adhere to that meaning in the following terms. This is consistent with the previous studies that show that keeping a central bank’s policy consistent is optimal monetary policy [16, 29, 9].

6 Conclusion

In today’s complex society, more and more people are obtaining higher educations to solve social problems. However, whether their professional knowledge help them to explain their solution to our society has not been explained well.

Using the dataset from FOMC minute and a information extraction technique, we found that the policy maker use their knowledge to explain
Table 1: The top three similar words of "labor" with cosine similarity around the financial crisis in 2008.

<table>
<thead>
<tr>
<th>Year</th>
<th>2005</th>
<th>2006</th>
<th>2007</th>
<th>2008</th>
<th>2009</th>
<th>2010</th>
</tr>
</thead>
<tbody>
<tr>
<td>1st</td>
<td>ebullient(0.308)</td>
<td>exchange-rate(0.31)</td>
<td>disorderly(0.272)</td>
<td>occur(0.259)</td>
<td>tips(0.293)</td>
<td>disorderly(0.333)</td>
</tr>
<tr>
<td>2nd</td>
<td>incremental(0.245)</td>
<td>opportunity-cost(0.278)</td>
<td>emerge(0.265)</td>
<td>resulting(0.246)</td>
<td>nonresidential(0.273)</td>
<td>attention(0.26)</td>
</tr>
<tr>
<td>3rd</td>
<td>disorderly(0.218)</td>
<td>financial(0.25)</td>
<td>exchange-rate(0.26)</td>
<td>housing(0.244)</td>
<td>fragile(0.265)</td>
<td>foreclosed(0.258)</td>
</tr>
</tbody>
</table>

in economic conditions or may come from the FRB policy strategy changes.

6.1 Note related to the class projects

The FRB also provides other documents such as transcripts of FOMC. Although they are informative, they are not analyzed in the report. Since most these documents are published five years after the meeting, they are not for communication with the public. However, we provide the data and the code for them in the repository. \(^{17}\)

References


\(^{17}\)For detail, please read README in the repository.


