Researching the Causal Effect of r/WallStreetBets on Stock Price and Volume

CSDS 442

Professor Andy Podgurski

Takoda Denhof, Brandon Rudolph, Chris Toomey

Due 7 May 2021

Researching the Causal Effect of r/WallStreetBets on Stock Price and Volume

An academic paper authored by: Takoda Denhof, Brandon Rudolph, Chris Toomey

Department of Computer and Data Sciences, Case Western Reserve University

Cleveland, OH 44016-7076

May 7, 2021

***Abstract***

**In this paper, we research if there is a causal relationship between whether a stock is mentioned on the Reddit forum r/WallStreetBets (WSB) and the price of a stock. We also investigate if the number of mentions has any causal relationship with price. In order to do this, we use IP Weighting and Propensity Score Matching to mitigate confounding bias for the question of whether a stock is mentioned has an effect on the stock price, and linear regression for the question of whether the number of mentions matters.**

**We found that there is a causal relationship between whether a stock is mentioned and the price, with the strongest relationship shown using IP Weighting, but there is no causal relationship between the number of times a stock is mentioned and its price.**

**I. BACKGROUND**

***A. WallStreetBets (WSB) and Causal Questions***

In January of 2021, the stock price of video game retailer GameStop increased over 1,800% in a matter of weeks. Although the security’s price has since dropped, the story made national headlines due to the influence a Reddit forum known as r/WallStreetBets, or WSB, had on the stock’s price. The social media webpage encouraged retail traders on the brokerage app Robinhood to purchase the security, prompting a short-squeeze which resulted in multi-million dollar losses for hedge funds such as Melvin Capital, who lost $846 million as a result [1].

A company’s stock price is meant to represent the present and future value of its business: better financial fundamentals are rewarded with higher relative valuations on stock price. But ultimately, a stock’s price is determined by the laws of supply and demand, and GameStop isn’t the only company featured on WSB who saw an astronomical rise in price due to a heavy influx of buyers. The movie theater chain AMC, electronics provider Blackberry, and retailer Bed, Bath & Beyond are just a handful of stocks which have seen heightened levels of volatility these past few months.

The purpose of this report is to investigate whether there is a causal effect of a company being mentioned on the WSB forum (the treatment) on that company’s stock price (the outcome). Furthermore, this report will investigate whether there is a linear causal relationship between the number of times a stock is mentioned on the forum and its change in price. This particular regression will examine only the treated subjects (since the untreated subjects have a mention total of 0).

***B. Overview of Short Selling***

Although not the only possible confounder (a variable which affects both the treatment and outcome), short-selling is the most notable confounder for this study due to the substantial coverage of the topic during the height of the WSB saga. One of the most discussed characteristics which drove a high volume of mentions for companies such as GameStop or AMC was the high levels of short-interest in these companies [2].

Short-selling is a trading technique where an investor borrows a stock, sells that stock at the current market price, and then buys back the stock at a future date. The difference between the price at the beginning of the short (when you borrow and sell the stock) and when the short is covered (purchase back the stock to return) is the profit. If the stock’s price falls, the investor makes money-if it rises, the investor loses money. Generally speaking, short-selling is an investment technique that is reserved for hedge funds or institutional investors and is not readily available to retail traders.

Short interest refers to the percentage of tradable shares that are being held short. A higher level of short interest in a company indicates that investors believe that stock’s price is likely to fall [3]. Since the original motivation for the WSB movement was to make hedge funds lose large amounts of money by driving up the price of shorted securities by increasing demand (forcing hedge funds to cover their position for a loss), we used short interest as the basis for other confounders.

***C. Confounding and Possible Confounders***

Confounding refers to outside observed variables which influence both the likelihood of receiving a treatment and the outcome of that treatment. In the context of the WSB example, short selling increases the likelihood a stock is mentioned on WSB (the treatment) and has a possible effect on a company’s stock price (since high short interest is associated with overpriced stocks).

Another possible confounding variable is the sector/industry a company operates in. Some sectors are known to be more volatile than others (technology stock vs. utilities for example). Since we know that a certain type of company is favored on WSB, namely consumer cyclicals and technology, a stock’s sector influences the likelihood of receiving the treatment. Furthermore, since a company’s sector influences the level of returns that company’s stock is likely to receive, it also influences the outcome.

Market capitalization, the total market value of a company, is another good example of a confounding variable we’ll want to consider. Larger, well known companies are more likely to be featured on WSB (such as Apple, Amazon, Tesla, etc.) due to their familiarity with everyday investors. As such, market capitalization has an effect on the likelihood of receiving the WSB treatment and, since large and small companies tend to generate different returns [4], have an effect on the outcome.

Finally, we’ll consider a stock’s prior 3 month return as a confounding variable to address. The WSB community values volatility in stock analysis and selection [5]. As such, stocks with high past returns may be more likely to appear on the WSB feed since they are prominent in the retail investor space. Furthermore, since past return can be argued to have an effect on the future of a stock’s price, we can conclude that past performance covers all bases to serve as a confounding variable.

***D. Data Acquisition and SwaggyStocks.com***

Thousands of stocks are traded on the Nasdaq and New York Stock Exchange. To streamline our data collection process, we limited our controls to companies listed on the S&P 500 (500 largest companies in the United States) and Russell 2000 (2,000 small-capitalization companies) indices [6].

Treatment is defined as a stock appearing on the main dashboard of SwaggyStocks with over 5 mentions on February 11, 2021 which we were able to access using the Wayback Machine [7]. SwaggyStocks is a website that provides users the ability to view stock trends on WSB. SwaggyStocks scrapes top-rated comments from top-rated posts on WSB, analyzing the text for ticker mentions using natural language processing techniques. We use SwaggyStocks to analyze WSB for us, extracting both a company’s symbol and number of mentions, and generating the data we need to distinguish the treated from untreated in our study.

Company specific and trading data was collected using an API in Python called yfinance [8]. This API collects information on company stocks stored within the Yahoo Finance database. In addition to specific pricing data, yfinance allows us to collect data on key company financial figures such as size, short interest, and sector. In our dataset, the percent difference between a stock’s closing price on February 4, 2021 (one week prior to the treatment) and February 18, 2021 (one week after the treatment) defines the outcome for our study.

All in all, our final dataset consisted of just below 160 treated stocks coupled with over 2500 untreated stocks. In addition to pricing, each stock had a value for their market capitalization, short interest, and sector information (all which are used for the propensity score calculation). Only treated stocks had a value for number of mentions, since that aspect of the analysis was restricted to treated subjects only.

***E. Study Design and Causal Effect Measures***

In this study, we are trying to mitigate the potential confounding bias and isolate the effect of WSB mentions on the price of stocks. There are unmeasured variables, such as macroeconomic trends, that cannot be accounted for. We use inverse probability weighting and propensity score matching to account for confounding bias. Then, in order to test for the marginal effect of each additional mention, we use linear regression. We also want to make sure the size of the amount of data is not a concern, and that we are confident about the causal measure, so we will use bootstrapping to get more samples and average the causal measures from each.

The causal effect measures we will be using are the causal risk difference and the causal risk ratio. The causal risk difference is defined as . The causal risk ratio is defined as . In this case, Y is the stock price and A(a) is whether a stock was mentioned on WSB. We use these two metrics because they let us measure the strength of the causal effect on both the additive and multiplicative scales.

***F. Directed Acyclic Graph***

The Directed Acyclic Graph, or DAG, provides a visualization of causal relationships between the treatment, outcome, and other interfering factors. The DAG in figure 1 displays our understanding of the relationships between WSB, a stock’s price, and other confounding variables which have the potential to affect both.

The DAG also mentions the effect that macroeconomic factors can have on the securities market as a whole and by extension a stock’s price. The S&P 500 Index is a possible surrogate for the performance of the stock market and the effect of macroeconomic trends.

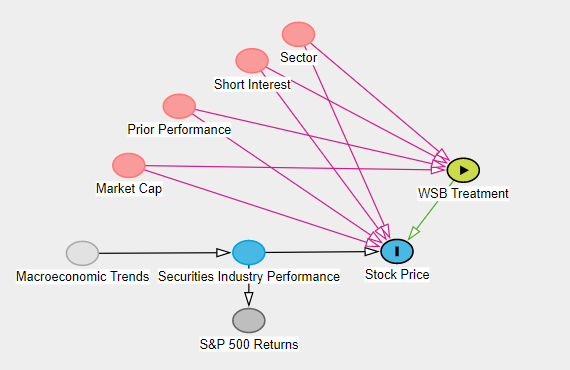


Figure 1: Causal DAG

**II. METHODS**

***A. Propensity Scoring***

Propensity score refers to the likelihood a subject will receive a given treatment in a causal effect study. With respect to this paper, propensity score will mean the probability a stock is listed on WSB. We must consider propensity scores since our study does not randomly select the stocks that receive treatment.

Propensity score is calculated by fitting a logistic regression using our confounders, L, to predict whether a company received the treatment of being mentioned on WSB. The logistic regression equation as defined in “What If” authored by Hernan and Robins can be seen in figure 2 [9].

In the context of our research, the input L is compiled using the variables of the logarithm of a company’s market capitalization, the 3 month prior price performance, the level of short interest, and the sector.

In order to include the company’s sector in the logistic regression, we generated dummy variables to input to the regression. Sector is a categorical variable, so generating a group of dummy variables (where the variable takes on the value 1 if a stock is in a particular sector and 0 otherwise) provides the logistic regression a numeric value for analysis.

Running each stock through the trained logistic regression returns a value between 0 and 1 which is meant to represent the probability of treatment for that respective stock. The Python library of scikit-learn, specifically the LogisticRegression function, was used in our generation of propensity scores. [8]

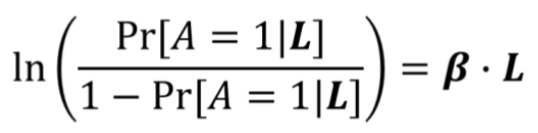


Figure 2: The logistic regression equation [9]

***B. Inverse Probability Weighting Regression***

Inverse probability weighting, or IPW, uses the propensity score for each subject to assign each subject a weight. The weight of a subject is defined as the inverse of a subject’s propensity to receive a specific treatment. That is, weights for untreated subjects (A=0) is defined as (P(A=0|L))-1 or (1-P(A=1|L))-1 and weights for treated subjects (A=1) is defined as (P(A=1|L))-1 . As a result, higher weights are assigned to subjects with lower probabilities of receiving whatever treatment was present.

The IPW regression takes the same inputs as the unweighted regression (regressing a stock’s return against the binary treatment indicator) but uses weighted least squares regression as opposed to ordinary least squares regression (which is used in the unweighted regression).

Weighted least squares considers the weight of each input variable when calculating the model’s slope and intercepts (while OLS does not). That is, weighted least squares calculates the models coefficients by minimizing 𝝨 Wi \* [Yi - (θ0+θ1Ai)]2 where θ0 is the intercept, θ1is the slope/causal risk difference, and W is the associated weight.

***C. Propensity Score Matching***

Propensity score matching refers to forming a matched population in which the treated and the untreated are exchangeable. This is done by matching untreated and treated with similar propensity scores. The resulting subset of the original population, or matched population, consists of treated-untreated pairs. We developed our own algorithm to generate the two populations. The matched population that is produced is a consistent estimator of effect measures.

In order to use propensity score matching in this experiment, we used the calculated propensity scores to match the stocks into untreated-treated pairs, creating a sort of counterfactual group for the treated subjects. Since each matched pair share similar confounding propensity scores, and since one recieved the treatment while the other did not, the matched treated population and matched untreated population are deemed exchangeable. From this, we generate our causal effect measures.

***D. Linear Regression***

A linear regression is a linear model approach to modeling the relationship between the dependent and independent variables. This is done by modeling a linear line of best fit of the form where X is the explanatory variable and Y is the dependent variable.

With respect to this paper, a marginal effect of mentions regression will be used to determine the effect of the scalar variable return on treated stock, the independent variable used will be the number of mentions on WSB. This regression will be used to determine if the number of mentions on WSB per stock has an effect on return on those who are already said to be treated. This differs from the IPW regression and propensity score matching since we are not classifying stocks into groups of treated/untreated, taking on values of 0 or 1.

***E. Bootstrapping***

Bootstrapping is an analysis tool when sample size is a concern and we wish to generate a standard error for our estimates. When dealing with a random experiment, one way to increase the number of observations is to increase the number of people (or whatever is being tested) participating in the experiment. This is not as easy in an observational study, as the researcher typically does not intervene in these situations. A common way to increase the size of samples in these situations is to use bootstrapping. Generally, this means sampling from the original sample with replacement in order to obtain more data sets.

In this experiment, we did not believe we had a sufficient amount of observations that were mentioned on WSB, and wanted to increase our confidence in the results. Since we could not increase the samples organically, we turned to bootstrapping. In this method, we sampled our original data set, whether or not a stock was mentioned on WSB, with replacement in order to obtain lots of “different” data sets. From here, we could calculate test statistics for each data set, and use that distribution to calculate the standard error of the test statistics used in analysis.

**III. RESULTS**

***A. Propensity Scoring and Weighting***

In this section, we will analyze the distribution of the proposed confounders from section I.C. above. In figure 3, we see four graphs, each featuring the distribution of both the treated and untreated subjects for each confounder. We see that the logarithmic market cap measure and price change over the past 3 months for treated subjects tends to be greater than the untreated. Contrary to our original belief, the short interest of the treated subjects was not an overwhelmingly defining feature between the subject groups. Finally, we see that treated subjects disproportionately feature technology, consumer cyclical, and communication service sectors.

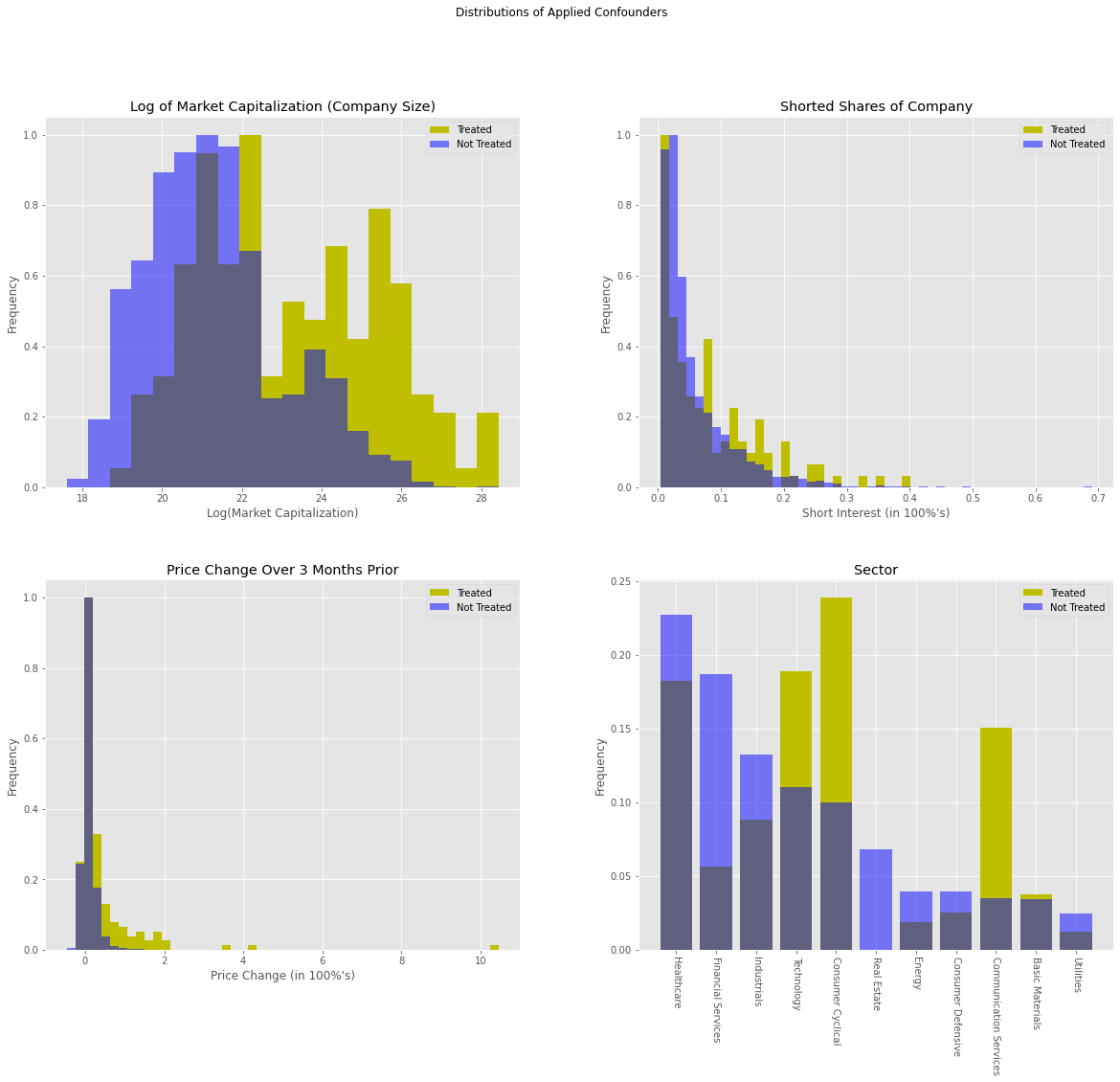


Figure 3: Distributions of Applied Confounders

As for the propensity scores and weighting, figure 4 shows the distributions of the propensity scores and weights for each subject in both subject groups. The two graphs are mathematically connected since the weight of a subject is the inverse of its propensity for the treatment it received. We can see that since the treated subjects are positively skewed (closer to a propensity score of 0 despite receiving the treatment), weights for the treated will generally be much greater than for the untreated. Although it is not ideal that the treated propensities are positively skewed, they are still clearly closer to 1 compared to the untreated subjects.

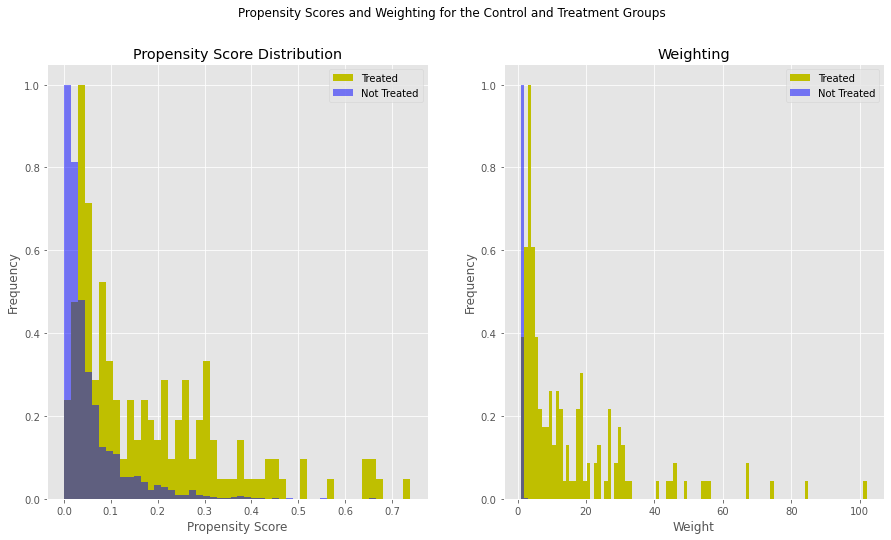


Figure 4: Propensity Scores and Weighting for the Control and Treatment Groups

***B. Unweighted (Plain Vanilla) Regression***

Bootstrapping results for intervals of 𝛽0 and 𝛽1 can be seen in figure 5 below.

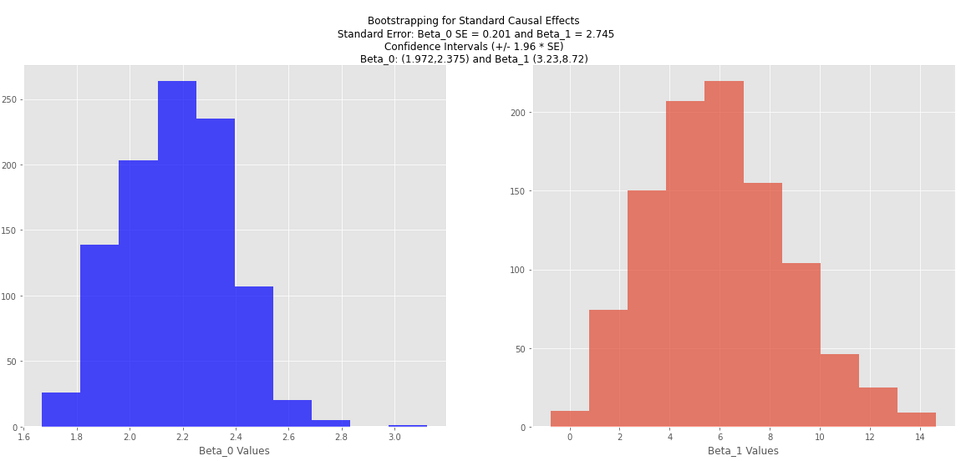


Figure 5: Bootstrapping for Standard Causal Effects

***C. IPW Regression***

The IPW regression uses the weights calculated in section III.A during the regression stage. The return regressions for both the weighted and unweighted models can be seen in figure 6 (based on a logarithmic scale to make difference visible). Bootstrapping results for the IPW model’s 𝛽0 and 𝛽1 can be seen in figure 7.

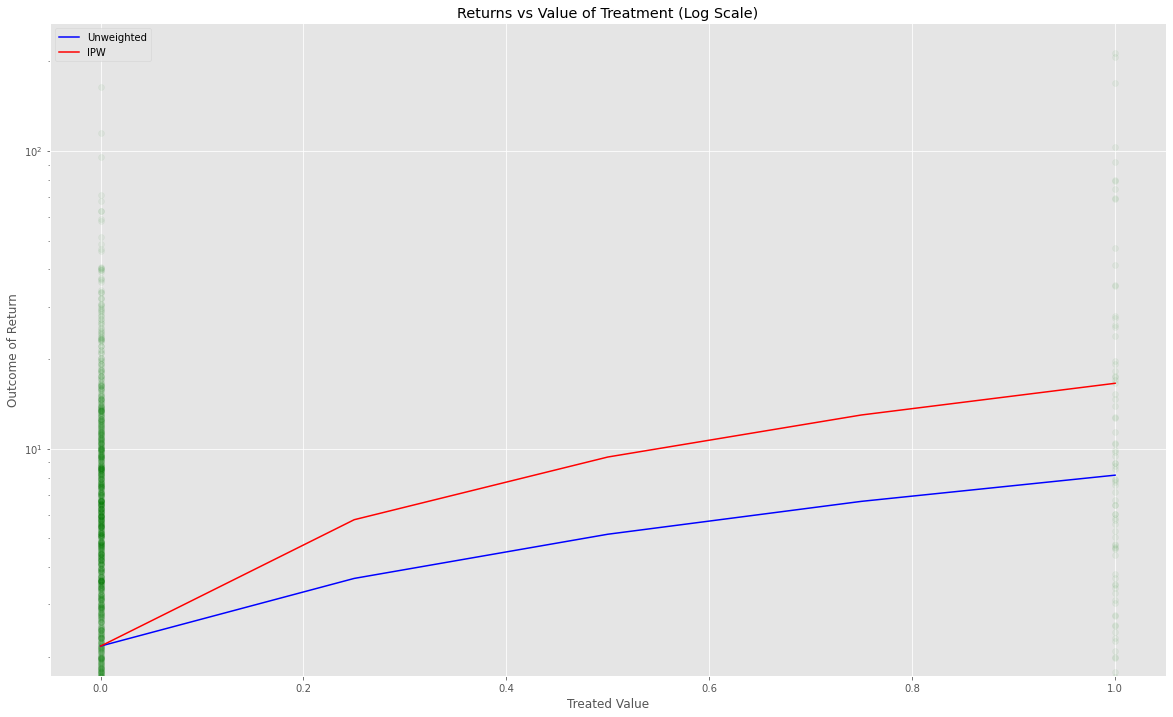
******

Figure 6: Returns vs Value of Treatment (Log Scale)

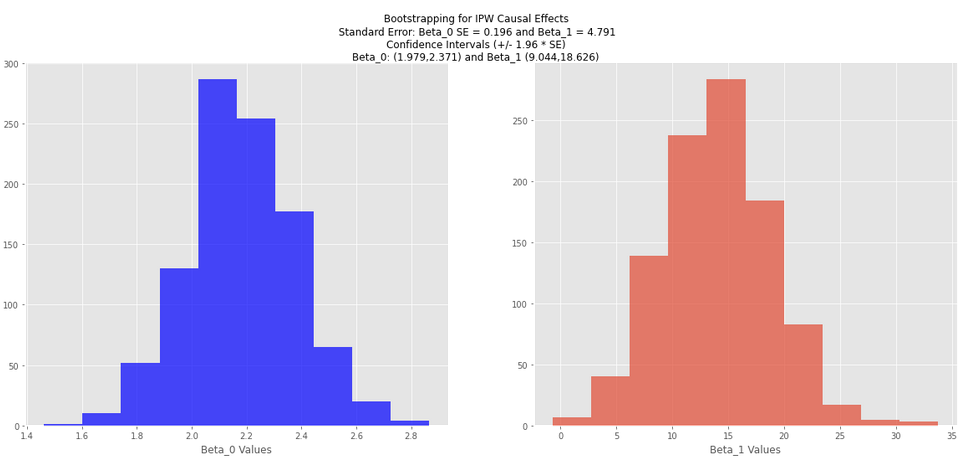


Figure 7: Bootstrapping for IPW Causal Effects

***D. Propensity Score Matching Regression***

Figure 8 displays the propensity score pairing and population distributions, along with the associated outcome results. We see that as propensity scores increase, it’s more difficult to find a match (since untreated propensities were more heavily skewed than treated). Bootstrapping results for the propensity score matching model’s 𝛽0 and 𝛽1 can be seen in figure 9.

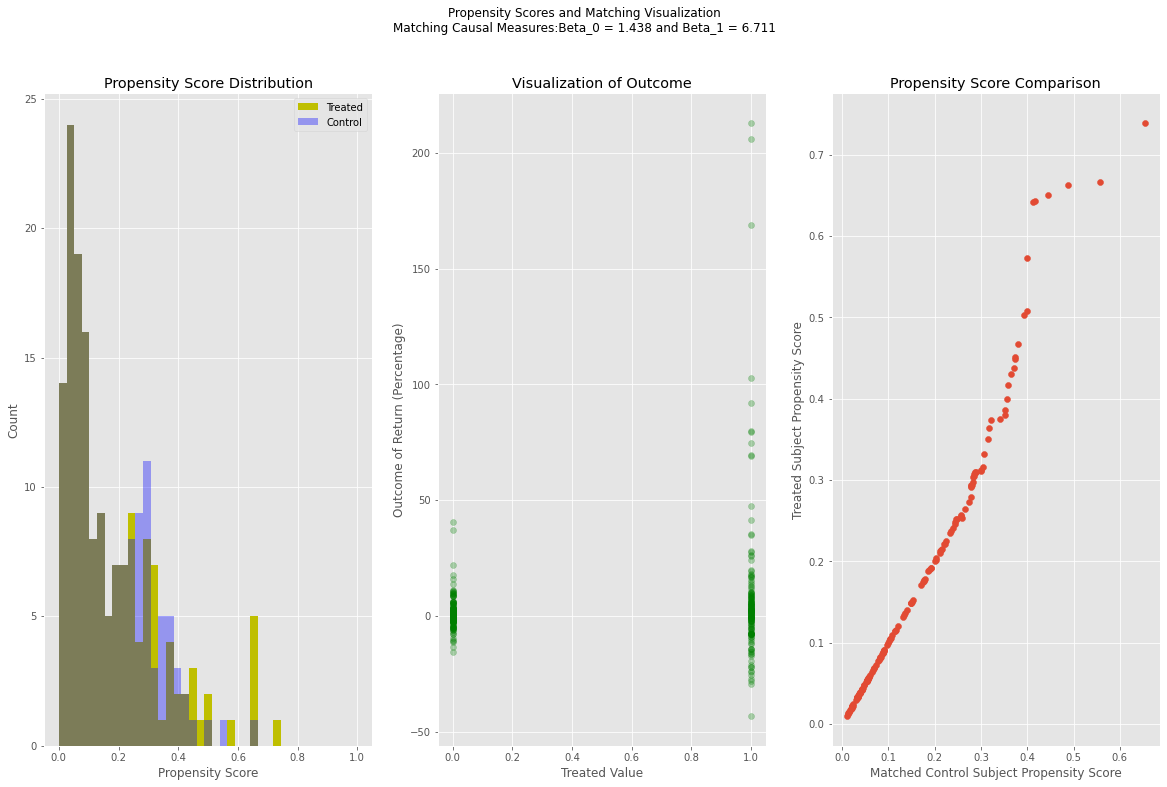


Figure 8: Propensity Scores and Matching Visualization

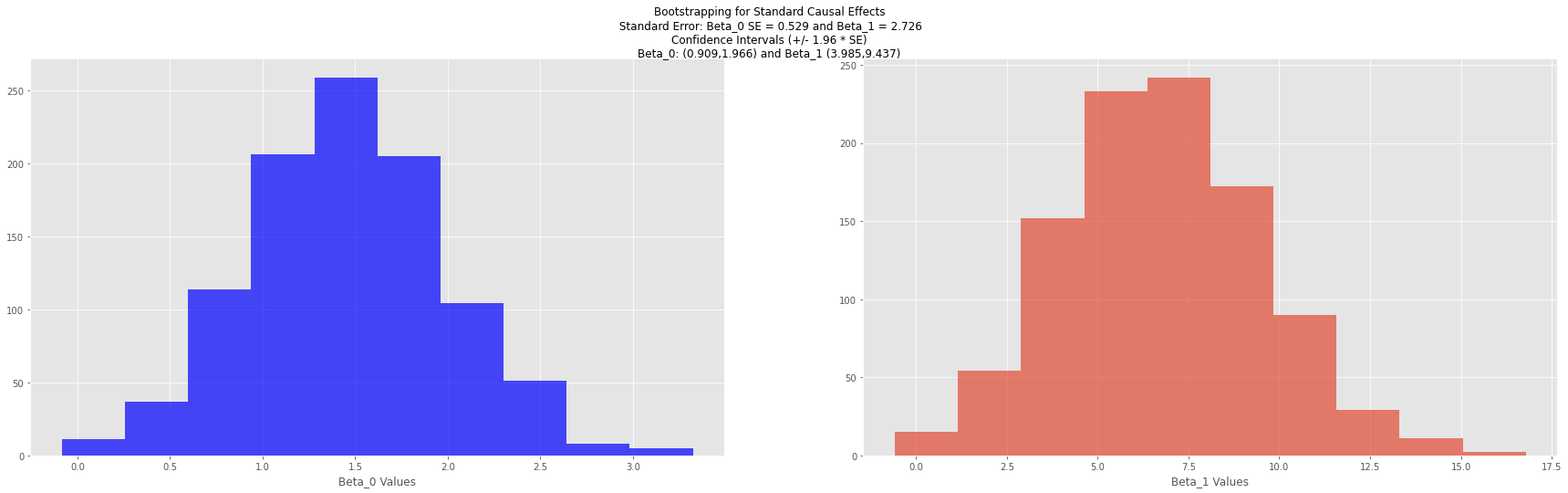
******

Figure 9: Bootstrapping for Standard Causal Effects

***E. Marginal Effect of Mentions Regression***

Figure 10 displays a visualization of a treated company’s return against the number of mentions in WSB (accompanied by the least squares regression line) alongside the associated residual chart. Bootstrapping results for the linear regression model’s 𝛽0 and 𝛽1 can be seen in figure 11.

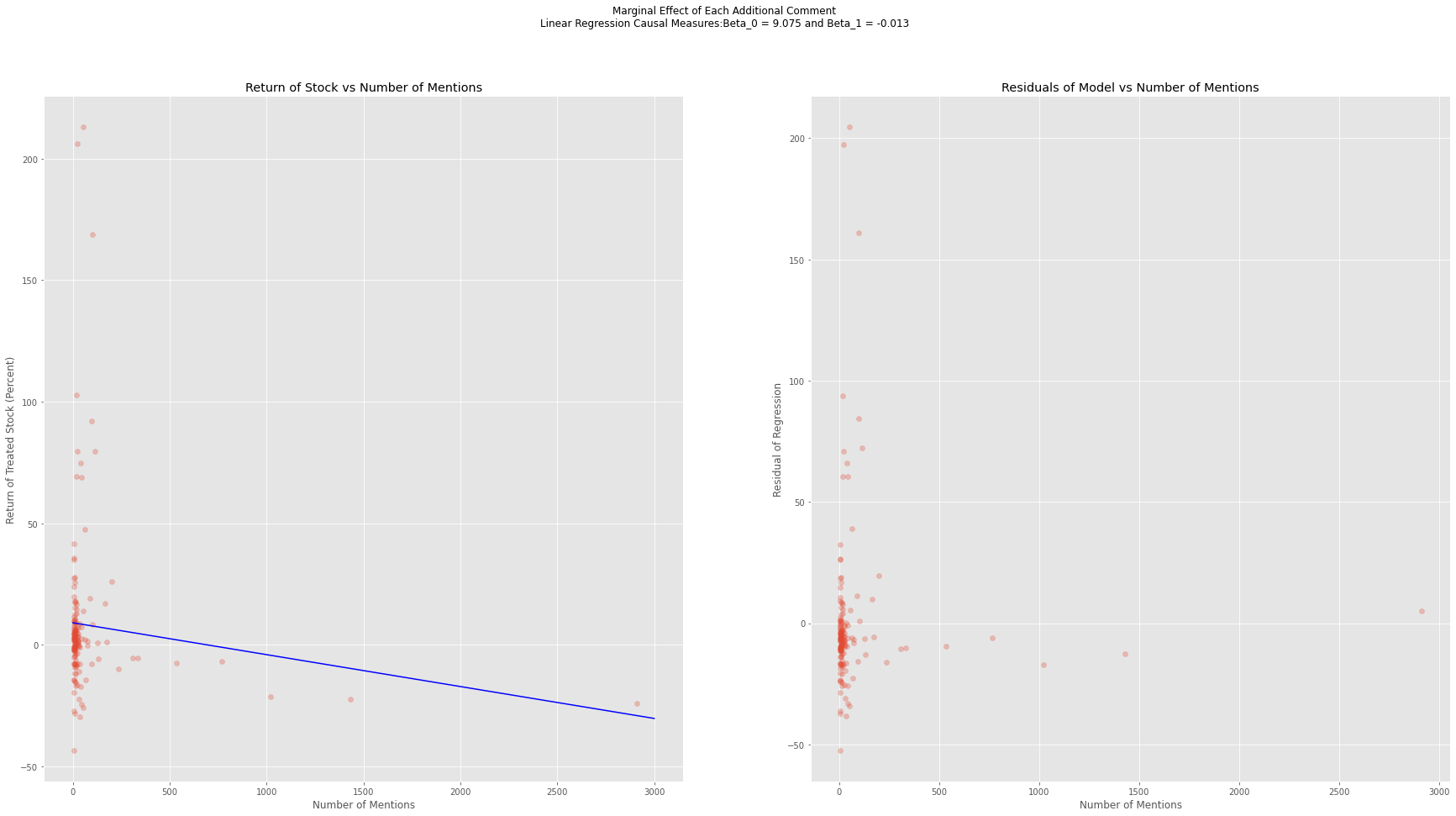


Figure 10: Marginal Effect of Each Additional Comment

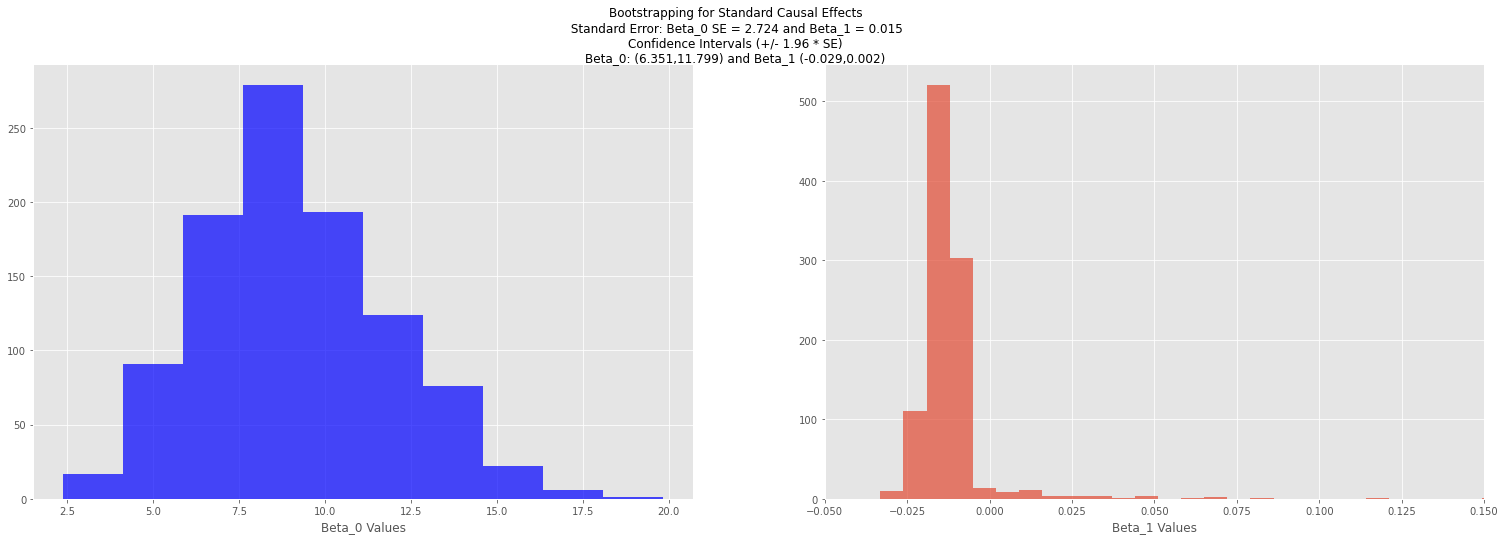
****

Figure 11: Bootstrapping for Standard Causal Effects

**IV. DISCUSSION**

***A. Overview of Results***

All three models which tested the effect of the treatment of WSB on a stock’s price (plain vanilla regression, IPW, and propensity score matching) saw an additive causal effect measure (𝛽1) significantly greater than 0. Using the bootstrapping technique described earlier in this paper, confidence intervals were calculated by adding and subtracting 1.96 times the standard deviation of the bootstrap distribution (standard error of the statistic). This generated a 95% confidence interval for the 𝛽0 and 𝛽1 values, which can be found in sections III.A and III.B above.

Of the three (plain vanilla regression, IPW, and propensity score matching), IPW’s regression had the greatest additive and multiple effect measure values of 14.408% and 7.634, respectively. That is, the return of treated stocks under the IPW model were 14.408% greater than the untreated stocks and 7.634 times greater than the untreated. The corresponding effect measures for the propensity score matching model were 6.711% and 5.667.

The other aspect of our analysis, the marginal effect of mentions, did not see the same positive and significant results as the other models. The slope coefficient which encapsulated the unit effect each additional mention had on a company’s stock price, 𝛽1, was not significant at the 95% level. The interval of (-.029, .002) included 0, indicating that the 𝛽1 value is not significantly different from 0 and forces us to reject the notion that the number of mentions on WSB has any effect on a stock’s price.

***B. Shortcomings of Analysis***

Knowledge of subject matter is required when performing causal effect analysis, especially if identifying confounders is involved [9]. Although all authors of this paper performed extensive research on both the securities industry and the WSB phenomenon, it stands to reason that there are additional confounders left unaccounted for in our analysis. The visualization of our propensity score in figure 4 further supports this argument because, although there is a stronger skew toward the right portion of values for the treated subjects as opposed to the untreated, scores for both treatment groups remain positively skewed.

***C. Other Academic Publications***

Previous publications on the subject of r/WallStreetBets are primarily limited to the forum prior to the short-squeeze phenomenon seen early in 2021. The most explicit paper featuring WSB, “WallStreetBets: Position or Ban” out of the Georgia Institute of Technology [5], researched the psychological framework of the community itself as opposed to the profitability/causal effects the community has on financial markets.

**V. CONCLUSION**

Ultimately, this paper found a causal relationship between a treatment of a company’s presence on the WSB forum and the subsequent price of the stock. By generating propensity for treatment probabilities through confounders of short interest, company size, sector, and previous performance. The inverse probability regression showed an especially strong causal effect, as the causal risk difference was 14.408 and the causal risk ratio was 7.624. Propensity score matching showed a similar, but slightly less strong, result. Regardless, we can confidently conclude that there is a causal relationship between a company being mentioned on WSB and stock price. We conclude this based on the causal risk difference is significantly different from 0 and the causal risk ratio is significantly different from 1.

On the other hand, studying the number of mentions within the treated population did not have a significant causal relationship with stock price. The 95% confidence interval for the effect of each additional mention, (-.029, .002), encapsulates 0, indicating the causal risk measure isn’t significantly different from 0.

**VI. REFERENCES, DATA, AND CODE**

**References:**

[1] I. Lee, “The hedge funder whose GameStop short bets led to a 53% January loss took HOME $846 million in 2020,” 10-Feb-2021. [Online]. Available: https://markets.businessinsider.com/news/stocks/melvin-capital-gabe-plotkin-took-home-846-million-in-2020-2021-2-1030067950#:~:text=The%20hedge%20funder%20whose%20GameStop,million%20in%202020%20%7C%20Markets%20Insider. [Accessed: 06-May-2021].

[2] D. Burrows, “GameStop: How WSB Beat Hedge Funds at Their Own Game,” *Kiplinger*, 31-Jan-2021. [Online]. Available: https://www.kiplinger.com/investing/stocks/602158/gamestop-how-wallstreetbets-wsb-beat-hedge-funds. [Accessed: 06-May-2021].

[3] CFI, Ed., “Short Interest - Definition, How to Short a Stock, Formula,” *Corporate Finance Institute*, 14-Mar-2019. [Online]. Available: https://corporatefinanceinstitute.com/resources/knowledge/trading-investing/short-interest/. [Accessed: 06-May-2021].

[4] “Why Market Cap Matters,” *Fidelity*. [Online]. Available: https://www.fidelity.com/learning-center/trading-investing/markets-sectors/why-market-cap-matters. [Accessed: 06-May-2021].

[5] C. Boylston, B. Palacios, P. Tassev, and A. Bruckman, “(PDF) WallStreetBets: Positions or Ban.” [Online]. Available: https://www.researchgate.net/publication/348860912\_WallStreetBets\_Positions\_or\_Ban. [Accessed: 06-May-2021].

[6] T. Smith, “How the S&P 500 and Russell 2000 Indexes Differ,” *Investopedia*, 15-Mar-2021. [Online]. Available: https://www.investopedia.com/articles/investing/012215/how-sp-500-and-russell-2000-indexes-differ.asp. [Accessed: 06-May-2021].

[7] WSB - Ticker Sentiment,” *Swaggy Stocks*. [Online]. Available: <https://web.archive.org/web/20210211213513if_/https://swaggystocks.com/dashboard/wallstreetbets/ticker-sentiment>.

[8] Pedregosa, F., G., Varoquaux, A., Gramfort, V., Michel, B., Thirion, O., Grisel, M., Blondel, P., Prettenhofer, R., Weiss, V., Dubourg, J., Vanderplas, D., Passos, M., Brucher, M., Perrot, and E., Duchesnay. "Scikit-learn: Machine Learning in Python".*Journal of Machine Learning Research* 12 (2011): 2825–2830

[9] Hernán MA, Robins JM (2020). Causal Inference: What If. Boca Raton: Chapman & Hall/CRC.

**Unreferenced Data:**

*January 2021 Daily Short Sale Volume Files*. [Online]. Available: http://regsho.finra.org/regsho-January.html. [Accessed: 06-May-2021].

“2021 List of All Russell 2000 Companies: Free Excel Sheet Download,” *Sure Dividend*, 09-Apr-2021. [Online]. Available: https://www.suredividend.com/russell-2000-stocks/. [Accessed: 06-May-2021].

**Datasets and Code**

Datasets used in the analysis portion of this project (dataset generated using the WSB, yfinance, and symbol data) will be submitted along with this final report.