The University of Memphis
Computer Science
COMP 7150-Fundamentals of Data Science
The Impact of Public Speeches by Federal Reserve on Interest Rates
Stephen Lee, Computer Science, Fall 2019

INTRODUCTION

The Federal Reserve System was created by an act of Congress in 1913, and they are tasked with a so called “dual mandate” to 1) promote full employment and 2) ensure price stability. In practice, the most traditional tool that they have to achieve these goals is by setting the interest rate at which large banking institutions can lend to each other overnight. This rate is known as the federal funds rate, and it is decided by the Federal Open Markets Committee (FOMC) in meetings that occur about every six (6) weeks. The decision is important because it ultimately filters through the economy, from the big banks, to consumers like you and me. It does so in the form of mortgage rates, credit card rates, interest rates on business loans, and so on, whereby the “cheaper” it is for banks to borrow money overnight, the lower the interest rate that they can afford to give consumers.

In between meeting dates, members of the Federal Reserve Board of Governors – a group that is always allowed to vote on the interest rate decision – may give speeches to the public during scheduled events. One might wonder, do these speeches contain information about their upcoming decisions?

After the financial crisis in 2008, the federal funds rate hit historic lows near what is now called the “zero lower bound”. This name alludes to the belief that interest rates could not effectively go below zero. In this new environment, the FOMC needed to engage in new and controversial monetary policy, since they couldn’t credibly lower the interest rate any more. “Forward guidance” was one such tool, whereby the FOMC would make statements about its future actions in order to spur the market to behave with confidence about what the future would hold. Thus, if they had a policy of telling the public what they would do in the future, are they effective with this new tool? More specifically, can we analyze their speech and learn information about their future interest rate decisions?

LIT REVIEW

Blei, Ng, and Jordan (2003) introduced the Latent Dirchlet Allocation (LDA) model for discovering hidden or “latent” topics from a corpus of text data. In short, LDA is a hierarchical Bayesian model where each text document is represented as words drawn from a mix of topic probabilities. These topics are often called “latent” since they are inferred rather than observed. Since its publication, it has been cited over 28,000 times, including in Hansen et. al. (2017) and Hansen and McMahon (2016), where they apply an LDA model to text data from the public transcripts of FOMCs meetings. In one paper, they exploit a policy change to examine how official communication changes under different levels of expected transparency. In the other paper, they look to see if the release of the FOMC meeting transcripts has any effect on macroeconomic indicators like stock market indices or bond yields. Perhaps most related to my work, Hayo et. al. (2010) explicitly try to explain FOMC interest rate decisions with communication indicators, although they do not employ any natural language processing techniques in their analysis. My proposed work will fill a gap in the literature by using a new set of publicly available data to explore if these interest rate changes can, to any degree, be forecasted based only on the text of official communications.

DATA

I scraped the Federal Board of Governors website and collected the text of every public speech given by a member of the Board from March 2006 to September 2019. Additionally, I gathered the interest rate decisions[1] made by the Federal Open Markets Committee in every meeting between March 2006 and September 2019.

1 Measured in basis points.

Next, I processed the text data by removing all characters that were not letters, converting all remaining letters to lowercase, removing all “stopwords”[2], stemming the words according to the Porter Stemmer[3], and finally making bigram phrases[4]. This process allowed me to reduce the dimensionality of the final vocabulary, while still maintaining more linguistic context than you get by just using the words themselves. Intuitively, this is because a bigram also contains information about the ordering and proximity of words.

Finally, I combined all speeches that occurred between consecutive meetings into a single “document”. For example, the FOMC met on July 31, 2019 to make an interest rate decision. About six (6) weeks later, they met again on September 18, 2019 to make another interest rate decision. In this case, I appended every speech given after July 31 but before the September 18th into a single text “document” that represented all of the FOMC’s communication before making that decision. While one may argue that various committee members each have a unique and separate opinions, it is therefore unjustified to combine them into a single document. In fact though, this is an advantage of my dataset as members of the Board of Governors make a conscious effort to have a unified message. Thus, while there will no doubt be some variation in topic and opinion, the members of the Board publicly make every effort to speak with a unified voice, so as to set clear expectations for the market. Figures 1 and 2 below shows a sample of raw text from a speech, and the corresponding, cleaned bigram, respectively.

Figure 1: A sample of raw speech data.

Figure 2: A sample of the same text after cleaning.

After combining the cleaned text with the corresponding meeting date and interest rate decision, I fit a Latent Dirchlet Alocation (LDA) model to the data. While the details are outside the scope of this paper, in essence, an LDA model finds groupings of words that tend to appear together across documents: these

2 Stopwords are words that appear so frequently that they don’t carry much information. Examples of common stopwords include the words “a”, “the”, “and”, and so on. I used a standard list of stopwords available at https://www.ranks.nl/stopwords.

3 Porter stemming removes the suffix of words according to a well-defined process in order to allow words with the same root map to same object. For example, the words financial, financier, and financials would all get mapped into the same word: “financi”.

4 A bigram phrase is the concatenation of two neighboring words. So the phrase, “the dog sat still”, in bigrams, would be [“the dog”, “dog sat”, “sat still”]. This allows to capture some additional linguistic context, while still maintaining a relatively small overall vocabulary. This can be generalized further into ngrams.

are often interpreted as “latent topics” since they are not explicitly defined, but rather inferred from the statistical analysis. Figure 3 below shows the results of fitting an LDA model with 12 topics. There are several important things to note. First, the choice of topic numbers is exogenously chosen, and can certainly be somewhat arbitrary. The tradeoffs in topic modeling is to use a large number of topics with typically low interpretability, or to choose a smaller number of topics with more intuitive interpretations.[5] Second, note the common appearance of the phrase “neutral rate”. This phrase is extremely common in the macroeconomic literature as it signifies, in essence, the long run equilibrium rate of interest. While it cannot be measured directly, much discussion and mathematical modeling goes into trying to make good estimates for this number. Given the prevalence of this phrase, in a follow up analysis, I would remove it and consider it a stop word. Finally, note the light grey highlighted rows. As I will show in the next section, these “topics” were selected as the most useful in predicting interest rate decisions.

Topic 0	neutral rate	bank independ	latin america	financi educ	macroprudenti polici	averag inflat	central clear	financi decis	leverag loan	protect consum
Topic 1	headlin inflat	oper risk	senior manag	neutral rate	basel framework	financi turmoil	price discoveri	labor cost	credit deriv	lender last
Topic 2	neutral rate	macroprudenti polici	headlin inflat	margin requir	asset bubbl	lender last	central clear	interest price	simpl rule	task forc
Topic 3	develop countri	longrun inflat	senior manag	neutral rate	oper risk	credit deriv	recent event	afford hous	world bank	ten year
Topic 4	neutral rate	central clear	margin requir	volcker rule	longterm debt	countercycl capit	incom wealth	macroprudenti polici	recent event	protect consum
Topic 5	neutral rate	price bubbl	credit product	central clear	oper risk	margin requir	labor cost	commun organ	frbu model	equilibrium real
Topic 6	neutral rate	central clear	mortgag servic	macroprudenti polici	longrun inflat	financi educ	financi disrupt	senior manag	capit assess	headlin inflat
Topic 7	credit deriv	mortgag servic	senior manag	financi educ	intern control	develop countri	forward rate	basel framework	oper risk	incent compens
Topic 8	neutral rate	macroprudenti polici	capit assess	develop countri	advers scenario	longterm unemploy	central clear	bank independ	macroprudenti tool	otc deriv
Topic 9	neutral rate	asset bubbl	treasuri market	equilibrium real	incent compens	financi educ	monei fund	student loan	financi turmoil	toobigtofail problem
Topic 10	neutral rate	oper risk	monei fund	margin requir	equilibrium real	corpor govern	central clear	securit market	basel framework	headlin inflat
Topic 11	neutral rate	task forc	incom wealth	treasuri market	central clear	rural area	new normal	sheet normal	fed balanc	neutral real

Figure 3: The results of a fitted LDA with 12 topics.

Finally, for each document in the corpus (i.e. each set of communication between FOMC interest rate decisions), I calculate the amount of time spent discussing each topic. Figure 4 below shows a time series of the interest rate decisions, along with the percent of communication devoted to each topic in the preceding weeks before the meeting.

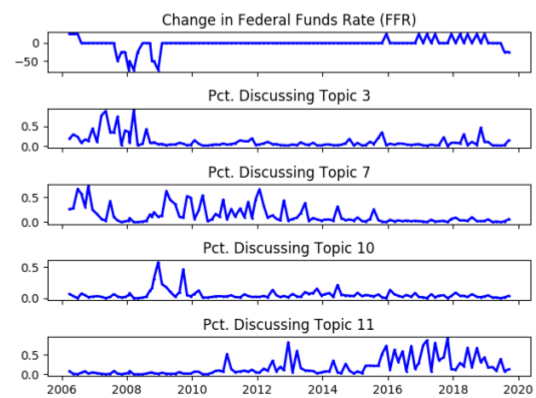

Figure 4: A time series of interest rate decisions and the time spent discussing relevant topics.

Finally, we can view the head of the dataset in Figure 5 below, noting that the final dataset has 111 observations.

```
>>> df.head()
     Topic0    Topic1    Topic2    Topic3    Topic4    Topic5    Topic6
0  0.255410  0.124658  0.043252  0.145469  0.046604  0.045981  0.041588
1  0.736499  0.016011  0.015795  0.022318  0.021972  0.026582  0.015738
2  0.510512  0.000000  0.010388  0.012247  0.000000  0.018453  0.000000
3  0.750645  0.015778  0.017898  0.019498  0.019051  0.017919  0.013994
4  0.645092  0.018795  0.016427  0.015077  0.016312  0.051070  0.000000

     Topic7    Topic8    Topic9   Topic10   Topic11  Change
0  0.060733  0.039576  0.030504  0.037390  0.128835   -25.0
1  0.014164  0.016824  0.022129  0.014010  0.077958   -25.0
2  0.000000  0.000000  0.000000  0.000000  0.394297     0.0
3  0.013389  0.017200  0.012581  0.013869  0.088178     0.0
4  0.000000  0.000000  0.012136  0.010300  0.185212     0.0
```

Figure 5: The first five (5) entries of the final dataset.

[5] Additionally, there are more rigorous methods to determining the ideal number of topics, most notably using a so called “coherence” value. This was calculated for a range of topic numbers from 9 to 40, but the results are omitted for brevity since it ultimately wasn’t used to decide the number of topics to use. That said, the coherence score did also prefer models with a number of topics between 9 and 14.

ANALYSIS

My first step in analyzing the data was to reduce the dimensionality of the covariates since I only have 111 observations, but have 12 topics. Using a L1 penalized regression[6] (also called LASSO), topics 3, 7, 10, and 11 were selected as the most relevant. These four (4) topics were then used in a standard OLS regression. The results from this exercise are shown in Figure 6 below.

Dep. Variable:	Change	**R-squared:**	0.311
Model:	OLS	**Adj. R-squared:**	0.285
Method:	Least Squares	**F-statistic:**	11.95
Date:	Mon, 02 Dec 2019	**Prob (F-statistic):**	4.73e-08
Time:	19:58:39	**Log-Likelihood:**	-459.16
No. Observations:	111	**AIC:**	928.3
Df Residuals:	106	**BIC:**	941.9
Df Model:	4		

	coef	**std err**	**t**	**P>|t|**	**[0.025**	**0.975]**
const	-0.8405	3.537	-0.238	0.813	-7.854	6.173
Topic3	-29.8516	9.124	-3.272	0.001	-47.941	-11.762
Topic7	20.5297	9.256	2.218	0.029	2.180	38.880
Topic10	-77.3146	19.405	-3.984	0.000	-115.786	-38.843
Topic11	20.9176	7.536	2.776	0.007	5.977	35.858

Omnibus:	26.904	**Durbin-Watson:**	1.035
Prob(Omnibus):	0.000	**Jarque-Bera (JB):**	53.707
Skew:	-0.978	**Prob(JB):**	2.18e-12
Kurtosis:	5.791	**Cond. No.**	13.9

Warnings:
[1] Standard Errors assume that the covariance matrix of the errors is correctly specified.

Figure 6: Results from an OLS regression of the change in interest rate (in basis points) on the LASSO selected topics.

We see here that using the four (4) LASSO selected topics in a simple linear model explains approximately 31% of the total variation in interest rate changes, based on the R^2 score. Interestingly, we find that each of the coefficients on the included topics are statistically significant to 5% levels.

In order to test how well the OLS performs out of sample, I perform a leave-one-out cross validation. Here, I iteratively withhold a single observation, and fit a model on the remaining 110 observations. Then, I use the fitted model to make a prediction on the withheld observation. Figure 7 below shows the graphical results of these predictions.

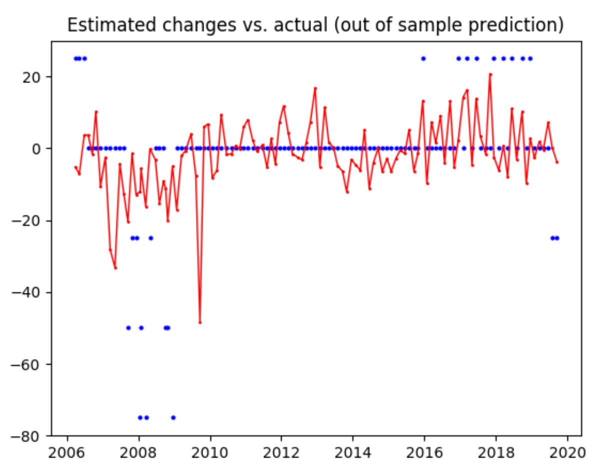

Figure 7: Leave-one-out predictions (red) compared to the actual interest rate changes (blue).

6 Note, the LASSO regression acts as a penalty on non-zero coefficients. Mathematically, it is the optimization problem $min\ \left\{\sum\left(y-\widehat{\beta x}\right)^2+\alpha\sum|\beta|\right\}$, where $\beta's$ are the coefficients, and α is the weight we give to non-zero coefficients. In this case, I chose $\alpha = 0.3$.

For robustness checks, I transformed the interest rate decision into a categorial problem[7] and then trained a neural network with two (2) hidden layers to predict the ultimate decision based on the full topic distributions (i.e. without using LASSO to select variables). While the details of this are included in other work for brevity[8], the trained neural network yielded an F1 score of 0.78 for the classification problem on the withheld test set. Similarly, this analysis is robust to other choices for the number of topics.

CONCLUSION

This work suggests that public speeches from the Federal Reserve Board of Governors do in fact seem to contain information that is useful in predicting future interest rate decisions. It is important to note, however, that this analysis does not show any causality. In any case, since the individuals that make the speeches are the same people that are involved in making the interest rate decisions, it is most likely the case that any information they embed in their speeches is reflective of their beliefs about what they will do in the future. This is to say that it’s very unlikely that giving speeches would *cause* the FOMC to make any particular decision. Rather, we may prefer to consider that the FOMC first has some beliefs about what decision it will make in the future, and those beliefs then *cause* them to deliver a speech with a particular wording.

Despite this subtle clarification, the results are no less interesting. While more study is needed to rule out spurious correlations, these results are supportive of the idea that the FOMC, intentionally or not, is constantly broadcasting information about what its next move will be. Given the fact that these speeches have never been formally studied in the literature, and the fact that the interest rate decisions ultimately impact the lives of billions of people around the world, my tentative results indicate that there may be untapped information available to market participants about what the future holds.

REFERENCES

1. Hansen, Stephen, Michael McMahon, and Andrea Prat. "Transparency and deliberation within the FOMC: a computational linguistics approach." *The Quarterly Journal of Economics* 133, no. 2 (2017): 801-870.
2. Hansen, Stephen, and Michael McMahon. "Shocking language: Understanding the macroeconomic effects of central bank communication." *Journal of International Economics* 99 (2016): S114-S133.
3. Blei, David M., Andrew Y. Ng, and Michael I. Jordan. "Latent dirichlet allocation." *Journal of machine Learning research* 3, no. Jan (2003): 993-1022.
4. Hayo, Bernd, and Matthias Neuenkirch. "Do Federal Reserve communications help predict federal funds target rate decisions?." *Journal of Macroeconomics* 32, no. 4 (2010): 1014-1024.

[7] Specifically, I mapped a rate increase to the categorical vector [1,0,0], no change to [0,1,0], and a rate decrease to [0,0,1].

[8] See Quiz 7.