Capitol Gains: Investigating Insider Trading in the United States House of Representatives Using Time Series Analysis

James Burke Pagano

Advisor: Dr. Daniel Rigobon

Submitted in partial fulfillment
of the requirements for the degree of
Bachelor of Science in Engineering
Department of Operations Research and Financial Engineering
Princeton University

April 2024

I hereby declare that I am the sole author of this thesis.

I authorize Princeton University to lend this thesis to other institutions or individuals for the purpose of scholarly research.

James Burke Pagano

James Burke Pagano

I further authorize Princeton University to reproduce this thesis by photocopying or by other means, in total or in part, at the request of other institutions or individuals for the purpose of scholarly research.

James Burke Pagano

James Burke Pagano

Abstract

This thesis investigates closed-door congressional meetings as a potential source of insider trading in the United States House of Representatives. Employing two sample t-tests, ARIMA models, and regression discontinuity design (RDD), the trading activity of congresspeople around dates when they may receive nonpublic information is analyzed. Through these methods we find that there is a statistically significant immediate increase in trades made by representatives on the day of closed-door meetings. This phenomenon is not as pronounced in open meetings or during normal trading days. This finding contributes evidence to the idea that congresspeople use the confidential information in these meetings for personal financial gain. From these results, this thesis suggests the need for stricter transparency and regulation regarding congresspeople's trading activities.

Acknowledgements

First I would like to thank my advisor, Dan Rigobon, and my preceptor, Rajita Chandak. Thank you for giving me the freedom to explore this topic in American politics through the lens of an Operations Research and Financial Engineering thesis. Week after week and meeting after meeting you invested your time to allow me to produce something I am proud of. Your guidance and support have made this long process one of the most rewarding experiences of my life.

Thank you to all my friends who I have met along the way here at Princeton. Everyday I wake up excited to tackle life with you. Whether on Prospect, in Firestone, or anywhere in between, the memories I have made here have perfectly balanced out all the trials and tribulations of the rigorous academic experience at Princeton. I truly wouldn't have it any other way. This thesis would never have been completed if it weren't for the countless laughs and memories we have shared in this place.

Finally, to my incredible parents, this thesis represents the culmination of my academic career and none of that would be possible with out you guys. Thank you for investing in and supporting my education for as long as I can remember. I am continuously inspired by the mantra you have repeated throughout my life: "to whom much is given much is required."

To my friends and family

Contents

List of Tables

List of Figures

Chapter 1

Introduction

In the United States, politicians are under constant scrutiny to ensure they are acting with their constituents' best interest in mind, and not solely for personal gain. Since the 1970s, in the wake of Watergate Scandal, there has been a movement to make the proceedings of the US government more visible to the American people. Simultaneously, people have begun to dive deeply into any and every public piece of information they can get their hands on in order to understand politicians' underlying beliefs and make an informed decision when it comes to their vote. An interesting avenue for understanding politicians' true motivation is to look at their personal finances during their tenure, specifically their trading history.

This thesis examines the trading activity of members of the House of Representatives in the legislative branch of the United States federal government to see if there is a relationship between when these politicians may receive nonpublic information and when they perform trades. If politicians receive confidential information that the public does not have access to and trade a public company's stock based on that nonpublic information, then they are breaking the law by "insider trading" [14]. Investigating the trading activity in this way allows for more insight into whether or not the people in office are interested in addressing the problems they campaign on, or if their primary motivation is the pursuit of wealth and power.

In order to understand the gravity of the issue that giving congresspeople financial advantages presents, a brief history of insider trading in American government and how access to nonpublic information has been abused by politicians in the past must be understood.

Just two years after the ratification of the Constitution of the United States and the establishment of the federal government, the American people got their first glance at politicians abusing their power for financial gain. In 1789, Alexander Hamilton, the first Secretary of the Treasury, led the charge of establishing the United States as a trustworthy country that could honor its debt [19]. To do this, Hamilton put forward a plan which many did not expect. Although the state-backed revolutionary war bonds had become almost worthless, Hamilton said the newly formed government would honor the country's debts at face value. When this plan was put forward to Congress, many congressmen bought up bonds from unknowing bondholders who were ready to receive a small fraction of the face value, financially gaining from their knowledge of this plan [19].

The practice of using nonpublic information for personal financial gain has been around since the beginning of the Union, even before the US stock market existed. So this begs the question: Why does this issue still exist today? Unfortunately preventing this kind of practice requires a law, and laws get passed by the very people committing these acts. As a result, it took more than 200 years after the event above for Congress to pass the "Stop Trading on Congressional Knowledge" Act of 2012, appropriately named and abbreviated as the STOCK Act [2]. This law came after an independent study from the Wall Street Journal showed that many senators' stock picks significantly outperformed the market gained national attention and legislators' hands were forced to increase scrutiny on their personal finances [18]. Most notably, the STOCK Act affirmed insider trading on nonpublic information as a criminal offense for members and employees of Congress [2]. In addition this law requires that members of Congress report any transaction they make that is larger than $10,000 within 45 days of its execution (this data will be the foundation of this thesis). Since the law was passed in 2012, there have been many violations of the law's mandatory reporting deadline and many reports that politicians have traded on nonpublic information [22]. However, no criminal charges have been brought against any of these politicians.

Chapter 2

Literature Review

Given the relative recentness of the STOCK Act, which required politicians to publicly disclose their financial transactions, the majority (but not all) of the research relating to politicians and their personal finances has been through investigative journalism which gets published much faster than academic work. As a result, in this section there is a brief overview of the existing literature, published in a wide spectrum of spaces, on politicians and their ethics regarding their personal finances.

2.1 Evaluations of Politicians' Personal Finances

One way that has been used to examine where politicians' true interest lies is by comparing what companies politicians invest in and seeing if these companies align with their political leanings. In 2017, Aiken and his team found that liberal politicians were more likely to invest in "socially responsible" stocks and more conservative politicians were less likely to invest in the same stocks [7]. This trading tendency aligns with political stances that the two dominant US political parties tend to hold on environmental, social and governance topics.

On the surface, this finding seems to indicate that politicians are campaigning in accordance with their true beliefs and that they are advocating for the best path forward for the country. However, this alignment of political and financial agendas can be seen as another compelling reason to investigate the finances of politicians. Instead of simply highlighting the positive that politicians' public beliefs align with their personal finances, this paper only further shows that congresspeople are acting not solely for the good of the people they represent, but for personal gain. Politicians are even more incentivized to try and pass legislation that aligns with their ideology because it will not only help their constituents and their likelihood of getting reelected, but it will increase the return on their personal investments.

2.2 Congresspeople's Returns on Public Equities

William Belmont and Bruce Sacerdote offer another way to evaluate the ethics of politicians, namely by evaluating the relative success of congresspeople's personal investments [8]. By examining the trading activity of congresspeople using data published through the reporting guidelines of the STOCK Act, the researchers examine the holdings of the congresspeople against the market as a whole. They find that, not only do House members and senators not beat the market, but their returns underperform the market in aggregate.

This study is perhaps counterintuitive and may on the surface seem to offer evidence that congresspeople are not insider trading, but this is not necessarily true. In order for an act to qualify as insider trading, it only requires the person trading the security possesses nonpublic information; it does not matter if the person makes a profit [33].

2.3 COVID-19 Trading Scandal

On March 20, 2020 many news outlets began reporting on an incident that suggested United States senators were trading on insider information they received in a classified briefing [20]. Within days of a classified briefing about the spread of COVID-19 in the United States, four senators sold hundreds of thousands of dollars of stocks that would have been negatively impacted by the pandemic. One senator even bought stock in companies that make personal protective equipment [30]. These transactions were made months in advance of when the virus spread dominated US news outlets and the public was made aware of the implications of COVID-19 on the economy. Although none of these senators were charged with their (blatant) violations of the STOCK Act, it is clear from this event that this matter of insider trading on nonpublic information in Congress warrants further investigation.

2.4 Regression Discontinuity Design

In order to examine how times when confidential information is disclosed impacts the trading patterns of members of the House of Representatives, it is necessary to turn to other disciplines that analyze data is similar ways. One common topic of research where time series data is scrutinized before and after a singular event is in epidemiology. Epidemiologists often evaluate the success of public health interventions by seeing how the number of new infections of a disease on a daily basis changes before and after their event. They do this using a method known as interrupted time series analysis which is also known as regression discontinuity design. James Lopez Bernal's paper titled, "Interrupted time series regression for the evaluation of public health interventions: a tutorial" provides a framework on a a useful technique to evaluate continuous time series data [9]. By splitting up the data into two distinct segments before and after some kind of intervention, the method explains how to quantitatively isolate the immediate impact of an event and the subsequent long-term trend. The insights from this paper are valuable in shaping the approach this thesis uses to examine the relationship between when nonpublic information is received in closed-door congressional meetings and how the trading activities of the representatives change as a result, even though the paper looks at the technique through the lens of public health interventions.

Some methods use ordinary least squares to model the time series data whereas other methods use Autoregressive Integrated Moving Average (ARIMA) models for the data. Both of these methods, although used for public health interventions in the paper, will be informative in understanding the effect of nonpublic information on the trading activity, if the days when these disclosures happen are known.

Chapter 3

Underlying Mathematics

3.1 General Approach

In order to paint a full picture of the relationship between when congresspeople receive confidential information and trade on that information, this thesis looks at the House of Representatives trading frequency before and after events when they may receive nonpublic information.

Each member in the House of Representatives is assigned committees to serve on. In these committees, representatives are briefed on information regarding national and international issues relating to the committee's oversight in order to be best prepared to present relevant legislation to the full Congress [29]. Committees allow certain members of Congress to become experts and fully understand an area or industry to introduce bills and resolutions [5]. Although most of the time when these committees meet they are open for the public to watch, sometimes committees hold "closed-door" meetings where confidential information is shared and the public is not disclosed the contents of the meetings. Congressional calendars on when these closed meetings occur and who the corresponding committee members are that have received confidential information are available online. The assumption is made that in these meetings congresspeople receive nonpublic information and the dates of these meetings are used in our analysis. Simultaneously, the trading activity from members of the House is examined. It is also assumed that congresspeople share information with other congresspeople about what happens in these meetings. For the period from 2015 to 2020 (the 114th, 115th and 116th meetings of Congress) the quantitative effect of these meetings on trading is investigated.

3.2 Mathematical Method

This thesis aims to model the effect of closed-door congressional committee meetings through a group of statistical methods to analyze the change in congressional trading activity before and after these events.

3.2.1 Two Sample T-Test

The first statistical test performed is a simple two sample t-test on the data surrounding the days when congresspeople receive nonpublic information. To do this, two samples will be constructed. The first sample is the difference between the number of trades between the day before and the day of a closed-door meeting. The second sample is the difference between the number of trades between the day before and the day of an open meeting. The following equation is used to calculate the t-statistic:

$$t = \frac{\bar{X}_1 - \bar{X}_2}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}} \tag{3.1}$$

$\bar{X}_1$: sample mean of difference for closed meetings
$\bar{X}_2$: sample mean of difference for open meetings
s_1^2, s_2^2 : sample variances
n_1, n_2 : sample sizes

The t-statistic informs on whether or not the dataset of differences of closed meetings has a significantly different mean from the dataset of differences of closed meetings. From the t-statistic a p-value is obtained which gives the probability that the two samples are different by an amount by chance alone. To see if the trading activity in the House of Representatives is different after times when politicians receive confidential information, a p-value is found from the following equation:

$$p = 2 \times (1 - \text{TCDF}(|t|))\,, \tag{3.2}$$

TCDF : the cumulative distribution function of the t-distribution

For reasons to be explained later, the two sample t-test may not be the most fitting or powerful test that can be performed in this situation. As a result, we will dive in further to our statistical catalog to try to find more evidence to make a stronger claim.

3.2.2 ARIMA Model

Another avenue for modeling time series data is to use an Auto-Regressive Integrated Moving Average (ARIMA) model. This model takes time series data and transforms it to make it stationary, meaning the series does not have a trend and stays around a constant mean with relative consistency [11]. The data can be looked at for the whole window and the model can be applied from 2015-2020. Our approach uses an ARIMA model to assess the impact of open and closed-door meetings on trading activity. To do this, indicator variables for the days of these meetings are used to capture their effect.

The ARIMA model is made up of two parts, an Autoregressive model (AR) and Moving Average model (MA). An AR model accounts for all previous values that a variable has taken on. This model can be written as such [11]:

$$y_t = c + \phi_1 y_{t-1} + \phi_2 y_{t-2} + \ldots + \phi_p y_{t-p} + \varepsilon_t$$

y_t : value of the time series at time t
ϕ : coefficients estimated from the data
c : constant
p : order of the AR model
ε : white noise error term

The MA model uses linear combination of past white noise error terms instead of the previous data points. This model can be written as such:

$$y_t = c + \varepsilon_t + \theta_1 \varepsilon_{t-1} + \theta_2 \varepsilon_{t-2} + \ldots + \theta_q \varepsilon_{t-q}$$

y_t : value of the time series at time t
θ : coefficients estimated from the data
c : constant
q : order of the MA model
ε : white noise error term

The AR and MA model can be combined to get the ARIMA model with parameters p, d, and q where d is the differencing component is used to transform a non-stationary time series into a stationary one [26]. Differencing is the the difference between consecutive data points in order to remove any trends in the data. The d value corresponds to how many differences are needed to achieve stationarity [11]:

$$y'_t = c + \phi_1 y'_{t-1} + \ldots + \phi_p y'_{t-p} + \theta_1 \varepsilon_{t-1} + \ldots + \theta_q \varepsilon_{t-q} + \varepsilon_t$$

y'_t : differenced time series at time t

Note: all other variables references in this equation are previously defined

The indicator variables for the days of open and closed-door meetings are then added to this traditional ARIMA model as so:

$$y'_t = c + \phi_1 y'_{t-1} + \ldots + \phi_p y'_{t-p} + \theta_1 \varepsilon_{t-1} + \ldots + \theta_q \varepsilon_{t-q} + \beta_1 X_{1t} + \beta_2 X_{2t} + \varepsilon_t \quad (3.3)$$

y'_t : the differenced time series at time t
X_{1t}, X_{2t} : indicator variables for open and closed-door meetings at time t
β_1, β_2 : coefficients capturing the impact of the meetings

By adding X_{1t} and X_{2t} into the ARIMA model, the model will now be able to clearly show the effects of open and closed door meeting on trading.

In addition to the ARIMA model above, a slight change can be made where the indicator variables are lagged one day in order to allow for the analysis of delayed effects of open and closed-door meetings. Here is the slight modification of the equation (note the different subscripts on the indicator variables):

$$y'_t = c + \phi_1 y'_{t-1} + \ldots + \phi_p y'_{t-p} + \theta_1 \varepsilon_{t-1} + \ldots + \theta_q \varepsilon_{t-q} + \beta_1 X_{1t-1} + \beta_2 X_{2t-1} + \varepsilon_t \quad (3.4)$$

In order to find the optimal parameters for p, d, and q in the ARIMA model, a technique called grid search is used. In grid search, all possible combinations of parameters are tested to see which one is optimal for the model. This study judges optimality of a model by Akaike Information Criterion (AIC). AIC measures the accuracy of a model by balancing the fit of a model and its complexity [12].

$$\text{AIC} = 2k - 2\ln(L), \tag{3.5}$$

L : the likelihood of the model
k : the number of parameters that are estimated

The results from the optimal ARIMA models will inform about the individual significance of the coefficients for the indicator variables on open and closed-door meetings through their z-statistics and p-values. The z-statistic is calculated like this:

$$Z = \frac{\hat{\beta}}{SE(\hat{\beta})}, \tag{3.6}$$

$\hat{\beta}$: coefficient for the indicator variable,
$SE(\hat{\beta})$: standard error of the coefficient

The p-value for the z-statistic is calculated from the standard normal distribution as so:

$$p_Z = 2 \times (1 - \Phi(|Z|)), \tag{3.7}$$

Z : Z-statistic,
Φ : cumulative distribution function of the standard normal distribution

In addition it will be informative to look at the collective impact of the variables through an f-test. An f-test can inform about whether or not the the ARIMA model is better at predicting the trading activity when both exogenous variables are included. The f-statistic is calculated as so [35]:

$$F = \frac{(\mathrm{RSS_R} - \mathrm{RSS_U})/q}{\mathrm{RSS_U}/(n-k-1)}, \tag{3.8}$$

RSS_R : residual sum of squares of the restricted model (without the exogenous variables)
RSS_U : residual sum of squares of the unrestricted model (with the exogenous variables)
q : number of exogenous variables
n : number of observations
k : number of parameters in the unrestricted model

The p-values for the f-statistic is calculated using the f distribution as so:

$$p_F = 1 - F\mathrm{CDF}(F), \tag{3.9}$$

F : F-statistic
FCDF : the cumulative distribution function of the F-distribution

Through this method, the effect of open and closed door meetings on trading activity can be isolated and analyzed.

3.2.3 Regression Discontinuity Design

Regression discontinuity design (RDD) is a method of looking at how time series data changes after the onset of an event [21]. In the context of this thesis, the time series data is the daily frequency of trades made by members of the House of Representatives and the cutoff event is the date of the meeting in Congress when members of a committee have received nonpublic information. Regression discontinuity design is an effective method of testing causal hypotheses like in this situation because unlike a traditional experiment a random assignment of treatment is not possible [13]. This method is most commonly used in the fields of statistics, econometrics, and political science and in a way our research looks at the intersection of these disciplines.

Implementation

Instead of just looking at the change in means before and after the cutoff (like in the t-test above), regression discontinuity design gives a more digestible result of statistical difference in trading immediately after the event and the continued effect of the new information [15].

For the the situation at hand, the change in the intercept and the slope at the time of "intervention" will be examined to inform what kind of changes in trading behavior are occurring. Changes in the intercept indicate that there is a difference in the trading activity immediately after the event (level change) and changes in the slope means that that there is a difference in the rate of trading activity (trend change) [9]. Although both of these results will be looked at, the level change will be the most obvious indicator of a jump in trading after the closed-door meeting, as most meetings would not cause any long term changes in trading behavior based on one meeting. Here is the equation used to model the situation:

$$Y_t = \beta_0 + \beta_1 T + \beta_2 X_t + \beta_3 T X_t \tag{3.10}$$

Y_t : outcome at time t
X_t : indicator variable that takes 0 for pre-event data and 1 for post-event data
T : time elapsed since the start of the analysis

The level change is represented by the coefficient β_2, and a t-test is performed on it to determine its statistical significance. The t-statistic is calculated as follows:

$$t = \frac{\hat{\beta}_2}{SE(\hat{\beta}_2)} \tag{3.11}$$

t : t-statistic for the level change
β_2 : coefficient associated with the X_t term
$SE(\beta_2)$: standard error of the coefficient β_2

The degrees of freedom for this t-statistic depend on the sample size, which is how many days are included in the window.

The same process for calculating the t-test is used on the coefficient β_3 to calculate statistical significance of the trend change. Once the t-statistics are obtained, the p-values are calculated in the same way as in the two sample t-test.

Finally the results from the House of Representatives trading activity will be compared to that of the total market volume to account for larger trading trends in the market.

Example RDD Analysis

In order to give a complete understanding of how regression discontinuity design is used in this thesis, a singular event is highlighted to show the methodology of the tests being performed.

For visualization's sake the complete trading activity in the House of Representatives from 2015 to 2020, which is the length of the full dataset we will be using, is shown below. The red dotted lines represent dates when there were confidential hearings in a committee in the House.

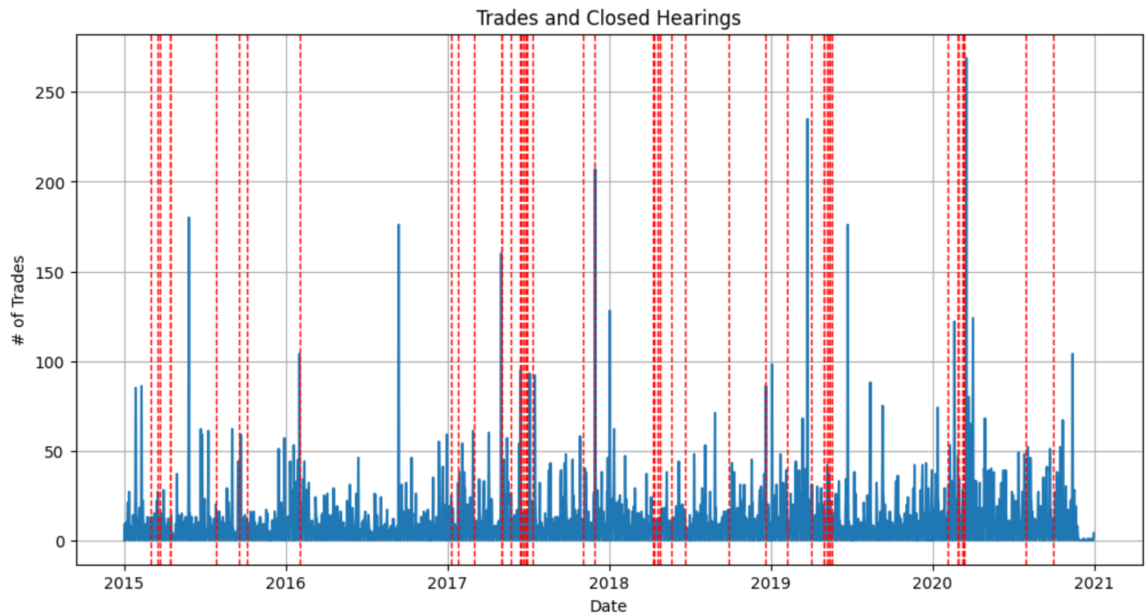

Figure 3.1: Trades by Members of the House of Representatives from 2015-2020

To perform the analysis, each of these events are zoomed in on and the trading activity among representatives in the House before and after these meetings is observed. For simplicity and conciseness, this section only walks through the methodology for one of the events and then a summary of all the events is published in the results section.

Let's start by looking at the days surrounding February 6, 2020. On this day, the House Appropriations Committee held a closed-door meeting where information about what was discussed was not shared to the public. To perform the analysis the window of five days on either day surrounding the closed-door hearings in the House of Representatives are observed. This choice of five days is an assumption about when the change in trading activity will be able to be noticed. Later other windows will be examined to see how mathematical results may change. Let's first start by looking at the trading activity in the five (trading) days before and after this date:

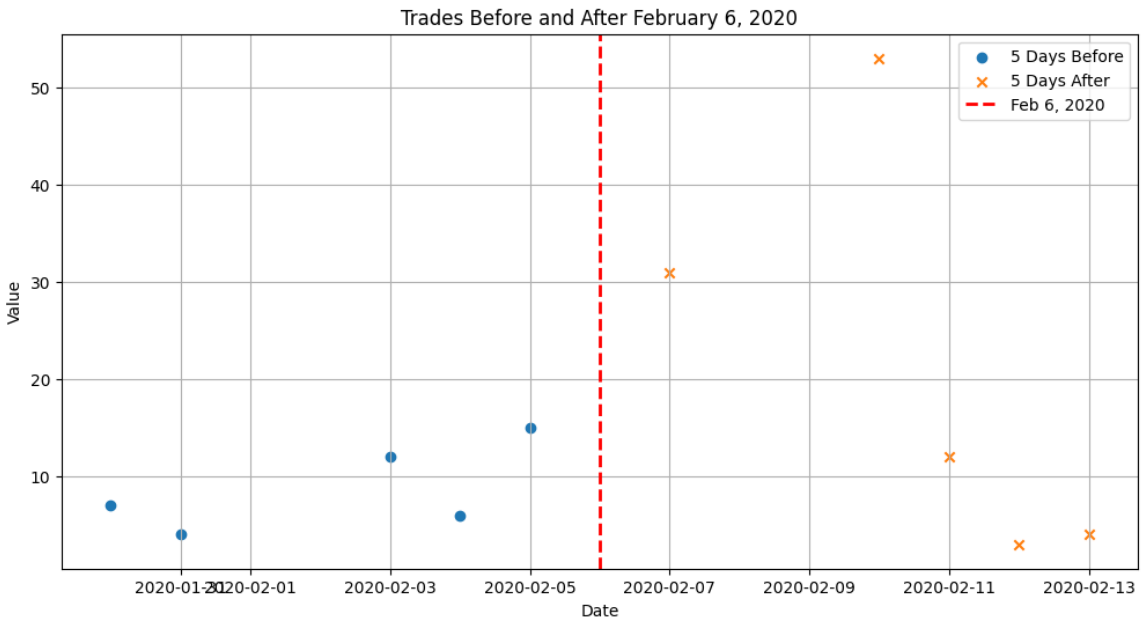

Figure 3.2: Trades Around February 6, 2020

Next regression discontinuity design is employed. The window of 10 trading days are linearly modeled to see if there is a jump in the number of trades after the event on February 6, 2020. Here is the graph:

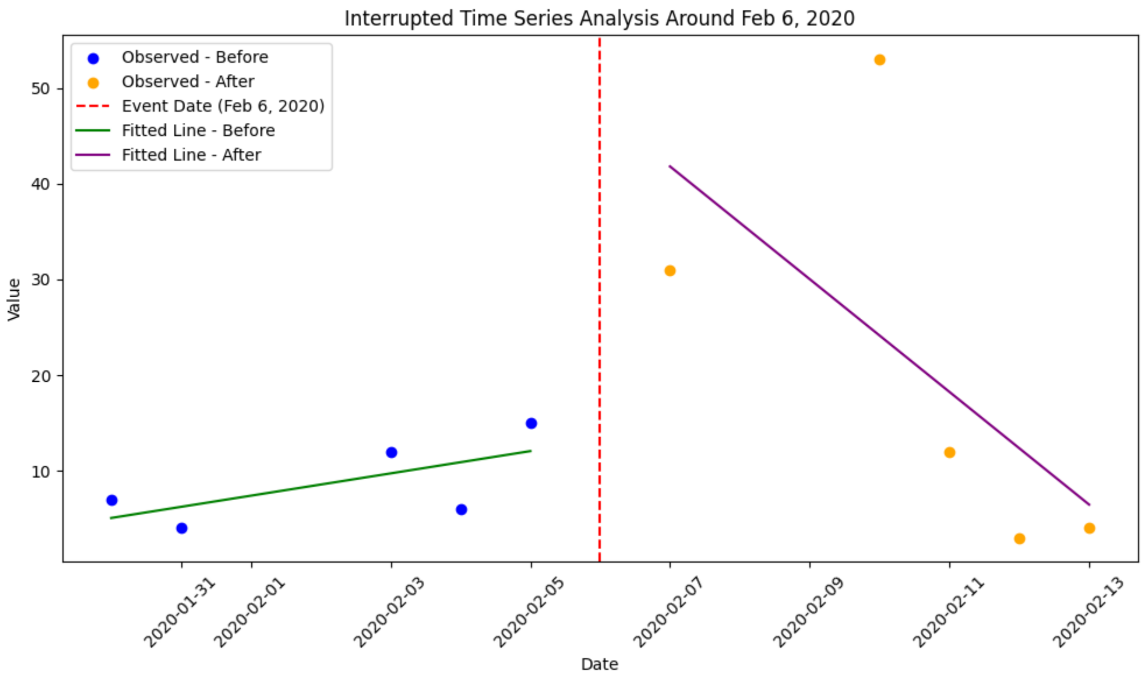

Figure 3.3: RDD Around February 6, 2020

Here are the results for the level change and trend change coefficients from the least squares regression:

Meeting Date	**Level Change**	**P-Value**	$\alpha = 0.05$	$\alpha = 0.10$
February 6, 2020	28.18	0.2298		

Table 3.1: Level Change and P-Value Summary 2/6/2020

Meeting Date	**Trend Change**	**P-Value**	$\alpha = 0.05$	$\alpha = 0.10$
February 6, 2020	-1.95	0.41698		

Table 3.2: Trend Change and P-Value Summary 2/6/2020

These two tables indicate that there was an immediate increase of 28.18 trades after the meeting on February 6, 2020 and that the rate of trading decreased by -1.95 trades/day. However, as seen by the p-values of 0.2298 and 0.41698, neither of these changes are statistically significant at $\alpha = 0.05$ or $\alpha = 0.10$, so no statistically powerful conclusions can be drawn.

In addition to this result it is useful to see how trading activity was in the market at large during this time. Here is the result of the level change for the five days before and after February 6, 2020 for the market at large:

Meeting Date	**Level Change**	**P-Value**	$\alpha = 0.10$	$\alpha = 0.15$
February 6, 2020	-0.281	0.482		

Table 3.3: Market Volume (in billions) 2/6/2020

There was no significant level change in the total trading activity around this date.

Chapter 4

Data

4.1 Congressional Calendars

The first key dataset in this study is a compilation of all the closed-door and open committee meetings in the House of Representatives over the relevant time period. To get this, CSV files are downloaded from congress.gov that contain information on every committee hearing and its relevant date. From here the CSV files are converted to a single excel file for ease of use. Once in an excel document, the committee meetings are filtered by name; only meetings that indicated they were "closed" in the title are kept in the dataset for closed meetings. From this pared down list the meetings are manually checked to make sure the remaining meetings were all closed-door meetings. Now the closed-door data set is usable. For visualization's sake a table of all the closed-door meetings in committees in the House of Representatives from 2015 to 2020 is summarized here:

Meeting Title	Date	Committee
Ongoing Intelligence Activities	March 2, 2015	Intelligence
Ongoing Intelligence Activities	March 17, 2015	Intelligence
Ongoing Intelligence Activities	March 23, 2015	Intelligence
CIA Budget	April 14, 2015	Intelligence
Special Activities	April 15, 2015	Intelligence
Iran	July 28, 2015	Intelligence
Defense Intelligence Agency	September 18, 2015	Intelligence
Access Request	October 6, 2015	Intelligence
Budget Views & Estimates	February 1, 2016	Intelligence
Business Meeting	January 10, 2017	Intelligence
Ongoing Intelligence Activities	January 24, 2017	Intelligence
Budget Views & Estimates Letter	March 2, 2017	Intelligence
Business Meeting: Access Request	May 2, 2017	Intelligence
Ongoing Intelligence Activities	May 4, 2017	Intelligence
Russia Investigation Hearing	May 23, 2017	Intelligence
Ongoing Intelligence Activities	June 14, 2017	Intelligence
Ongoing Intelligence Activities	June 15, 2017	Intelligence
Ongoing Intelligence Activities	June 21, 2017	Intelligence
Ongoing Intelligence Activities	June 22, 2017	Intelligence
Ongoing Intelligence Activities	June 27, 2017	Intelligence
Ongoing Intelligence Activities	June 28, 2017	Intelligence
Ongoing Intelligence Activity	June 29, 2017	Intelligence
Intelligence Authorization Act	July 13, 2017	Intelligence
Full Committee Hearing	November 2, 2017	Intelligence
Full Committee Hearing	November 30, 2017	Intelligence
U.S. Strategic Command	April 11, 2018	Appropriations
Budget Hearing	April 12, 2018	Intelligence
Budget Hearing 2	April 19, 2018	Intelligence
Department of Defense Budget	April 25, 2018	Appropriations
Budget Hearing	April 26, 2018	Intelligence
Ongoing Intelligence Activities	May 22, 2018	Intelligence
DIA Roles & Mission	June 21, 2018	Intelligence
Business Meeting	September 28, 2018	Intelligence
Business Meeting	December 20, 2018	Intelligence

Table 4.1: Closed-Door Meetings 2015-2018

Continued from previous page

Meeting Title	Date	Committee
Organizational/Business Meeting	February 6, 2019	Intelligence
Budget Request	April 3, 2019	Intelligence
Budget Request	May 1, 2019	Intelligence
CIA Program Budget Request	May 2, 2019	Intelligence
Budget Request	May 8, 2019	Intelligence
Budget Request	May 9, 2019	Intelligence
Budget Request	May 14, 2019	Intelligence
Defense Subcommittee Markup	May 15, 2019	Appropriations
Full Committee Business Meeting	May 20, 2019	Intelligence
U.S. Strategic Command	February 6, 2020	Appropriations
U.S. European Command	February 27, 2020	Appropriations
World-Wide Threat	February 27, 2020	Appropriations
World-Wide Threat	February 27, 2020	Appropriations
NGA Budget Hearing	March 10, 2020	Intelligence
U.S. Southern Command	March 10, 2020	Appropriations
NSA Budget Hearing	March 11, 2020	Intelligence
U.S. Central Command	March 11, 2020	Appropriations
U.S. Africa Command	March 11, 2020	Appropriations
FBI Budget Hearing	March 12, 2020	Intelligence
Intelligence Authorization Act	July 31, 2020	Intelligence
Business Meeting	September 30, 2020	Intelligence

Table 4.2: Closed-Door Meetings 2019-2020

4.2 Trading

4.2.1 Obtaining Data

Obtaining the trading data was not as straightforward as simply downloading it from a government website. Despite one of the main goals of the STOCK Act being to increase transparency of public officials personal finances, just one year after the bill was initially passed, Congress amended the STOCK Act in 2013 and restricted the mandate for online publication of financial disclosures. This amendment removed the stipulation that the system must facilitate searching, sorting, and downloading of data that was in this report [32]. This change made it more challenging to aggregate and survey easily usable data on congressional trading. Fortunately, other research has been done in recent years that uses similar data where people have already manually consolidated trades in usable formats. One example of this research is a study titled: “Do senators and house members beat the stock market? Evidence from the STOCK Act” which was mentioned previously in the literature review [8]. In our search for reliable dataset we reached out to Dr. Sacerdote and he was generous enough to allow us to use the dataset that he and his team constructed. His team’s meticulous, manual consolidation of trading activity in the House of Representatives as well as his standing in the academic community inspire confidence that the data is trustworthy and properly preprocessed. The dataset contains every trade of public equities made by congresspeople, their spouses, and their dependents from January 2012 to December 2020 and will be the used to create time series for the frequency of trades in Congress in our study.

4.2.2 Cleaning Data

Because the dataset on closed-door meetings only contains information from 2015-2020 in the House, the trading data is trimmed to align the windows of data and only includes trades by representatives. Additionally the intent of this research is not to single out individual politicians in order to provide evidence that they are insider trading; instead this study is looking for trends across the House of Representatives as a whole to see if any widespread phenomenon in trading are present. As a result, the daily trades are aggregated and the names connected with each trade are removed from the dataset as well.

4.3 Market Volume

In order to control for the market activity and the fluctuation that occurs in trading, the market volume of the Standard and Poor's 500 (S&P 500) will be simultaneously analyzed with the trades of congresspeople. The S&P 500 is a stock market index that is made up of the 500 largest companies that are listed on stock exchanges in the United States [4]. In order to get the market volume of this data, it was downloaded from Yahoo Finance [36]. On Yahoo Finance the date range is set to be January 1, 2015 to December 31, 2020 and the historical data on market volume is downloaded.

Chapter 5

Results

5.1 Two Sample T-Test

The first test performed is a two sample t-test where the first sample is the difference between trades on the day before to the day of closed-door meetings and the second sample is the difference between trades on the day before to the day of open meetings. Here is the result of the two sample t-test:

T-Statistic	**P-Value**	$\alpha = 0.05$	$\alpha = 0.10$
2.604	0.0114	✓	✓

Table 5.1: T-Statistic and P-Value Summary

There is a statistically significant result that the differences in trading values associated with closed-door meetings has a higher mean than the differences for open-door meetings.

5.2 ARIMA Models

To look at the trends of trading activity around closed door and open door meetings at large, the entire dataset of trading days is looked at. The ARIMA model is then used to model the time series data.

5.2.1 Non-Lagged Indicators

First the dataset of trades is modeled where the exogenous indicator variables for the open and closed-door meetings is on the day of the meetings. Here is the plot and resultant table of results:

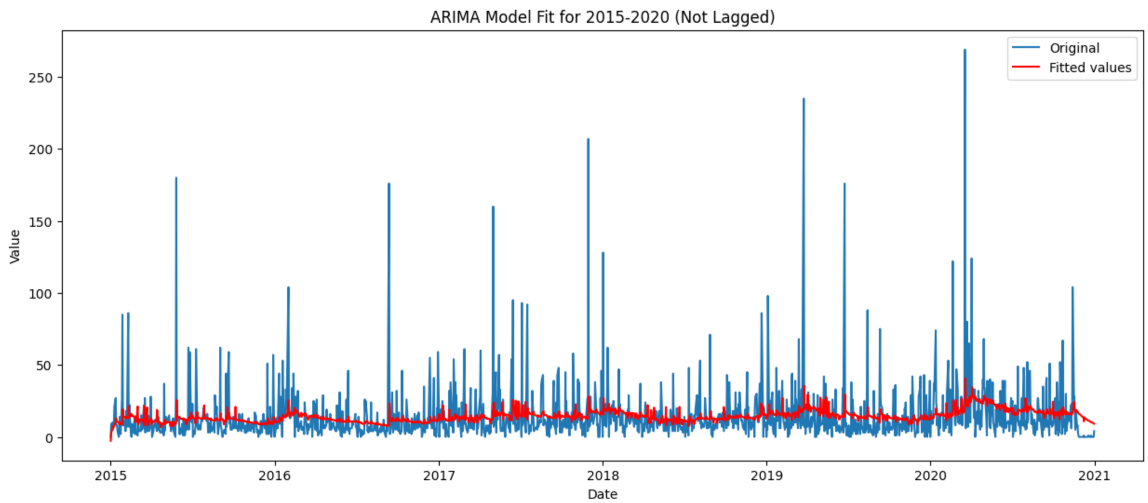

Figure 5.1: ARIMA Model (Non-Lagged)

Table 5.2: ARIMA Results (Non-Lagged)

Model:	ARIMA(1, 1, 1)					
	coef	std err	z	P> |z|	[0.025	0.975]
OpenMeeting	-2.7073	2.220	-1.219	0.223	-7.059	1.644
ClosedMeeting	8.3572	1.420	5.886	0.000	5.574	11.140
ar.L1	0.0679	0.027	2.509	0.012	0.015	0.121
ma.L1	-0.9768	0.006	-174.263	0.000	-0.988	-0.966
sigma2	358.6266	2.972	120.683	0.000	352.802	364.451

F-Statistic	**P-Value**	$\alpha = 0.05$	$\alpha = 0.10$
6.821	0.00112	✓	✓

Table 5.3: F-Test that Coefficients for Meeting Types are Zero (Non-Lagged)

The coefficient for the closed-door meeting days is 8.35 and statistically significant (the p value is less than 0.05). This suggests that on the day of the closed door meeting that the trading activity increases compared to other days. The coefficient for open door meetings is not statistically significant which suggests that there is no clear impact on trading on the days these public meetings occur. Additionally given the statistically significant f-statistic, this suggests that the inclusion of these exogenous variables improves the model's ability to explain the variations in the trading data. In other words, we reject the hypothesis that both the open and closed-door meeting terms should have a coefficient of 0. More specifically, closed-door meetings significantly affect trading activity on that day.

5.2.2 Lagged Indicators

Now the indicator variables are lagged one day after the day the meetings occur to see if there is a significant effect when there is an additional day allowed for congresspeople to make trades.

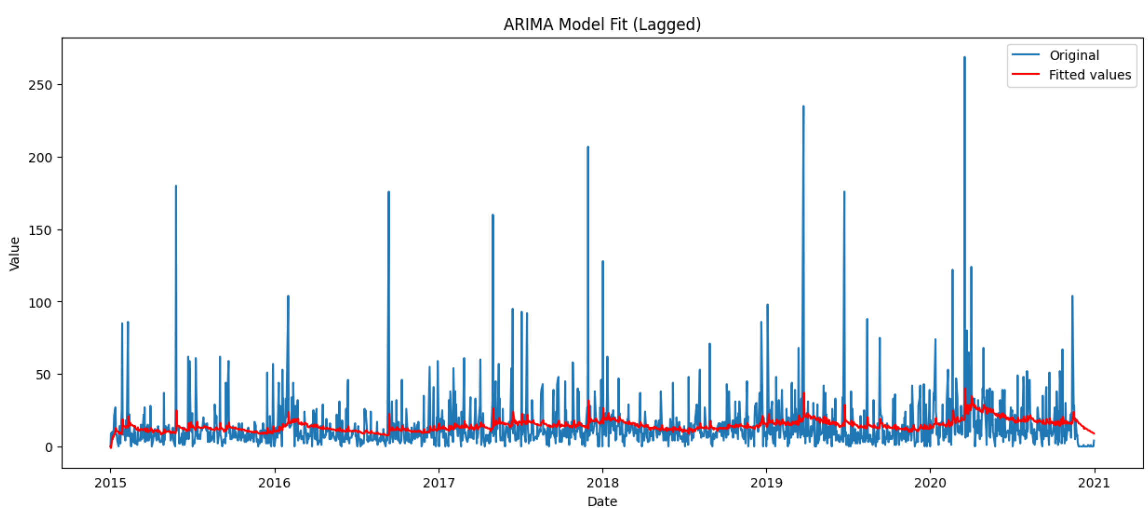

Figure 5.2: ARIMA Model (Lagged)

Table 5.4: ARIMA Results (Lagged)

Model:	ARIMA(1, 1, 1)					
	coef	std err	z	P> |z|	[0.025	0.975]
LaggedOpenMeeting	-1.1669	1.466	-0.796	0.426	-4.040	1.706
LaggedClosedMeeting	0.9651	2.751	0.351	0.726	-4.427	6.357
ar.L1	0.0623	0.027	2.320	0.020	0.010	0.115
ma.L1	-0.9758	0.006	-175.212	0.000	-0.987	-0.965
sigma2	361.5433	2.683	134.758	0.000	356.285	366.802

F-Statistic	**P-Value**	$\alpha = 0.05$	$\alpha = 0.10$
0.385	0.680		

Table 5.5: F-Test that Coefficients for Meeting Types are Zero (Lagged)

Neither the open or closed door meetings have statistically significant coefficient when the indicator variables are lagged one day. This suggests that the impact that these meetings have on trading may be only felt on the day of the closed-door meetings and is not long lasting. Here the f-statistic is not statistically significant; likewise, including one-day lagged indicators for both open and closed-door meetings does not improve the model's ability to explain the variations in trading activity.

5.3 Regression Discontinuity Design

The final test performed is an analysis using regression discontinuity design (RDD). The level change between a regression performed on the the five trading days before and after congressional meetings is examined first.

5.3.1 10-Day Window

Closed-Door Meetings

Here are the results from the closed-door meetings that had statistically significant level changes on representatives' trading activity. *Asterisks next to checkmarks entail a negative level change.*

Event Date	**Level Change**	**P-Value**	$\alpha = 0.10$	$\alpha = 0.15$
March 2, 2015	13.000	0.070	✓	✓
May 23, 2017	22.760	0.092	✓	✓
June 15, 2017	-39.500	0.121		✓*
April 26, 2018	8.627	0.101		✓
February 6, 2019	-27.507	0.075	✓*	✓*
February 27, 2020	77.955	0.057	✓	✓

Table 5.6: Level Change for RDD (Closed Meetings)

Of the 50 closed-door meetings, 6 had statistically significant level changes between the regression on the trading days before and after the closed-door meetings. Of the 6 statistically significant results, 2 of the level changes were negative, indicating a statistically significant decrease in the trading activity after these meetings.

Now to check for confounding factors in the market at large, here are the results of performing the same analysis on total trading activity, but using market volume of the S&P 500 instead of representative's trades on the dates above.

Event Date	**Level Change**	**P-Value**	$\alpha = 0.10$	$\alpha = 0.15$
March 2, 2015	-0.352	0.371		
May 23, 2017	-0.031	0.960		
June 15, 2017	1.490	0.074	✓	✓
April 26, 2018	-0.295	0.413		
February 6, 2019	0.624	0.303		
February 27, 2020	2.286	0.191		

Table 5.7: Level Change for RDD (Market Volume)

None of the days when there was a positive level change in congressional trading activity also had a positive change in market volume suggesting that the increase trading activity was not seen with the market at large.

Open Meetings

Now the open door congressional committee meetings will also be examined using the same methodology. Here are the results from the statistically significant level changes in open meetings:

Event Date	**Level Change**	**P-Value**	$\alpha = 0.10$	$\alpha = 0.15$
April 3, 2015	-18.308	0.127		✓*
April 18, 2015	-10.150	0.070	✓*	✓*
July 9, 2015	55.130	0.023	✓	✓
February 23, 2016	-22.150	0.084	✓*	✓*
January 1, 2017	-63.400	0.025	✓*	✓*
March 21, 2017	25.500	0.055	✓	✓
March 22, 2017	25.269	0.052	✓	✓
April 4, 2017	-20.245	0.110		✓*
April 5, 2017	-53.545	0.015	✓*	✓*
May 3, 2017	-120.761	0.060	✓*	✓*
May 17, 2017	-36.836	0.082	✓*	✓*
May 18, 2017	-43.190	0.049	✓*	✓*
July 17, 2017	27.262	0.030	✓	✓
July 18, 2017	-63.270	0.108		✓*
July 19, 2017	-60.739	0.110		✓*
September 14, 2017	-22.050	0.130		✓*
December 1, 2017	-164.900	0.120		✓*
February 5, 2018	42.200	0.039	✓	✓
March 1, 2018	-16.060	0.067	✓*	✓*
April 17, 2018	8.025	0.114		✓
May 10, 2018	22.805	0.105		✓
June 6, 2018	26.209	0.090	✓	✓
February 7, 2019	-31.200	0.038	✓*	✓*
February 26, 2019	29.060	0.086	✓	✓
February 28, 2019	-40.435	0.072	✓*	✓*
March 13, 2019	46.806	0.073	✓	✓
March 26, 2019	-197.565	0.038	✓*	✓*
September 10, 2019	58.700	0.033	✓	✓
June 3, 2019	-30.158	0.016	✓*	✓*

Table 5.8: Level Change for RDD (Open Meetings)

Continued from previous page

Event Date	Level Change	P-Value	$\alpha = 0.10$	$\alpha = 0.15$
June 11, 2019	6.667	0.086	✓	✓
June 13, 2019	-6.000	0.114		✓*
September 26, 2019	11.790	0.064	✓	✓
February 11, 2020	-43.445	0.071	✓*	✓*
March 3, 2020	-35.465	0.051	✓*	✓*
July 15, 2020	-29.328	0.110		✓*
July 29, 2020	-45.701	0.067	✓*	✓*

Table 5.9: Level Change for RDD (Open Meetings) - Part 2

Of the 205 open meetings that took place between 2015 and 2020, there was a statically significant level change in 37 of the meetings. Of those 37 meetings, 24 of the meetings have a negative level change, indicating a decrease in the trading activity after those meetings.

5.3.2 20-Day Window

To test to see if there is a sustained effect of the meetings on trading activity the window of RDD analysis is expanded to include the 10 trading days before and after meetings

Closed Meetings

Event Date	Level Change	P-Value	$\alpha = 0.10$	$\alpha = 0.15$
March 02, 2017	-19.947	0.139		✓*
May 4, 2017	-49.823	0.147		✓*
June 15, 2017	-26.838	0.027	✓*	✓*
June 21, 2017	-37.951	0.161		✓*
December 20, 2018	26.221	0.124		✓
May 8, 2019	17.478	0.109		✓
May 14, 2019	-20.488	0.056	✓*	✓*
May 15, 2019	-21.826	0.038	✓*	✓*

Table 5.10: Level Change for RDD (Closed Meetings, 20-Day Window)

When the regression discontinuity design analysis is performed on the ten days before and after closed door meetings, there is a statistically significant change in trading activity after 8 meetings. However 6 of these events are associated are a significant decrease. This result of a majority decreasing is similar to what we've seen for open meetings for the smaller window.

Open Meetings

The same RDD model is performed for the ten days before and after open meetings. Here are the statistically significant results:

Event Date	**Level Change**	**P-Value**	$\alpha = 0.10$	$\alpha = 0.15$
February 11, 2015	-25.253	0.143		✓*
February 13, 2015	-28.493	0.114		✓*
March 18, 2015	-9.276	0.136		✓*
March 19, 2015	-13.937	0.050	✓*	✓*
July 8, 2015	23.570	0.136		✓
July 9, 2015	33.089	0.029	✓	✓
February 3, 2016	-37.398	0.065	✓*	✓*
March 14, 2016	-14.687	0.023	✓*	✓*
February 28, 2017	-25.281	0.056	✓*	✓*
March 1, 2017	-21.345	0.108		✓*
April 5, 2017	-24.099	0.066	✓*	✓*
May 3, 2017	-68.421	0.034	✓*	✓*
May 17, 2017	-27.571	0.037	✓*	✓*
May 18, 2017	-30.067	0.025	✓*	✓*
June 12, 2017	39.007	0.061	✓	✓
June 20, 2017	-37.258	0.081	✓*	✓*
September 28, 2017	-22.677	0.126		✓*
November 15, 2017	-17.238	0.113		✓*
November 29, 2017	70.493	0.097	✓	✓
December 1, 2017	-83.927	0.058	✓*	✓*
February 5, 2018	26.902	0.004	✓	✓
February 26, 2018	11.207	0.059	✓	✓
March 1, 2018	-10.187	0.072	✓*	✓*

Table 5.11: Level Change for RDD (Open Meetings, 20-Day Window)

Continued from previous page

Event Date	**Level Change**	**P-Value**	$\alpha = 0.10$	$\alpha = 0.15$
May 9, 2018	10.887	0.120		✓
May 10, 2018	10.969	0.120		✓
May 15, 2018	-12.061	0.097	✓*	✓*
May 16, 2018	-11.947	0.096	✓*	✓*
June 6, 2018	15.234	0.075	✓	✓
July 11, 2018	17.575	0.057	✓	✓
February 7, 2019	-16.812	0.092	✓*	✓*
February 26, 2019	21.952	0.044	✓	✓
March 26, 2019	-102.615	0.033	✓*	✓*
March 27, 2019	-73.710	0.128		✓*
May 7, 2019	20.951	0.062	✓	✓
May 16, 2019	-20.010	0.062	✓*	✓*
June 5, 2019	-16.007	0.036	✓*	✓*
September 10, 2019	43.308	0.005	✓	✓
September 17, 2019	-32.604	0.042	✓*	✓*
October 16, 2019	-16.535	0.074	✓*	✓*
November 15, 2019	16.557	0.120		✓
December 3, 2019	20.983	0.116		✓
July 15, 2020	-22.127	0.104		✓*
July 29, 2020	-28.886	0.042	✓*	✓*

Table 5.12: Level Change for RDD (Open Meetings, 20-Day Window) - Part 2

Of the 205 open door meetings that took place between 2015 and 2020, there was a statically significant level change in 43 of the meetings. Of those 43 meetings, 28 of the meetings have a negative level change, indicating a decrease in the trading activity after those meetings. This is a similar pattern as seen in the open meeting RDD treatment for five days before and five days after the meetings.

Chapter 6

Discussion

6.1 Two Sample T-Test

The two sample t-test on the differences between the one day before and the day of closed door meetings shows that the mean of the change in trades is higher on days with closed-door meetings than on days with open meetings. This test provides evidence to support the claim that members of the House of Representatives are trading differently based on the the type of information they receive in the closed-door meeting and that they are more likely to increase their trading on days when these meetings occur. This is problematic as the STOCK Act prohibits using non-public information that representatives obtain through their positions for personal benefit [2].

In order to perform a two sample t-test, the assumption is made that the data in the two datasets (before and after the event) are independent of each other must be made [24]. This is not necessarily true in this case. Furthermore, the t-test also requires that the random variables from each population (in this case the trading activities on days surrounding different types of meetings) are identically distributed, which is not necessarily true. Trading volume on different days can be affected by a number of different factors like individual investor behavior. Considering these factors, it becomes challenging to confidently apply these assumptions to trading activity data. As a result of these weak assumptions the two sample t-test results are not as robust as we want for our analysis.

Another assumption that is made is that the data is normally distributed. In statistics the rule of thumb for assuming normality is that the sampling distribution of the mean must be constructed of 30 or more observations. For the open meetings dataset, the sample size of 204 greatly exceeds that number, but for closed meetings the data set is 50 points, which only slightly exceeds the threshold. Given the size of the closed-door sample it may also be more sensitive to outliers which is not uncommon when looking at trading data.

Practically, this test is limiting because it does not allow us to see the delayed effects on trading that may be present from these meetings. In other words, trading activity may continued to be impacted in the few days after the meeting, which is not accounted for in this t-test. To compensate for this the ARIMA model and the method in regression discontinuity design will also look at a longer period following the days of the open and closed door meetings.

6.2 ARIMA Models

Looking at the two ARIMA models, one that includes non-lagged indicator variables for meetings (Figure 5.1) and one that has indicator variables that are lagged by one day (Figure 5.2), we can see that the trend of the trading activity being affected by the type of meeting continues. The statistically significant positive coefficient for closed-door meetings in the non-lagged ARIMA model shows that there is an immediate impact seen on the day of closed door meetings on trading (Table 5.2). This finding offers indirect support to the notion that the information provided to congresspeople in closed-door meetings prompts an increase in trading that is not typically observed in all meeting types, like those where the information is publicly available.

When we compare this result to the lagged ARIMA indicators neither open or closed-door meetings show a statistically significant impact on trading activity the day after the meetings (Table 5.4). This suggests that the effect that the meetings have on trading is temporary, only significantly occurring on the day of the closed-door meeting. This result, in conjunction with the result of the non-lagged indicators, paints a fuller picture than the t-test alone about how this phenomenon in trading activity based on meeting type manifests.

6.3 Regression Discontinuity Design

6.3.1 Closed-Door Meetings, 10-Day Window

The results from the regression discontinuity design method of this thesis suggest that certain closed-door meetings have a statistically significant increase in trading activity from before the meeting to the after, but that there is not a dominant trend in trading activity changes across all closed-door meetings. To thoroughly analyze this result, we will explore the broader context of the meetings where a statistically significant increase in trading activity is observed.

March 2, 2015

From the five days before to the five days after March 2, 2015, there is a statistically significant increase in trading activity. The title of the closed-door meeting on that day is "Ongoing Intelligence Activities." Upon further inspection there is only one major world event surrounding that day in history that pertains to the US Intelligence committee: an ISIL (also known as ISIS) attack that destroyed three major cities in Iraq on March 5-8, 2015 [23]. Looking back on the original event, the largest transaction made in the five days after March 2, 2015 was a transaction on March 3, 2015 where a representative bought shares of Targa Resources (NYSE: NGLS) which is a one of the largest natural gas and natural gas liquids companies in the United States. Following a logical progression, this assault would negatively impact Iraq's primary export, crude oil, leading to an uptick in the price of US energy stocks like NGLS [17]. Although there is no definitive evidence of insider trading, this is an interesting phenomenon and when looked at with the level change result there are some indicators of foul play.

May 23, 2017

After May 23, 2017 there is a statistically significant increase of more than 22 trades. The title of the meeting that took place on this day is "Russia Investigation Hearing" and it comes six days after the Justice Department appointed a special counsel to oversee the investigation into Russian interference in the 2016 election. We can infer that this meeting was related to this news. In the five trading days following May 23rd there is over $35,000 worth of recorded transactions in stocks related to the energy industry. Russia's four largest exports are crude petroleum, petroleum gas, and refined petroleum [1]. If this meeting suggested that the (trade) relations between the United States and Russia would be greatly impacted by the investigation or results of the investigation, energy stocks would be the the most likely avenue of insider trading. Similar to the trading activity around March 2, 2015, there is a somewhat suspicious increase in trading of energy stocks which again provides more circumstantial evidence of insider trading.

February 27, 2020

The event on February 27, 2020 has the largest level change of any of the statistically significant events for a closed-door meeting, a jump of more than 77 trades, and perhaps it is the most likely circumstance of insider trading. The title of the meeting on this day is "World-Wide Threat." This meeting took place just weeks before the United States declared COVID-19 a nationwide emergency. As the general public continued to attend school and work as usual, members of the House of Representatives were trading stocks frantically, with a preponderance of these trades taking place in the healthcare and pharmaceutical industries. Examples of stocks traded by representatives in the five days after February 27, 2020 include Pfizer (NYSE: PFE), Alexion Pharmaceuticals (NYSE: ALXN), and Abbott Laboratories (NYSE: ABT), which would all have volatile stock prices during a pandemic and likewise have the greatest profit potential. This event is similar to the event described in the literature review where there was a jump in trading activity in these industries among senators who attended a confidential briefing related to the COVID-19 pandemic.

6.3.2 Open Meetings, 10-Day Window

For open meetings there is a higher percentage of statistically significant differences in the level change in the five days before the meeting and the five days after the meeting. This represents a statistically significant difference in 18% of meetings, as opposed to 12% statistical significance in the closed meetings. Interestingly, a healthy majority of statistical differences are negative, indicating that congresspeople decreasing their trading activity after these meetings.

This drop in trading activity following open meetings could suggest that members of Congress have access to the information discussed during these meetings ahead of time and could be acting on it in anticipation of its public disclosure. This could be a less obvious way of using insider information for personal financial gain by making moves before the meeting happens. This is one theory explaining the observed decrease, but the precise cause remains ambiguous based on this analysis.

6.3.3 20-Day Windows

When the window is expanded to include the 10 trading days before and after open and closed-door meetings, the majority of statistically significant level changes are negative in both types of meetings (as seen in the 10-day open meeting window). This indicates that when the window is expanded that trading activity around any date will appear to return to normal. In other words, the impact of the closed door and open door meetings is the same because there are many distant days from when the information is revealed to the representatives included in the analysis. This result is consistent with the rest of the findings that indicate that for the most part the impact of closed-door trading on meetings is immediate and short lived.

Chapter 7

Conclusion

In this thesis we find that closed-door congressional committee meetings are associated with an immediate increase in trading activity by members of the House of Representatives. This indicates that these politicians are potentially using the exclusive information they receive at these events in attempt to profit for personal gain. Upon further inspection we see that specific closed-door meetings are correlated with events that demonstrate potential instances of trading based on nonpublic information (insider trading). We note that all evidence presented in this paper is purely circumstantial.

In addition to the trading trends we see around closed-door meetings, we also observe a decrease in trading activity following some open meetings. This might indicate that congresspeople engage in preemptive trading based on anticipation of meeting outcomes or possibly because they receive information before these meetings are held publicly.

7.1 Policy Implications

These results compound the evidence and contribute to the growing amount of literature that supports the idea that there needs to be increased transparency in government, especially with regard to trading activity. There are many ways in which increased visibility could be implemented. The first is that the reporting guidelines from the STOCK Act could be amended to require trades to be reported immediately after they are executed. Because we know that the effect of closed-door meeting is short, the current 45 day period that is allotted for congresspeople to report their trading is simply too long. This new tighter reporting guideline would allow for the general public to act in near real time and copy trades. While this would not prevent insider trading, it would be a solution that gives the public the opportunity to act in parallel when congresspeople trade.

Another way to create more visibility in the trading activity around closed-door meetings is to give more information (more than just the meeting title) about the topics discussed in these meetings. If more details around the topic of the meeting are shared, then the public would be able to gather even more circumstantial evidence about potential insider trading. Obviously this solution is more challenging because there are certain topics discussed in these spaces that are highly confidential and would create a security risk if the information is revealed to the public. Even so, if slightly more information about the meeting is revealed it would allow the public to get a better sense of if congresspeople are insider trading with the information from the meeting. This solution would more allow for bad actors to be more easily identified and it would deter congresspeople from trading on information from these meetings.

In addition to policies that add more transparency to the financial dealings of congresspeople (and their dependents), there are other solutions that the government could enforce that would prevent insider trading. In many companies, especially in the financial industry, employees are not allowed to trade on certain days known as "blackout days" [28]. These blackout days usually come in the days before earnings are released when people in management have access to nonpublic information that will have an effect on their stock price. The government could employ a similar method by not allowing for trading on days when closed-door meetings occur.

An alternative method to ensure that crimes relating to insider trading in Congress are properly investigated would be to establish an independent governing body that monitors the trading activities of congresspeople. Unfortunately it would be unlikely that this would be implemented because it would involve Congress giving more authority to courts and the judicial branch to oversee their activity.

The most popular method proposed to prevent insider trading is to have all congresspeople place their investments in a blind trust. In 2023, the "Ban Congressional Stock Trading Act" bill was introduced into the Senate, which would require that congresspeople either divest from their holdings or place their investments in a blind trust [34]. Despite widespread support from the American people, 86% saying they support the bill, the legislation never reached the Senate floor and there has been little mention of it since.

The largest obstacle to proposing policy solutions to the problem of insider trading in Congress is that the congresspeople need to be the ones to pass the legislation. This requires the them to give up some power over their personal finances, which is a concession they are unlikely to make. As shown throughout this thesis, their patterns of trading around confidential information suggests that congresspeople are trading on said information, and hence we can infer that it is profitable to do so. However, this is not the end of the road for reforming the trading policy in Congress. In the COVID-19 trading scandal where senators were exposed in the media for executing trades related to the spread of COVID months ahead of the spread of the virus in the United States, the publicity associated with the scandal had an effect on their careers. Two of the senators who ran reelection campaigns lost in razor thin margins; their opponents were able to leverage their disregard for ethics as an area of attack [3]. While Congress may be resistant to passing legislation to limit their power and allow others to charge them with violations, the American people have the ability to hold these public servants accountable with their votes, and as seen in 2020 this can be very effective.

7.2 Future Work

As in many theses, this study was limited by two main factors: time and data. With more time, this study would ideally refine its models to better capture the difference in trading activity that occurs around closed-door meetings.

7.2.1 Refining the ARIMA Model

The current ARIMA model offers a foundational analysis of the impact that meetings have on trading activity. In the future, other features that are not currently in the regression that are correlated with the occurrence of meetings and affect trading patterns should be examined, as they would would bias the coefficients. For example overall market trends should be considered to check the trading activity in Congress against larger markets like the S&P 500, as done in the regression discontinuity portion of the analysis. Features could also be added to account for major international or political events (elections, trade agreements, etc.) or other macroeconomic indicators (inflation, unemployment, etc.). It would also be interesting to include indicators for committee meetings types (Intelligence, Appropriations, etc.) in order to see if some of these closed-door meetings yield information that is more valuable to trade on.

7.2.2 Challenges in Regression Discontinuity Design

The model of linear regression discontinuity design is well suited for looking at the impact of a discrete event on time series data, but it does not come without its shortcomings. The first limitation is that the way the model is currently constructed, the day of the meeting is excluded in the analysis. This is a challenge because it does not allow us to look at when congresspeople make trades on the day of the meeting. When the results of the regression discontinuity design are looked at with the results of the t-test and the ARIMA model, this issue is mitigated because the day of the meeting is included in those tests.

Another major challenge is that it is hard to select an appropriate window to perform the RDD analysis on. Ideally a smaller window will allow for examining the short impact of the new information received during the meetings, but when too few data points are included it is difficult to get statistically significant results, given the t-statistic's dependence on sample size. As a result, the window must be larger, but this makes the immediate impact of the event smaller on the linear regression. Constructing the model this way mitigates the trends seen in the t-test and ARIMA model results which indicate that the effect of meetings on trading is more immediate and dwindles as soon as one day after the meeting.

Finally, the analysis is also limited by the risk of false positives from running many hypothesis tests. Some results may appear statistically significant by simply chance alone. This phenomenon can lead to erroneously identifying patterns or effects that do not exist.

7.2.3 Advancing Insights

With extra time, the thesis could have taken a more mathematical approach the the policy implications section that simulates their potential impact. For example we could explore how instituting blackout periods may change what insider trading looks like with regard to trading patterns and how these periods of no trading affect congresspeople's returns. Looking at the results of a potential policy impact could inspire Congress to implement one of the solutions by offering concrete mathematical backing.

Ultimately many studies that would be most interesting and best at finding insider trading in congress are limited by data. This study examines data from 2015 to 2020 because it was thoroughly preprocessed. However, in the future to achieve timely insights into insider trading, enhanced data accessibility is essential. To achieve this, a substantial overhaul of the STOCK Act reporting mechanisms is necessary to ensure transparency and immediate public access, a pivotal step toward further holding American politicians accountable.

Appendix A

Code

The code for this thesis can be found in the following public GitHub repository:

https://github.com/burkepagano/SeniorThesis/

Appendix B

Tables

Event Date	**Level Change**	**P-Value**	$\alpha = 0.10$	$\alpha = 0.15$
March 02, 2015	13.000	0.070	✓	✓
March 17, 2015	4.375	0.737		
March 23, 2015	0.108	0.994		
April 14, 2015	4.820	0.343		
April 15, 2015	6.455	0.185		
July 28, 2015	-3.090	0.495		
September 18, 2015	23.223	0.514		
October 06, 2015	-0.920	0.832		
February 01, 2016	-21.477	0.214		
January 10, 2017	-9.335	0.380		
January 24, 2017	19.130	0.200		
March 02, 2017	-25.125	0.340		
May 02, 2017	-24.965	0.200		
May 04, 2017	-75.030	0.297		
May 23, 2017	22.760	0.092	✓	✓
June 14, 2017	37.381	0.400		
June 15, 2017	-39.500	0.121		✓*
June 21, 2017	5.321	0.883		
June 22, 2017	28.715	0.450		
June 27, 2017	8.410	0.352		
June 28, 2017	-14.478	0.688		
June 29, 2017	-7.845	0.845		
July 13, 2017	13.950	0.729		
November 02, 2017	10.460	0.717		
November 30, 2017	-13.595	0.349		

Table B.1: Level Change for 10-Day RDD (Closed Meetings 2015-2017)

Event Date	**Level Change**	**P-Value**	$\alpha = 0.10$	$\alpha = 0.15$
April 11, 2018	-6.784	0.466		
April 12, 2018	-1.880	0.826		
April 19, 2018	-4.370	0.271		
April 25, 2018	8.627	0.101		✓
April 26, 2018	7.680	0.200		
May 22, 2018	5.470	0.375		
June 21, 2018	0.590	0.921		
September 28, 2018	-5.438	0.764		
December 20, 2018	47.640	0.181		
February 06, 2019	-27.507	0.075	✓*	✓*
April 03, 2019	10.910	0.418		
May 01, 2019	-6.716	0.617		
May 02, 2019	-8.360	0.643		
May 08, 2019	16.321	0.377		
May 09, 2019	-9.975	0.555		
May 14, 2019	-19.795	0.311		
May 15, 2019	-12.948	0.491		
May 20, 2019	19.108	0.292		
February 06, 2020	28.180	0.230		
February 27, 2020	77.955	0.057	✓	✓
March 10, 2020	-5.415	0.455		
March 11, 2020	-71.918	0.482		
March 12, 2020	-12.455	0.914		
July 31, 2020	-7.262	0.846		
September 30, 2020	14.843	0.449		

Table B.2: Level Change for 10-Day RDD (Closed Meetings 2018-2020)

Event Date	**Level Change**	**P-Value**	$\alpha = 0.10$	$\alpha = 0.15$
February 11, 2015	-33.500	0.333		
February 13, 2015	-21.108	0.625		
February 25, 2015	-5.776	0.384		
February 26, 2015	-7.175	0.305		
February 27, 2015	-0.508	0.953		
March 3, 2015	1.170	0.854		
March 4, 2015	-4.799	0.494		
March 5, 2015	-5.340	0.424		
March 16, 2015	11.000	0.424		
March 18, 2015	-9.209	0.308		
March 19, 2015	-15.200	0.207		
March 24, 2015	-1.715	0.909		
March 25, 2015	2.963	0.838		
March 26, 2015	-4.280	0.701		
April 3, 2015	-18.308	0.127		✓*
April 18, 2015	-10.150	0.070	✓*	✓*
April 22, 2015	0.746	0.896		
April 23, 2015	-2.495	0.876		
May 14, 2015	-5.375	0.273		
May 20, 2015	-49.224	0.491		
June 18, 2015	8.715	0.749		
June 24, 2015	-11.022	0.738		
July 8, 2015	37.799	0.136		
July 9, 2015	55.130	0.023	✓	✓
September 10, 2015	8.160	0.664		
February 3, 2016	-58.052	0.114		
February 4, 2016	-27.020	0.511		
February 10, 2016	0.425	0.983		
February 11, 2016	0.485	0.981		
February 12, 2016	-22.300	0.224		
February 23, 2016	-22.150	0.084	✓*	✓*
February 24, 2016	-4.239	0.744		
February 25, 2016	2.095	0.875		
February 26, 2016	3.000	0.843		
March 1, 2016	-10.775	0.198		
March 14, 2016	-11.600	0.298		
March 15, 2016	4.040	0.640		
March 16, 2016	5.022	0.515		
March 17, 2016	4.030	0.546		

Table B.3: Level Change for 10-Day RDD (Open Meetings 2015-2016)

Event Date	**Level Change**	**P-Value**	$\alpha = 0.10$	$\alpha = 0.15$
January 1, 2017	-63.400	0.025	✓*	✓*
February 15, 2017	0.134	0.991		
February 16, 2017	1.865	0.886		
February 28, 2017	-35.445	0.162		
March 1, 2017	-29.231	0.239		
March 8, 2017	5.679	0.658		
March 9, 2017	8.165	0.536		
March 16, 2017	-16.895	0.241		
March 20, 2017	27.300	0.183		
March 21, 2017	25.500	0.055	✓	✓
March 22, 2017	25.269	0.052	✓	✓
March 28, 2017	-23.885	0.318		
March 29, 2017	3.731	0.875		
April 4, 2017	-20.245	0.110		✓*
April 5, 2017	-53.545	0.015	✓*	✓*
April 26, 2017	-7.052	0.913		
May 3, 2017	-120.761	0.060	✓*	✓*
May 16, 2017	-25.390	0.333		
May 17, 2017	-36.836	0.082	✓*	✓*
May 18, 2017	-43.190	0.049	✓*	✓*
May 24, 2017	11.799	0.354		
May 25, 2017	-14.275	0.181		
June 7, 2017	-7.284	0.750		
June 8, 2017	-12.025	0.732		
June 12, 2017	47.715	0.318		
June 13, 2017	31.140	0.500		
June 20, 2017	-21.595	0.585		
June 23, 2017	9.777	0.250		
June 26, 2017	10.154	0.234		
July 11, 2017	-35.455	0.527		
July 12, 2017	3.970	0.940		
July 17, 2017	27.262	0.030	✓	✓
July 18, 2017	-63.270	0.108		✓*
July 19, 2017	-60.739	0.110		✓*
September 14, 2017	-22.050	0.130		✓*
September 28, 2017	-15.220	0.472		
October 24, 2017	22.775	0.334		
October 25, 2017	22.545	0.343		
November 1, 2017	-8.604	0.740		
November 15, 2017	-0.597	0.965		
November 29, 2017	115.007	0.194		
December 1, 2017	-164.900	0.120		✓*

Table B.4: Level Change for 10-Day RDD (Open Meetings 2017)

Event Date	**Level Change**	**P-Value**	$\alpha = 0.10$	$\alpha = 0.15$
January 18, 2018	15.795	0.525		
January 29, 2018	-12.854	0.199		
February 5, 2018	42.200	0.039	✓	✓
February 26, 2018	8.477	0.512		
March 1, 2018	-16.060	0.067	✓*	✓*
March 6, 2018	6.945	0.423		
March 7, 2018	6.112	0.417		
March 14, 2018	5.448	0.253		
March 15, 2018	2.630	0.570		
March 20, 2018	-11.470	0.426		
March 21, 2018	-2.321	0.876		
March 22, 2018	5.240	0.723		
April 13, 2018	12.892	0.183		
April 17, 2018	8.025	0.114		✓
April 18, 2018	3.716	0.506		
April 19, 2018	-4.370	0.271		
May 7, 2018	3.315	0.848		
May 8, 2018	3.875	0.783		
May 9, 2018	14.455	0.283		
May 10, 2018	22.805	0.105		✓
May 15, 2018	-12.490	0.383		
May 16, 2018	-8.746	0.505		
May 17, 2018	1.665	0.894		
May 23, 2018	5.142	0.416		
May 24, 2018	-0.200	0.975		
June 6, 2018	26.209	0.090	✓	✓
June 7, 2018	9.935	0.191		
June 13, 2018	6.761	0.694		
June 15, 2018	0.077	0.993		
June 20, 2018	8.045	0.224		
June 26, 2018	-4.170	0.533		
June 28, 2018	4.535	0.544		
July 11, 2018	24.179	0.167		
July 12, 2018	8.340	0.285		
July 19, 2018	0.350	0.983		
July 25, 2018	5.993	0.474		
September 5, 2018	3.888	0.603		
September 13, 2018	2.345	0.474		
December 13, 2018	-3.975	0.772		

Table B.5: Level Change for 10-Day RDD (Open Meetings 2018)

Event Date	**Level Change**	**P-Value**	$\alpha = 0.10$	$\alpha = 0.15$
January 1, 2019	24.535	0.576		
January 30, 2019	-15.425	0.452		
February 7, 2019	-31.200	0.038	✓*	✓*
February 12, 2019	16.220	0.258		
February 13, 2019	11.590	0.341		
February 26, 2019	29.060	0.086	✓	✓
February 27, 2019	-19.970	0.278		
February 28, 2019	-40.435	0.072	✓*	✓*
March 6, 2019	9.515	0.542		
March 7, 2019	-21.995	0.424		
March 12, 2019	30.465	0.286		
March 13, 2019	46.806	0.073	✓	✓
March 26, 2019	-197.565	0.038	✓*	✓*
March 27, 2019	-122.209	0.221		
March 28, 2019	-76.190	0.481		
April 2, 2019	11.115	0.441		
April 4, 2019	-11.550	0.260		
April 9, 2019	-4.760	0.754		
April 10, 2019	7.440	0.594		
April 29, 2019	2.577	0.875		
April 30, 2019	3.035	0.799		
May 7, 2019	24.965	0.202		
May 10, 2019	-14.415	0.582		
May 16, 2019	-10.885	0.574		
May 17, 2019	-7.085	0.714		
May 21, 2019	-3.340	0.802		
May 22, 2019	2.351	0.857		
May 23, 2019	1.010	0.930		
June 3, 2019	-30.158	0.016	✓*	✓*
June 4, 2019	-2.333	0.842		
June 5, 2019	-12.846	0.484		
June 11, 2019	6.667	0.086	✓	✓
June 12, 2019	-2.692	0.522		
June 13, 2019	-6.000	0.114		✓*

Table B.6: Level Change for 10-Day RDD (Open Meetings 2019 Q1 & Q2)

Event Date	Level Change	P-Value	$\alpha = 0.10$	$\alpha = 0.15$
July 10, 2019	-12.299	0.507		
July 11, 2019	-1.345	0.945		
July 24, 2019	0.806	0.928		
July 25, 2019	1.775	0.849		
September 10, 2019	58.700	0.033	✓	✓
September 17, 2019	-6.410	0.774		
September 18, 2019	12.052	0.567		
September 19, 2019	9.490	0.239		
September 24, 2019	7.150	0.331		
September 25, 2019	6.239	0.387		
September 26, 2019	11.790	0.064	✓	✓
October 16, 2019	-17.246	0.324		
October 17, 2019	7.310	0.652		
November 13, 2019	-3.239	0.806		
November 14, 2019	10.025	0.477		
November 15, 2019	-0.385	0.985		
November 19, 2019	0.800	0.962		
November 20, 2019	-1.478	0.936		
November 21, 2019	21.790	0.264		

Table B.7: Level Change for 10-Day RDD (Open Meetings 2019 Q3 & Q4)

Event Date	**Level Change**	**P-Value**	$\alpha = 0.10$	$\alpha = 0.15$
February 5, 2020	21.209	0.371		
February 11, 2020	-43.445	0.071	✓*	✓*
February 12, 2020	-20.201	0.384		
February 26, 2020	27.493	0.560		
March 3, 2020	-35.465	0.051	✓*	✓*
March 4, 2020	-16.336	0.390		
March 5, 2020	0.955	0.960		
March 24, 2020	-10.650	0.929		
March 25, 2020	73.582	0.490		
May 6, 2020	28.672	0.217		
May 28, 2020	-5.345	0.609		
June 4, 2020	8.430	0.576		
June 11, 2020	-16.780	0.445		
June 15, 2020	-3.977	0.830		
June 18, 2020	-7.610	0.625		
June 23, 2020	-5.780	0.551		
July 1, 2020	-7.724	0.561		
July 6, 2020	13.162	0.429		
July 7, 2020	-2.850	0.893		
July 8, 2020	-0.709	0.973		
July 9, 2020	6.100	0.774		
July 10, 2020	7.731	0.758		
July 13, 2020	32.662	0.174		
July 14, 2020	0.545	0.946		
July 15, 2020	-29.328	0.110		✓*
July 29, 2020	-45.701	0.067	✓*	✓*
September 11, 2020	14.023	0.620		
October 2, 2020	7.985	0.696		
October 15, 2020	30.795	0.390		

Table B.8: Level Change for 10-Day RDD (Open Meetings 2020)

Bibliography

[1] Russia (RUS) Exports, Imports, and Trade Partners. The Observatory of Economic Complexity. Accessed 1 Apr. 2024.

[2] Stop Trading on Congressional Knowledge Act of 2012. (112-105), 2012. Public Law 112-105, 126 Stat. 291.

[3] Democrats hammer Perdue, Loeffler over stocks, 2020. Accessed 30 Mar. 2024.

[4] S&P 500. S&P Dow Jones Indices, Mar. 2024. Available online at https://www.spglobal.com/spdji/en/indices/equity/sp-500/overview.

[5] The House Explained. Website, Accessed 7 Apr. 2024. https://www.house.gov/the-house-explained.

[6] David Abugaber. Using ARIMA for Time Series Analysis. https://ademos.people.uic.edu/Chapter23.html. Accessed: February 26, 2024.

[7] Adam L. Aiken and Jesse A. Ellis and Minjeong Kang. Do Politicians "Put Their Money Where Their Mouth Is?" Ideology and Portfolio Choice. *Management Science*, 66(1):379–400, 2019.

[8] William Belmont, Bruce Sacerdote, Ranjan Sehgal, and Ian Van Hoek. Do senators and house members beat the stock market? evidence from the stock act. *Journal of Public Economics*, 207:104602, 2022.

[9] Jose L. Bernal, Steven Cummins, and Antonio Gasparrini. Interrupted time series regression for the evaluation of public health interventions: a tutorial. *International Journal of Epidemiology*, 46(1):348–355, 2017.

[10] George E. Box, Gwilym M. Jenkins, and Gregory C Reinsel. *Time Series Analysis: Forecasting and Control*. John Wiley & Sons, 1994.

[11] Peter J. Brockwell and Richard A. Davis. *Introduction to Time Series and Forecasting*. Springer Texts in Statistics. Springer International Publishing, 2016.

[12] Kenneth P Burnham and David R Anderson. Multimodel Inference: Understanding AIC and BIC in Model Selection. *Sociological Methods & Research*, 33(2):261–304, 2004.

[13] Matias D. Cattaneo and Rocío Titiunik. Regression Discontinuity Designs. *Annual Review of Economics*, 14(1):821–851, 2022.

[14] Cornell Law School. Insider trading. Legal Information Institute, No Year. Accessed on 7 Apr. 2024.

[15] Joyce E. Ewusie, Charlene Soobiah, Erik Blondal, Josephy Beyene, Lehana Thabane, and Jemila S. Hamid. Methods, Applications and Challenges in the Analysis of Interrupted Time Series Data: A Scoping Review. *Journal of Multidisciplinary Healthcare*, 13:411–423, 2020.

[16] R. A. Fisher. *Statistical Methods for Research Workers*. Oliver and Boyd, Edinburgh, 1925.

[17] International Trade Council. Major Exports of Iraq. https://thetradecouncil.com/major-exports-of-iraq/, 2023.

[18] Jane J. Kim. U.S. Senators' Stock Picks Outperform the Pros'. *Wall Street Journal*, 2004.

[19] Kimberly Kindy. Congressional rules on trading had their start in 1789. *Washington Post*, 2012.

[20] Sylvan Lane. Four senators sold stocks before coronavirus threat crashed market. The Hill, 2020.

[21] David S. Lee and Thomas Lemieux. Regression Discontinuity Designs in Economics. *Journal of Economic Literature*, 48(2):281–355, June 2010.

[22] Dave Levinthal. 78 members of congress have violated a law designed to prevent insider trading and stop conflicts-of-interest, 2023.

[23] Michael Martinez and Mohammed Tawfeeq. Isis Bulldozes Ruins of Ancient Assyrian City of Nimrud, Iraqi Ministry Says. *CNN*, Mar 2015.

[24] D. S. Moore, W. I. Notz, and M. A. Flinger. *The Basic Practice of Statistics*. W. H. Freeman and Company, New York, NY, 6th edition, 2013.

[25] Ray Mrinmoy. Application of Time Series Intervention Modelling for Modelling and Forecasting Cotton Yield. *ResearchGate*, 2014.

[26] Robert Nau. Statistical Forecasting: Notes on Regression and Time Series Analysis, 2020.

[27] Oscar Torres-Reyna. Time Series. https://www.princeton.edu/ otorres/TS101.pdf, 2008.

[28] Bob Schneider. What Is a Blackout Period in Finance? Rules and Examples. https://www.investopedia.com/ask/answers/08/blackout-period.asp, 2024. Accessed 30 Mar. 2024.

[29] United States Senate. Senate Committees Frequently Asked Questions, 2023.

[30] Sonam Sheth. Sen. David Perdue Bought Stock in a Company That Produces Protective Medical Equipment the Same Day Senators Received a Classified Briefing on the Coronavirus. *Business Insider*, 2020. Accessed 26 Mar. 2024.

[31] Pennsylvania State University. Stat 510: Applied Time Series Analysis. https://online.stat.psu.edu/stat510/lesson/3/3.1, 2023. Accessed: 10/20/2023.

[32] U.S. Office of Government Ethics. Stop Trading on Congressional Knowledge Act of 2012 (STOCK Act), Accessed 2023. Accessed on November 30, 2023.

[33] U.S. Securities and Exchange Commission. Insider Trading Policy. https://www.sec.gov/Archives/edgar/data/1164964/000101968715004168/globalfuture$_8k-ex9904.htm$, 2015.

[34] U.S. Senator for Georgia Jon Ossoff. Sens. Ossoff, Kelly Introduce Bill Banning Stock Trading by Members of Congress. https://www.ossoff.senate.gov/press-releases/sens-ossoff-kelly-introduce-bill-banning-stock-trading-by-members-of-congress-2/, 2023. Accessed 12 Sept. 2023.

[35] J.M. Wooldridge. *Introductory Econometrics: A Modern Approach.* Cengage Learning India, 2020.

[36] Yahoo! Finance. S&P 500 (ˆGSPC) Historical Data. https://finance.yahoo.com/quote/%5EGSPC/history?p=%5EGSPC, 2024. Accessed on 29 Mar. 2024.