# Final Project Report: Analyzing the Impact of Federal Reserve Minutes and Policy statements on Stock Market Volatility

**Group 12 Members: Karan Patel, Lee Hoang, Jason Forral**

**Abstract**

This paper investigates the effects of Federal Reserve (FED) issued FOMC Policy Statements on the S&P 500 stock index. By performing Entity-Sentiment analysis on the statements and by extracting additional NLP related features from statements such as N-Grams, paragraph, word, and POS tags counts, we attempt to predict the direction of change (i.e., up or down) in the S&P 500 stock index on the day on which a given FOMC Policy Statement is issued using Naive Bayes and Linear SVC classifiers. Given our feature set and a sample of 143 labeled (with true class labels) FOMC Policy Statements, we find that both the Naive Bayes and Linear SVC classifiers are unable to predict the direction of change in the S&P 500 stock index on the day on which a given FOMC Statement is issued. The results may also suggest that the issuance of the FOMC Policy Statements by the FED has no significant effect on the S&P 500 stock index on an intraday basis.

## 1 Introduction

The Federal Reserve (FED) system is the central banking system of the United States which controls the monetary policy in the country. Monetary policy is of economic importance since it controls and dictates interest rates, how much credit is in the economic system and many more important aspects of the US economy. The FED publishes various forms of communications from time to time to inform the public about its decisions and stance on the monetary framework. One of the communications that the FED leverages is the FOMC Policy statements which are usually composed of a few short paragraphs reflecting FED's decision on interest rates and its view on the economy. Given that the monetary policy has huge direct and indirect implications on the publicly listed companies, the publication of a new FOMC policy statement can cause the S&P 500 to move sharply in a certain direction. We aim to leverage various NLP techniques such as entity-sentiment analysis and various other NLP related feature extraction methods to build a Naive Bayes as well as a Linear SVC classifier model to see if such features and models can accurately predict the direction of change in the S&P 500 stock index.

## 2 Related Work

Harmon (2018) previously showed that the FOMC statements made by the chair of the United States Federal Reserve (FED) from 1998 to 2014 impacted the stock market uncertainty as measured by the VIX volatility index. FOMC statements made by the chair of the Federal Reserve are critical to the stock market's interpretation of the current and future monetary policy which significantly impacts asset prices across different asset classes like Stocks, Bonds, Commodities, fiat currencies, etc. By gauging the level of uncertainty that a FOMC statement might generate, traders and investors can use such information to hedge their risks against stock market crashes or potentially realize a capital gain by making calculated trades on the volatility itself (via ETFs that track VIX like VXX). There is no question that the monetary policy set by the FED has far reaching economic impacts across the globe. The research paper linked above is just one of the many papers written on the topic of FOMC statement analysis and its impact on the stock market.

The author of the paper lays out the following three hypotheses regarding the FOMC statements:

1. The more a FOMC statement explicitly reaffirms the backing underlying the monetary policy framework, the more stock market uncertainty will increase.
2. The positive tone of a FOMC statement will suppress the stock market uncertainty created by the Fed chair explicitly reaffirming the backing underlying the monetary policy framework.
3. The fear expressed in the business media at the time of a FOMC statement will amplify the stock market uncertainty created by the Fed chair explicitly reaffirming the backing underlying the monetary policy framework.

After laying down each of the hypotheses, the author validates them by calculating the "backing ratio" of different FOMC statements to the change in VIX (which is one way of measuring the stock market volatility) during the time of the FOMC statement. The author calculates the backing ratio as follows:

Backing ratio = (number of paragraphs that make the backing explicit / total number of paragraphs)

The author shows that the more the FED chair explicitly reaffirms the backing underlying the monetary policy framework, the more this creates stock market uncertainty. Similarly, the author calculates the positive tone of FOMC statements (using a text analysis software called Linguistic Inquiry and Word Count) as well as business media fear (using Thomson Reuters MarketPsych Indices proprietary database) and compares them to the stock market uncertainty using VIX to reaffirm his second and third hypothesis as listed above.

As a result of the research performed by the author, he attempts to establish links between the following three different aspects of the FOMC statements to volatility in the stock market:

- The extent to which the FED chair reaffirms the "backing" underlying the monetary policy framework
- Magnitude of the positive/negative tone of the FOMC statement
- Fear in business media as an amplifying factor to the uncertainty

By linking different aspects of the FED chair FOMC statement to stock market volatility, the author contributes to the broader analysis of FED communication on stock market volatility. He demonstrates that explicit reaffirmations that are true but taken for granted create market uncertainty.

Some of the limitations of the paper was that the paper only looked at the FED chair's FOMC statements and not at various other forms of communications that FED leverages such as FOMC policy statements, FED chair's semi-annual monetary policy to Congress amongst others.

We choose this research paper because it clearly laid out and proved three distinct hypotheses that affect volatility. The author also clearly laid out the way he calculated and measured different aspects of FOMC statements such as positive tone, backing ratio, etc. The author also laid out a regression-based model that he used to prove out his hypothesis that was not too trivial nor too complicated, it was easy to understand and relate to.

## 3 Methodology

We implemented the code (using beautifulsoup4) for scraping the FOMC Policy statements from the Federal Reserve Board - FOMC Materials Filter website.

We conducted Entity-Sentiment analysis of FOMC policy statements using Google Cloud's Natural Language API to extract entities and their sentiment scores ranging from -1 (most negative) to 1 (most positive). This helped us identify and compute sentiment values for various important entities, like unemployment, for each FOMC statement.

We obtained the intraday change in S & P 500 from Polygon.io - Stock Market Data APIs.

Since we already know entities of importance (like inflation, GDP growth, unemployment) that would be economic in nature, we can analyze if we are able to identify entities that we expect to see. We also manually reviewed and filtered the final set of entities before we start using them as features for our binary classification models.

It was decided to use Multinomial Naive Bayes (NB) and Linear SVC models (developed using scikit-learn ) instead of a Deep Neural Net as a binary classifier since there were only 201 samples available. The two models would help predict if the S&P 500 went up or down (on an intraday basis) on the day on which the FOMC statement was released. The models would be trained and make determinations/classifications based on the extracted features from the FOMC policy statements such as the entity sentiment values on entities of importance as well as features from statements such as N-Grams, paragraph, word, and POS tags counts.

## 4 Experiments and Results

Following are the 10 most common entities across all 201 policy statements for which sentiment values exists as identified by GCP's Natural Language API:

| **Most common entities with sentiment values** | **Counts across 201 samples (i.e. speeches)** |
|---|---|
| committee | 90 |
| employment | 75 |
| price stability | 75 |
| mandate | 75 |
| inflation | 61 |
| investment | 61 |
| growth | 55 |
| household spending | 48 |
| stance | 47 |
| business | 43 |

Table 4.1Most Common Entities

We used the sentiment values on all the entities listed above as features to our binary classification models.

We conducted experiments with two different classification models: Multinomial Naive Bayes and Linear SVC. The features that we used included word count, paragraph count, n-gram probability count, pos tag (noun, verb, adjective) count, and entity sentiment of certain words. The goal was to determine the classifiers' accuracy in predicting whether the change in S&P 500 was positive or negative.

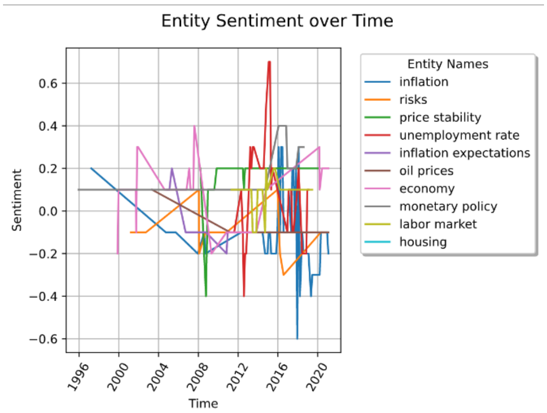


Figure 3.1: Sentiment over time for economic keywords.

The first experiment used the Naive Bayes classifier and the features. The second experiment used Linear SVC. Aside from using a different model, we kept other aspects of the second experiment the same.

For training the models, we split the labeled dataset of 143 FOMC statements such that we had into a training set, a validation set, and a test set. 80% of the dataset (114 samples) was set aside for training while 10% was used for the validation set and test sets each (14 for validation, 15 for test).

As a preliminary task, we conducted entity sentiment analysis on economic keywords (inflation, risks, price stability, unemployment rate, etc.) to determine how the sentiment of these entities changed over time.

## 5 Conclusions

Our Naive Bayes and Linear SVC models achieved average cross validation accuracy of 42.66% and 43.00%, respectively. The test accuracies for Naive

F-score:

| | precision | recall | f1-score | support |
|---|---|---|---|---|
| DOWN | 0.56 | 0.71 | 0.63 | 7 |
| UP | 0.67 | 0.50 | 0.57 | 8 |
| accuracy | | | 0.60 | 15 |
| macro avg | 0.61 | 0.61 | 0.60 | 15 |
| weighted avg | 0.61 | 0.60 | 0.60 | 15 |

Figure 5.1: Naive Bayes F-score

F-score:

| | precision | recall | f1-score | support |
|---|---|---|---|---|
| DOWN | 0.44 | 0.57 | 0.50 | 7 |
| UP | 0.50 | 0.38 | 0.43 | 8 |
| accuracy | | | 0.47 | 15 |
| macro avg | 0.47 | 0.47 | 0.46 | 15 |
| weighted avg | 0.47 | 0.47 | 0.46 | 15 |

Figure 5.2: Linear SVC F-scores

Bayes and Linear SVC were 60% and 46.66%, respectively.

Both cross validation scores were under 50%, but they were close to it. The F-score results revealed that the classifier had a higher F-score for DOWN predictions than UP predictions. Each time

| | Iterations | | | | | Average Score |
|---|---|---|---|---|---|---|
| | 1 | 2 | 3 | 4 | 5 | |
| **Naive Bayes** | 0.4 | 0.67 | 0.4 | 0.67 | 0 | 0.4266 |
| **Linear SVC** | 0.4 | 0.25 | 0.25 | 0.25 | 1 | 0.4300 |

Table 5.1 5-fold cross validation results for the Naive Bayes and Linear SVC classifier.

we ran the classifiers, we got inconsistent results. The values shown above reflect the results of two random runs for each classifier.

For both the Naive Bayes and Linear SVC classifiers, the test accuracies, cross validation scores, and F-scores fell within the 50% range, meaning that the classifier we made is as accurate as a coin toss. This may be due to the small amount of input data that we obtained; only 201 FOMC statements were available and out of 201 statements, only 143 had labeled data for changes to S&P 500.

In addition, using 80% of the dataset for training may have led to overfitting, a phenomenon that results when the test set accuracy is significantly lower than the training set accuracy. Interesting Insights

## 6 Interesting Insights / Pivots

While working on this project, we found that the sentiment for entities such as “economy” and “price stability” had spikes around the year 2008. This made sense to us as 2008 was the same year in which the Great Recession, a large financial crisis, occurred.

The original plan for this project was to use a Deep Neural Network from Tensorflow or PyTorch as a binary classifier. It would allow us to tune the number of layers as a hyperparameter. However, upon examining the FOMC statement dataset, we discovered that there were only 201 FOMC statements available, and of those 201statements, only 143 had data available for change in S&P 500. Given the large training requirements for a neural network, we decided to switch our binary classifier to either Multinomial Naive Bayes or Linear SVC from Scikit-Learn.

If more time had been made available, we would gather more FOMC statement data or other economic data to help bolster the initial 201 statements that we obtained. If more data was available, we may switch to using a different classifier such as a neural network.

## 7 Ethical Considerations

Most of the data used for this experiment is publicly available. The FOMC policy statement can be obtained for free from the FED website, but the S&P 500 index quotes typically require payment to be made to a data vendor to be accessed.

An investor with access to a lot of capital will have an advantage over others and will benefit the most since such an investor would be able to make larger bets and turn out a larger profit assuming that a working predictive model exists.

## 8 Future Directions

We recommend the following modifications and extensions to our binary classification model:

- For the binary classifier, include more useful features like Wall Street sentiment score, current unemployment %, etc. which can help the classifier be potentially more accurate at predicting change in the stock market.

- Train the binary classifier with different kinds of FED statements like FED minutes, FED Greenbooks, Memoranda of Discussion, etc.
- Create binary classifiers for all possible assets like commodities (gold, crude oil), bonds, individual stocks (GameStop, Apple, AMC), etc.

## 9 Contributions

Following is a link to our GitHub repository containing code that we wrote and used to download FOMC policy statements from FED's website as well as extract various features in addition to the code for our two binary classification models: https://github.com/pkaran57/FED-NLP-Project

## References

Derek J. Harmon. 2018. When the Fed Speaks: Arguments, Emotions, and the Microfoundations of Institutions. *Administrative Science Quarterly*, 64(3):542–575.