***Vaughn Hajra, Econ 475***

***Formal Project Proposal***

* 1. ***Proposal:*** This document can be thought of as the very beginning of your research paper (ie an early and incomplete draft) and should include the following:
  2. a. A detailed outline of your paper;
  3. b. A refined statement of you research question;
  4. c. A discussion of your empirical approach, including the source of your data and description of your econometric model;
  5. d. An annotated bibliography of books and journal articles you have used in your research for the paper. You need to include full, clear citations; and
  6. e. Article summaries for at least three of your empirical references.

**Refined Research Question**

To what extent did the sentiment of President Donald Trump’s tweets mentioning publicly traded companies correlate with the subsequent stock market performance those companies? How does the specific sentiment value (very strong, slightly strong, or neutral) for positive and negative tweets relate to stock outcomes, when controlling for the reach of the tweets and company industry?

This project will analyze the influence of President Trump’s tweets on the stock performance of publicly traded US companies during the entirety of his presidential term. By filtering a dataset of these companies against Trump’s tweets and then conducting sentiment analysis on those tweets, I aim to understand the potential impact. Previous studies have conducted very similar methods, but the datasets were limited, and by including the entirety of Trump’s presidency, I hope to add to the existing research. Additionally, I hope to include control variables (related to the specific reach of the tweets, and the industry of the mentioned companies) that previous studies did not do, with the idea being of capturing the true effect for positive and negative tweets. Additionally, I plan to use multiple sentiment analysis databases, and run the same analysis using all three which could lead to some interesting results when comparing the sentiment scores.

*Continued on next page*

**Empirical Approach – Events Study**

***Datasets & Other Resources***

[*Trump Twitter Archive*](https://www.thetrumparchive.com/)

[*S&P 500 companies (with industry)*](https://github.com/datasets/s-and-p-500-companies/blob/main/data/constituents.csv) *Dataset*

[*Stock Price on Specific Date Using Excel*](https://exceljet.net/formulas/get-stock-price-on-specific-date#:~:text=This%20can%20be%20done%20with,the%20worksheet%20into%20multiple%20cells.)

[*Sentiment Analysis in R*](https://www.tidytextmining.com/sentiment)

As far as the data processing is concerned, Here is the proposed workflow:

1. Cross-reference the S&P 500 companies with President Trump’s tweets
2. Build a list of tweets, the companies mentioned, and the industry
3. Add stock data for each individual tweet, using multiple points in time both before and after the tweet was made so an events study method can be done
4. Run sentiment analysis on the tweets
5. Run OLS to calculate the alpha and beta effects used for calculating abnormal returns (see next section for modeling specification)
6. Calculate abnormal returns, average abnormal returns, and cumulative abnormal returns
7. Run the appropriate test (most likely t-test unless the data are non-normal) for hypothesis testing
8. (Optional additions) – run multiple tests, broken down by specific industry if the dataset is large enough
9. Run the analysis with multiple sentiment analysis word banks

***Econometric Approach***

This econometric model is based on others I’ve seen, but I would like to find a way to add controls to it. The events study by Brans is one I really like and that is the econometric method I will be listing here.

First, calculate abnormal returns: , where *ARit* is abnormal return, the *Rmt* is the return of market index, and *Rit* is the actual return. Alpha (risk involved with stock) and beta (sensitivity between individual stock and market) are OLS estimators of the market model. My hope is I can add the industry controls in the OLS estimators for the stock returns, or also run separate analyses by industry.

Next, calculate abnormal returns: where *n* is sample size, and also aggergate to get the cumulative abnormal returns: .

To test the significance, I hope to also perform a similar t test to the Brans paper, where the t statistic is the portfolio mean return on day t, divided by its standard deviation. Like I previously mentioned, although the test itself is very similar, I can add to the existing research by using a larger sample size and better control variables. The Brans study also weighted the portfolio mean, which would also be worth considering.

**Paper Outline**

**Abstract**

* Give a brief summary of the main findings, methodology, and implications

**Introduction**

* Present a background of why this study is important
  + Implications of potential market manipulation
  + Potential need for regulation
  + Oversight into major political figures’ social media?
  + How COVID changed how people communicate and take in information
* Summarize existing literature, their findings, and how they contradict each other
  + How methodology of specific papers may cause different conclusions
    - How this led me to the method I took

**Theoretical Framework**

* Present the model and framework from a theoretical standpoint
  + Focus on sentiment analysis and how data was collected and analyzed
  + Benefits of using multiple sentiment analyses

**Methodology**

* Discuss specific models and analysis run (see previous sections for specific data and econometric methods)
* Build upon theoretical framework and explain **why** everything was done the specific way

**Results or Findings**

* Detail the main results of the research
* Discuss the robustness of the model

**Discussion**

* Discuss findings, and why they may be the way they are
* Compare the findings to those done by other papers mentioned in the introduction
  + Consider (and potentially justify) why the differences may occur
* Discuss shortcomings of the study
  + Propose future improvements to build upon the current research

**References/Bibliography**

* List complete references and bibliography (and make sure sources are cited appropriately throughout the paper

**Appendices**

* Consider adding larger tables or graphs/charts to appendix rather than in the middle of the paper if they take up a lot of space
  + (Ensure the tables or figures are referenced appropriately during the paper)

**Acknowledgements**

* Thank the appropriate people who assisted in making this paper possible!