Powell's Power

A deep dive into the Fed's Impact on Financial Markets

By:
Siddhant Mehta – mehtasi6 – 10073242449
Eyad Elghobary – elghobar – 1006886835
Alexander Sanderson – sande150 - 1006905627

Introduction

In 2024, the US GDP reached $25 trillion USD in nominal terms! The US Fed has historically been responsible for communicating important monetary policy decisions in the world's largest economy. As such, the Fed helps maintain overall macroeconomic stability, inflation expectations, growth prospects, and unemployment through its speeches. Undoubtedly, the Fed's impact extends beyond the United States of America to Canada - with their closely linked supply chains - and to the rest of the world.

At the mast of the ship that is the Fed is Jerome Powell. Jerome Powell is a prolific investment banker and attorney who has served as the 16th Chairman of the Fed since February 5th, 2018 (Jerome. H Powell). He makes monthly interest rate announcements for hikes and cuts to influence price levels and growth in the economy. His outlook towards the future determines foreign and domestic investment in the US stock market.

Motivation and Justification

"*When Jerome Powell talks, the world listens*." - Tyler Paul, ECO349, Money, Banking & Financial Markets. In this course, we would begin every lecture by manually analyzing Powell's speeches and then seeing the impact on stock prices. Further, in November 2023, the New York Times released an article titled "Why Jerome Powell can be the most important man in Washington in 2024." This motivated our selection for the topic.

As students in ECO482, *Machine Learning applications in Macroeconomic Finance*, we were interested in **quantifying and correlating** Powell's words and stock market movements. Machine learning tools enable us to analyze the sentiment of Powell's speeches on a large-scale to visualize trends.

The market cap of US stocks is $50.8 Trillion USD while the world GDP is around $100.8 Trillion USD. As a share, the US stock market cap is 50% of world GDP! The opportunity to delve into the determinants of price fluctuations in the stock market motivated us.

For this project, we decided to monitor two indices : The S&P and the VIX. The S&P represents the 500 biggest companies in the USA. Approximately **55 million US households**, representing over 100 million individual investors, owned mutual funds invested in US equities, many of which track or are influenced by the S&P 500. Therefore, we believed that monitoring the S&P would reflect market movements as a whole (Pew Research Center). Similarly, the VIX index is a measure of market volatility of S&P options. Measuring the VIX gives us information about investor expectations.

Finally, we were motivated to test the *efficient market hypothesis*. This theory suggests that the stock price reflects all publicly available information. Powell's speeches on rate cuts and hikes and important pieces of information and so we were determined to see *if and how quickly* the market incorporates Powell's words into stock prices.

Research Questions

The research question for this project is as follows : Under the Efficient Market Hypothesis, in what way do Jerome Powell's speeches on monetary policy decisions impact market returns and volatility?

Our research question is significant as we can measure the strength of correlation between Powell's speech and stock prices. In addition to our main research question, we have some additional questions that direct this project and help conduct our research. They are as follows :

- What is the Sentiment Score of each of Powell's speeches?
- What words, phrases and statements are important for the Sentiment Score?
- How accurate is our sentiment analysis in predicting a market dip or rally?

Answer to Research Question

The S&P is our measure of market returns and the VIX is our measure of volatility in the market. When we run the sentiment analysis models on Jerome Powell's speeches, we find that the market truly does incorporate his words, validating the Efficient market hypothesis. Our research shows that the sentiment of his speeches is primarily neutral as he does want to cause any strong market crashes. However, when the prevailing sentiment is negative (for a rate hike) the market does indeed dip. During macroeconomic shocks like Covid, the Trump Impeachment and Russia - Ukraine crisis, there is a strong correlation of negative sentiment and market dips in the following two weeks.

Contribution

This research aims to contribute to the literature surrounding macroeconomics, finance, and behavioral economics by quantifying the effect that Jerome Powell has on expectation and actualized return as well as volatility of the stock market. The use of the Finbert model for sentiment analysis has primarily been for Financial statement analysis. We hope to extend its use to financial speeches.
We also hope to see the lagged effect of Powell's speeches. We hope to see the impact 1 day, 1 week and 2 weeks later. We use different time spans to test how quickly the market responds to Powell's word under the efficient market hypothesis.

Literature Review

For the Literature review, we hoped to go through papers and find the similarity and difference between their work and this paper's. The following were the results :

"Deep Neural networks applied to Stock Market Sentiment analysis." – Filipe Correia
Correia et al apply neural networks to the Dow Jones Industrial Average to assess different neural network techniques on price prediction using sentiment analysis data

Similarity: We use sentiment analysis and a system similar to Filipe Correia. The VIX index is shared between our work and Correia et al.
Difference: The sentiment analysis data utilizes social media, particularly Reddit, for its sentiment analysis data while we use Jerome Powell's speeches

"Can Central Bank speeches predict financial market turbulence? Evidence from an adaptive NLP sentiment index analysis using xgboost machine learning technique." - Petropoulos, Anastasios

Similarity: This study provides a similar aggregation of sentiment of central bank sentiment analysis as well as random forest boosting to assess the importance of key variables and prediction
Difference: Petropoulos et al aggregate datasets of central banks of multiple countries while our study focuses on the United States and its domestic market

"The informative value of central banks talks: a topic model application to sentiment analysis" - Priola, Molino et al

Similarity: Usage of LDA to find positive or negative sentiment. This is used to produce an index score for central bank confidence.
Difference: Priola et al predict monetary policy using sentiment analysis using multiple central banks.

Data & Methodology

Sources

The data that was required for our analysis was freely available on multiple APIs that were available as Python packages. To obtain Powell's speeches, the FRASER API was used. FRASER is the economic history division of FRED, the Federal Reserve of St. Louis. Specifically, the speeches were categorized under the "*Statements and Speeches of Jerome H. Powell*" section. For the S&P500 and VIX data, it was taken using the Yahoo Finance API. This allows for accurate data for both the S&P500 and VIX, which are used as our market index and market volatility expectations index respectively. The range of the dates for all data was between February 5th, 2018 and December 31st, 2023 to ensure consistency with the speeches.

Descriptive Statistics - Speeches

Figure 1: Word Cloud of Powell's Speeches

Word Cloud analysis of Powell's speeches can be seen in Figure 1. Word Cloud analysis emphasizes common words and phrases that have high occurrence relative to the entire document. There is an abundance of macroeconomic and financial terms such as inflation, monetary policy, and labor market. This is a small excerpt from the entire dataset, however many of the speeches have high similarities in word choice.

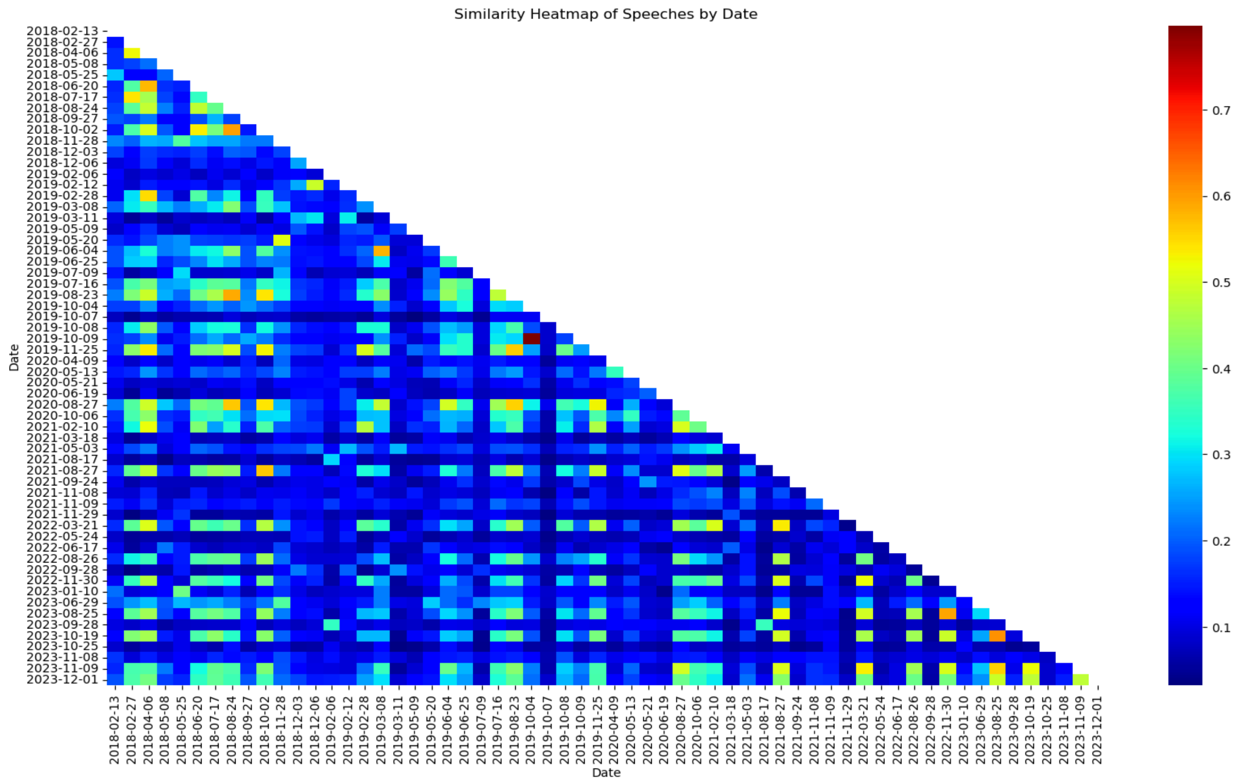


Figure 2: Heatmap of Powell's Speeches

This heatmap shows the correlations between speeches on specific dates. There is high correlation between many of the speeches, which stems from the large amount of financial and macroeconomic terms that are used across Powell's speeches. The discrepancy in the speeches stem from different events that affect macroeconomic conditions, however Powell keeps his focus on the broader economy.

Methodology

Web Scraping

Obtaining the data required scraping the FRASER database to obtain Powell's speeches starting with his inauguration on February 5th, 2018 to December 31st, 2023. Using the FRASER API, we were able to obtain a JSON dictionary that obtained the metadata of the speeches. The JSON dictionary was deeply nested. The important parts of the metadata contained the date and the link to the html text of the speech. We then obtained the texts by using an algorithm that scraped the url from the API using the requests library, then assigned a speech to its respective date in a dataframe.

Data Cleaning and Processing

The raw text data frame was cleaned using nltk stopwords and punctuation. Every cell that contained the text was treated as a string, so new line characters were replaced as well as lowering the text. Preprocessing eliminated meaningless phrases such as the date, time of speeches, and the unnecessary speech information at the top of the website using REGEX. The words were then tokenized, lemmatized, and then had collocations and phrases applied from Gensim before further analysis.

Sentiment Analysis
We opted to use 2 models for our sentiment analysis: VADER and FinBert. VADER is a pre-trained NLTK model that serves as a baseline for our analysis, and FinBert is an extension of Google's Bert model that is pre-trained on financial text. By using the 2 models on our data, we were able to identify unique sentiment specific to financial language as well as the most pivotal aspects of financial language to understanding sentiment.

Calculating sentiment using the VADER model was straightforward, we simply passed the cleaned speech to the model, and sentiment (polarity) score was calculated. The case wasn't as straightforward when it came to the FinBert model, as there was a limit to the number of tokens, 512, that could be passed to the model.

As a result, our approach was to break down the cleaned speech to chunks of 512 tokens using the **encode()** function. This function also added the necessary special tokens to each chunk and made sure that any chunk with less than 512 tokens would be padded up to 512 to ensure uniform size. Each chunk was passed to the model, where the chunk was given a logit score that was converted into a probability using the **softmax()** method. The overall sentiment of the speech was then calculated by taking the average of these probabilities across all chunks, providing a comprehensive sentiment analysis of the entire speech.

Using each model, we were able to assign a probability of the cleaned speech being negative, neutral, and positive., which we used to determine the sentiment of a speech.

Plotting Sentiment - Index Return Relationship
Using the sentiment scores from our two models, we decided to graph a time-score plot for each index. This approach enabled us to analyze the relationship between sentiment and stock returns, considering lagged effects over intervals of one day, one week, and two weeks. Looking at the lagged effect over different time intervals, allowed us to observe how market reactions to sentiment changed over time.

To graph the plots, we used a dataframe for each index: **merged_sp** and **merged_vix.** These dataframes contained the sentiment of a speech, grouped with the returns of the relevant index 1 day after, 7 days after, and 14 days after. To achieve this grouping, we used a few dataframes:

1) **daily_sentiment_df:** This dataframe contained the date the speech was made with the sentiment scores using both models.
2) **sp_new:** This dataframe contained a date with the lagged return of the S&P500 index over the different time intervals.
3) **vix_nex:** This dataframe contained a date with the lagged return of the VIX index over the different time intervals.

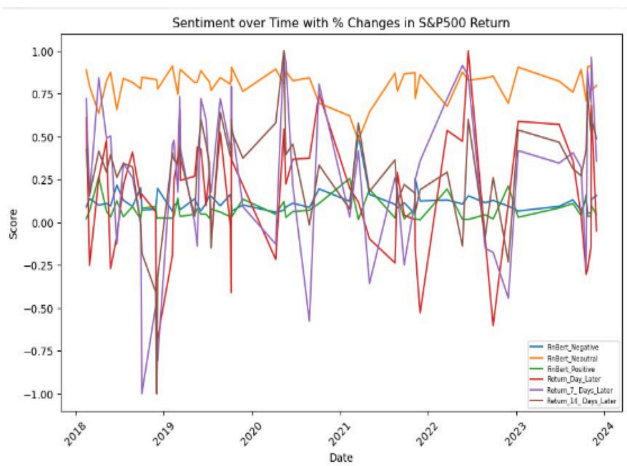
Given the above dataframes, we were able to merge the relevant ones by date, allowing us to create merged_sp and merged_vix dataframes.

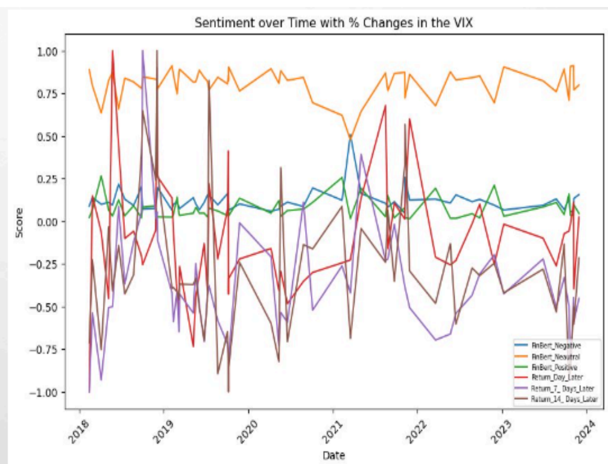
Random Forest

The Random Forest regression analysis was our way of associating certain sentiment with changes in the expectations of volatility and the actualized return of the market over certain periods. We split the data into a 75% training and 25% test split as there were only 60 speeches. We trained the different time spans on the three different sentiment values that were produced by Finbert. Using this we used gridsearch on a normal decision tree to analyze the depth that generated the largest score. We found that the depth that yielded the highest sklearn score was 1. However, to examine the split in greater detail, we analyzed the ensemble methods with a depth of two to see the split in greater detail as small changes in negative or positive sentiment could change the results of the model greatly. We applied the data to two ensemble tree methods, Random Forest and Gradient Boosted Trees. Both of these models were run with a max depth of two and were given an R-Squared score between negative infinity and 1 where higher is better. The Gradient Boosted Tree was run with 10 estimators and a learning rate of 0.1. This allowed us to compare the models to see which performed best on the test set.

Results

Sentiment - Index Return Relationship (FinBert Model)





These are the plots we got by looking at sentiment over time. We made sure that the score was scaled to the range [-1,1] to allow us to better examine how the index returns fare against the sentiment scores.

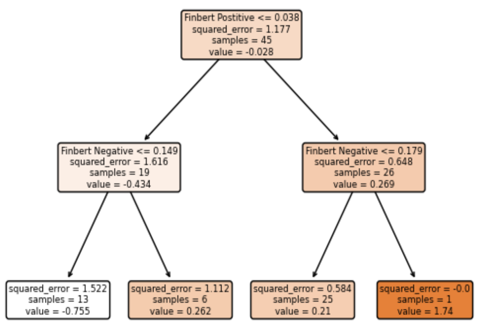
From these plots we can see that the prevailing sentiment on both graphs is mainly neutral, which we believe is likely a deliberate strategy by Powell in an attempt to ensure market equilibrium. We can also observe the relative correlation between sentiment and market returns. For the S&P 500, this correlation is positive; where positive sentiment outweighs negative sentiment, we generally observe surges in index returns, with the highest impact seen after a one-week lag. This suggests that the market typically reacts most significantly to Powell's speeches about a week after they are delivered. For the VIX, the correlation is negative; where positive sentiment outweighs negative, we generally see a downturn in index volatility, with the most pronounced effects after one and two weeks, indicating that the market adjusts to changes in sentiment with some delay as well.

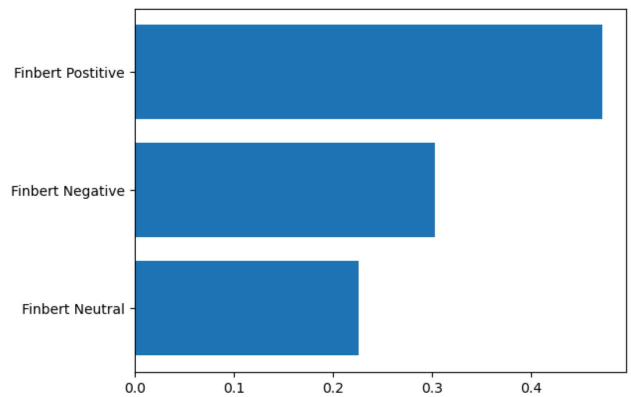
Using this information, we can see that when sentiment is more likely to be positive, there are increased returns for the S&P 500 which takes place with market growth. When sentiment is more positive, there are decreased returns for the VIX indicating reduced market volatility which also comes with market growth. Overall, our analysis suggests that the market growth anticipated by positive sentiment indeed tends to manifest as actual market growth.

Regression Tree Analysis

All the feature importance and decision trees were done with a max depth of two and the ensemble methods were compared. Random Forest performed better on the test sets with a higher sklearn score for all the time periods when compared to Gradient Boosted Trees with ten estimated trees. The results are estimated from the Random Forest model.

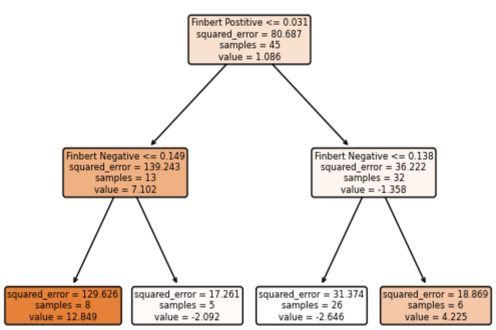
Short Term - S&P 500 One Day After Speech

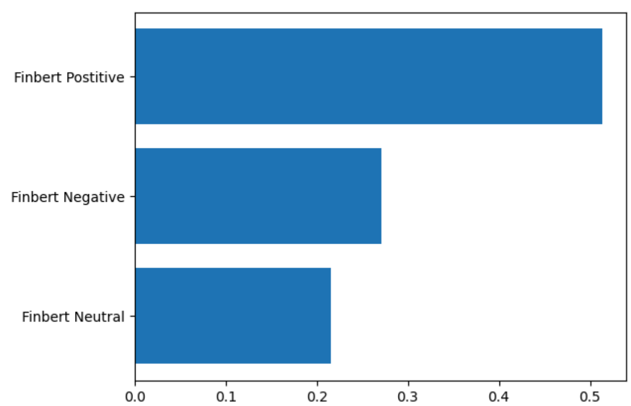




The results one day after the speech for the S&P 500 reveal that positivity influences changes the most, with negative sentiment second, and neutrality being the least influential. Higher positive sentiment leads to higher one day average returns in the market. The decision tree demonstrates how any change in the returns is due to small changes in the positive and negative sentiment of Powell's speeches.

Short Term - VIX One Day After Speech

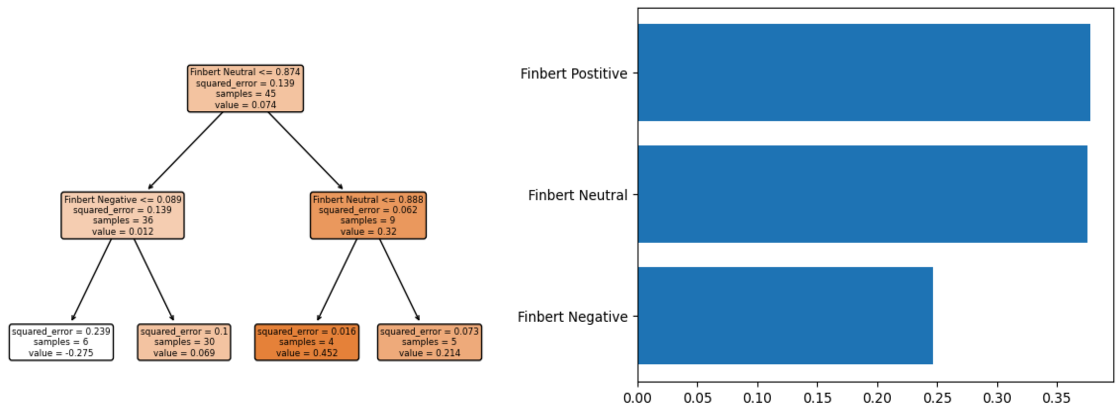




The Random Forest Model revealed that similarly to the S&P 500 in the short term, positive sentiment around Powell's speeches are the most influential to the VIX. Negative and neutral sentiment were second and third respectively. Positive sentiment is associated with decreasing the VIX, which means that investors do not expect the market to be volatile in the short term.

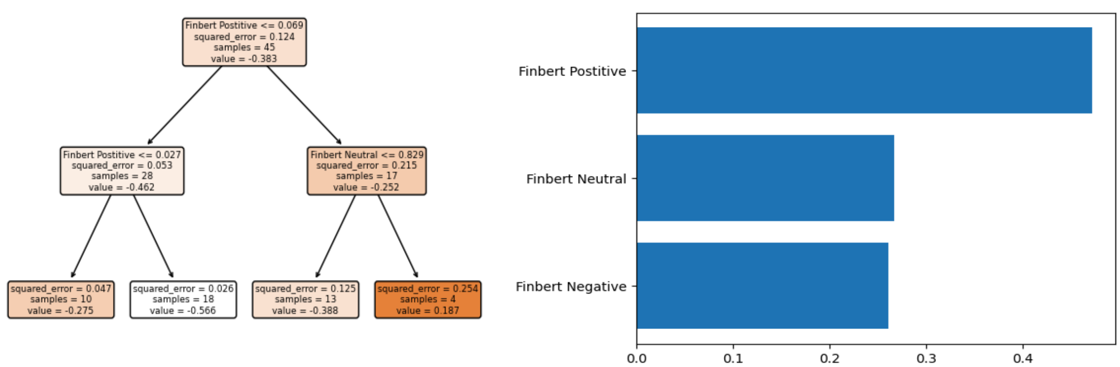
The results for the S&P 500 and the VIX one day after Powell's speech demonstrates his influence on the market and how the positive and negative sentiment in his speeches are small, but slight changes create significant changes in the market. This affirms the idea that Powell can actualize market rallies in the short term.

Long Term - S&P 500 One Week After Speech



Random Forest feature importance one week after Powell's speech revealed that positive and neutral sentiment are approximately equal in the influence on the S&P 500 weekly returns. The majority of his speeches are neutral.

Long-Term - VIX One Week After Speech



Analysis of the sentiment importance behind the VIX reveals that positive sentiment is the most influential to longer term expectations while neutrality and negativity are similar but almost half of the influence of positivity. This affirms that positive sentiment from Powell's speeches results in lower expectations of volatility for the market.

Random Forest analysis shows how Powell's speeches influence the market in the longer term. Powell attempts to stay neutral to ensure market stability and keep expectations anchored. He tries to be slightly positive to reassure investors when there are events that negatively affect the market, through long term growth and short term recovery in times of crisis.

Conclusion

Our results test the Efficient market Hypothesis to see how the stock market incorporates Powell's words by plotting changes in stock prices against the Sentiment score of models. Our results suggest that given Powell's influence on the market, we believe that he is in fact aware of his effect on the market and is cautious of the words that he chooses: he attempts to look past negative events and stay positive or at least neutral.

The sentiment analysis plots show a correlation between increased/decreased returns and positive/negative sentiments. The positive relationship between sentiment and market movements indicates the influence of the sentiment of Powell's speech on changes in the market. This plot also measures the period of responsiveness of the market within the 5-year interval that Powell has served as Chair of the Federal Reserve.

Finally, our regression analysis demonstrates the nuanced impact of Jerome Powell's speeches on market indices. In the short term positive sentiment from Powell's speeches had the most significant influence on both the S&P 500 and the VIX, leading to higher market returns and lower volatility, just one day after he speaks, Over a longer period of one week, positive sentiment continues to play a crucial role, especially in reducing market volatility expectations. These findings further highlight Powell's subtle yet impactful role in shaping market sentiments and influencing both immediate and anticipated market movements.

CITATIONS

- "Jerome H. Powell." *Federal Reserve History*, www.federalreservehistory.org/people/jerome-h-powell. Accessed 6 Apr. 2024.
- Pew Research Center. "More Than Half of U.S. Households Have Investment in Stock Market | Pew Research Center." *Pew Research Center*, 27 July 2020, www.pewresearch.org/short-reads/2020/03/25/more-than-half-of-u-s-households-have-some-investment-in-the-stock-market.
- Rohan-Paul. (n.d.). *LLM-Finetuning-large-language-models/other-language_models_bert_related/finbert_long_text_part_2.ipynb at main · Rohan-Paul/LLM-Finetuning-large-language-models*. GitHub. https://github.com/rohan-paul/LLM-FineTuning-Large-Language-Models/blob/main/Other-Language_Models_BERT_related/FinBERT_Long_Text_Part_2.ipynb