

Nijmegen School of Management
Department of Economics and Business Economics
Master's Thesis Economics (MAN-MTHEC)

Internet Message Board Sentiments & Stock Returns: An Analysis of the MAGA Movement

By Randy Moris (S4728858)

Nijmegen, 28 June 2025

Program: Master's Program in Economics
Specialisation: Financial Economics
Supervisor: dr. Jan Schmitz

Radboud Universiteit



- 1 • *Statement B:* Generative AI tools (e.g., ChatGPT, Copilot) were used to assist in coding, data
2 analysis, and/or refining the language of this thesis. Appendix C of this thesis provides a
3 detailed account of the use of Generative AI tools during the development of this thesis. By
4 submitting this thesis, I declare that I am fully responsible for the accuracy and completeness
5 of its content.

Abstract

In this current digital age investing in the stock market is increasingly more accessible to people across society, potentially leading to more activity on internet stock market message boards. I examine the impact of the content of over 69,000 messages, posted on the StockTwits platform, about 181 firms listed on the Russell 2000 index. The message sentiment and MAGA association are measured using a word matching model, to assign a single sentiment and MAGA association to each of the messages. I find that both negative message sentiments and positive MAGA associations immediately impact the daily stock returns of the Russell 2000 firms. While positive sentiments and negative MAGA associations possess a lagged effect on stock returns. These effects on stock returns are statistically significant but economically limited. Additionally, a heterogeneity analysis is carried out, indicating that the various industry sectors of the Russell 2000 firms influence the effects sentiments and MAGA associations have on stock returns. While times with exceptionally positive or negative message contents, do not significantly impact the effects on stock returns.

Table of Contents

1	Introduction	3
2	Literature Review & Hypothesis Development	4
2.1	Literature Review	4
2.2	Hypothesis Development	6
3	Methodology	7
3.1	Data Extraction & Dataset Construction	7
3.2	Variable Development.....	8
3.2.1	<i>Dependent Variables</i>	8
3.2.2	<i>Independent Variables</i>	9
3.3	Model Estimation & Regression Methods	12
4	Results	15
4.1	Summary Statistics	15
4.2	Primary Regression Results	16
4.3	Robustness Checks	19
4.4	Heterogeneity Analysis: Time & Industry Sector	22
4.5	Exploratory Analysis: 1-Day Lead Stock Returns	25
5	Discussion & Conclusion	27
6	References	29
7	Appendices	31
7.1	Appendix A: OLS Assumptions.....	31
7.2	Appendix B: Heterogeneity Analysis.....	34
7.3	Appendix C: Documentation of Generative AI Usage.....	35

1 Introduction

In this current digital age, investing in the stock market has become increasingly more accessible to all people across society. Leading to an increase of starting investors, with varying degrees of knowledge and experience in financial markets, that invest in stocks, cryptocurrencies and bonds. The increasing number of new investors, potentially lacking knowledge and experience in financial markets, could be one of the reasons for the increasing popularity and activity on internet stock market message boards. Internet message boards are online platforms on which users can freely share their thoughts, knowledge and questions with peers to strengthen or question their own beliefs and increase their knowledge. One of such internet stock market message boards is StockTwits, the only message board considered in this thesis, which provides a distinct page for individual firms on which users can post or react to messages and events. Besides sharing financial beliefs, StockTwits' platform can also be used to criticize or defend political beliefs possibly related to the firms at which the messages are directed. In times of political turmoil, fueled by for example the 2024 US presidential elections and multiple tariff announcements made by the president of the United States, and figurehead of the 'Make America Great Again' (MAGA) movement, Donald J. Trump such discussions tend to appear more frequently. Existing empirical literature on the impact of alignments in favor or against a political movement, measured in the messages posted on internet message boards, on financial markets or stock returns is scarce. Therefore, this thesis aims to contribute to the knowledge of the influence of political movements on financial markets and its economic impact, by examining the effects message sentiment and MAGA associations of StockTwits messages have on stock returns of firms listed on the Russell 2000 index.

The analysis is performed by first assigning sentiments and MAGA associations to 69,708 StockTwits messages posted in the time period from 26/06/2023 to 07/03/2025. I ran a fixed effects OLS regression with heteroskedasticity-robust standard errors, clustered at firm level as the primary regression method, to account for OLS assumption violations. The regression includes the sentiment and MAGA categories as key independent variables and the daily log stock returns as the dependent variable. To test the robustness of both the findings and the assignment method of the MAGA associations, panel and Poisson regressions are carried out and an alternative assignment method is constructed and tested. Additionally, I performed a heterogeneity analysis,

aimed at the differences between various industry sectors and times with exceptionally positive or negative sentiments and MAGA associations. As well as an exploratory analysis, aimed at the lagged effects of the key independent variables by considering the 1-day lead log returns as the dependent variable.

I find that both negative message sentiments and positive MAGA associations significantly and immediately impact the daily stock returns of the Russell 2000 firms. With the negative sentiments reporting a negative effect and positive MAGA associations reporting a positive effect on stock returns. The effects are significantly influenced by the number of daily messages, market values and trade volumes, while some of these interactions are only marginally significant for the positive MAGA associations. On the other hand, positive sentiments and negative MAGA associations report no immediate significant impact on stock returns, but appear to possess a lagged effect on stock returns as they report significant and positive effects on the following day's stock returns. Times with exceptionally positive or negative sentiments and MAGA associations do not appear to significantly affect the message sentiment and MAGA association relationships with stock returns. However, various industry sectors do indicate to have an affect these relationships, especially the interactions with the positive MAGA associations report significance for all but one of the industry sectors considered in this thesis.

2 Literature Review & Hypothesis Development

2.1 Literature Review

Existing literature regarding the effect of the content and sentiment of messages posted on internet message boards generally reaches a consensus in that the messages significantly affect a variety of behavioral and economic aspects. Sabherwal et al. (2011) for instance, found evidence of a pump and dump pattern related to the messages posted on TheLion.com. Indicating a tendency that internet message boards such as TheLion.com and StockTwits can be used for herding to influence the trading activity and behavior of investors and drive up the short-term price of a stock. Regarding the impact on economic aspects, Antweiler & Frank (2004) studied the effect messages posted on Yahoo! Finance and Raging Bull have on market volatility, trading volume and stock returns. They found evidence that the sentiment of these messages can help predict the market volatility and that a disagreement visible on these internet stock message boards significantly

increases the trading volume of the related stocks. More directly related to this thesis Antweiler & Frank (2004) also found that the messages and their sentiments possess a significant effect on stock returns of firms listed on the Dow Jones Industrial Average (DJIA), but have a limited economic impact. Similar evidence is found by Tumarkin (2002) in their study, who found that days with strong positive sentiments portrayed in the messages posted on Raging Bull are preceded by an abnormal increase in stock price, indicating a significant relationship between the two variables. Further evidence for the effect the message sentiments and the number of messages posted on the HotCopper platform have on stock returns is presented by Leung & Ton (2015), in their study of Australian small-cap stocks. They found that both sentiment and the number of messages is significantly and positively correlated to the stock returns of these Australian firms (Leung & Ton, 2015). Leading to the expectation that a similar result is found in this thesis, since the smaller-cap Russell 2000 firms possess similarities to the Australian small-cap firms.

To further extend the knowledge about the effects of the content of messages on internet stock message boards and to align this thesis to the somewhat recent developments in the political landscape, this thesis examines the impact associations to a political movement displayed in the StockTwits messages have on stock returns, focusing on the MAGA movement in particular. Existing literature regarding this particular effect is scarce, however Al-Maadid et al. (2020) found slight evidence, in their study about the effects business and political news have on stock returns, of a causal relationship between political news and stock returns of firms located in Gulf Cooperation Council (GCC) countries. Karime & Sayilir (2019) found similar evidence in their study of political news significantly affecting both stock returns and market volatility. These studies show the tendency of politics, with its news and potentially views or beliefs, having an impact on stock markets and returns. Furthermore, Bollen et al. (2011) in their study found that the mood of the public, measured on Twitter, is able to accurately predict daily price movements of the DJIA. Where Bollen et al. (2011) shows that the collective mood is able to predict price movements, collective views or beliefs based on a political movement as examined in this thesis, could potentially possess a similar effect. Focusing more directly on the MAGA political movement and its widely regarded figurehead Donald J. Trump. Wagner et al. (2018) in their study found evidence that Donald Trump was able to significantly impact stock prices through his presidential election of 2016. Seeing as the short-style StockTwits messages considered in this thesis are similar to tweets posted on Twitter, both studies of Born et al. (2017) & Burggraf et al.

(2020) found evidence of Donald Trump significantly impacting the financial markets through his tweets. Where Born et al. (2017) found a significant impact of Donald Trump's tweets on the short-term stock returns & trading volume of the firms Donald Trump tweeted about, Burggraf et al. (2020) found a similar significant impact on S&P 500 returns. Due to the significant effects the figurehead of the MAGA movement and the collective mood of Twitter messages have on financial markets and stock returns I suspect the MAGA associations visible in the StockTwits messages, to have a similar effect on the stock returns of Russell 2000 firms.

2.2 Hypothesis Development

The existing literature listed and detailed in the previous section all show evidence of sentiments observed in messages posted on internet message boards and the messages themselves, having a significant impact on different aspects of financial markets like stock returns and investor behavior. The question then arises of what drives these effects? A possible answer is offered by Shiller (2017 & 2020), who argues that narratives, being simple stories or explanations of events, have the potential to go 'viral' and influence people's investment behaviors to create a feedback loop. Where these affected behaviors influence the financial markets and the financial markets in turn reinforce or weaken the initial narratives (Shiller, 2017 & 2020). Furthermore Shiller (2017 & 2020) argues that narratives can better explain economic crises or anomalies than economic fundamentals, reiterating the explanatory power of narratives regarding financial markets. The ability of narratives influencing financial markets through investor behavior could explain the observed impact of message sentiment and MAGA association on stock returns. Since the sentiment of investors, observed in the messages, might act as an indicator for investor behavior. Investor sentiment and behavior themselves also appear to impact stock returns, as evidence is found by Baker & Wurgler (2007) and Greenwood & Shleifer (2014). Who respectively find that investor sentiment is able to predict stock returns (Baker & Wurgler, 2007) and investor expectations of returns are inconsistent with rational models and are thus possibly influenced by narratives or sentiment (Greenwood & Shleifer, 2014).

Overall, the effects found in the existing literature regarding the content, sentiment and political association of messages posted on internet message boards and the theories potentially explaining these effects provided by Baker & Wurgler (2007), Greenwood & Shleifer (2014) and Shiller (2017 & 2020) lead to the expectation of both the sentiment and MAGA association

observed in StockTwits messages having a significant impact on the stock returns of Russell 2000 firms. These expectations are summarized in the following two hypotheses, which are tested in the remainder of this thesis:

- **H1:** The sentiment of the messages on the StockTwits platform significantly affect the stock returns of Russell 2000 firms.
- **H2:** MAGA associated StockTwits messages significantly affect the stock returns of Russell 2000 firms.

3 Methodology

This section examines and details the methodology used in the different analyses of this thesis. Starting with the data gathering and dataset construction process. After that the development of key variables is detailed and concludes with a description of the regression models and estimation methods used in the different empirical analyses conducted in this thesis.

To assist me in the coding process of the message extraction model, MAGA dictionary construction and sentiment prediction, I enlisted the use of the generative AI tool ChatGPT. A detailed overview of the use of generative AI in this thesis is available in Appendix C.

3.1 Data Extraction & Dataset Construction

Before any data is gathered and the dataset can be constructed, a random draw of 200 firms is taken out of the full list of approximately 2000 firms listed on the Russell 2000 index. For each of these 200 firms the financial variables including daily closing stock price, market value, trade volume and dividend yield are extracted from the LSEG Workspace with the first and latest available dates being 26/06/2023 and 07/03/2025. These dates therefore also represent the start and end date of the time frame considered in this thesis.

The messages considered in the dataset consist of messages posted publicly on the StockTwits platform. This platform allows users to post messages on dedicated pages for individual firms, similar to Yahoo! Finance. Since there is no dataset publicly available containing such messages, I extracted the messages manually with the use of the Application Programming Interface (API) provided by StockTwits to adhere to their user & privacy guidelines. This API

however does possess some limitations, including a limit on the number of messages that can be extracted each hour and not being able set a time frame of messages to be extracted. To account for these limitations multiple holds were built into the code to not hit the extraction limit and the model is set up to extract 35 pages of messages for each of the 200 firms from the random sample. Each page consists of up to 30 messages, but the exact number varies between individual firms and pages leading to a scraping process taking over 6 hours with a total number of extracted messages reaching 133,228. This total number of observations however, includes messages posted before the time frame considered in this thesis, weekend dates on which no meaningful financial data is available and multiple dates for which one or more values of the financial variables is missing. Observations consisting of these dates are consequently dropped from the dataset. On top of that, some firms contained only a very limited number of messages and therefore I dropped the firms possessing less than 50 messages, which is a practice similar to the one used by Leung & Ton (2015) & Tumarkin (2002). Leading to a final dataset consisting of 69,708 observations of 181 individual firms.

3.2 Variable Development

3.2.1 Dependent Variables

The different analyses conducted in this thesis consider 2 key dependent variables, being the daily log returns for the primary model and the 1-day lead log returns in the exploratory analysis. The log returns are calculated by measuring the difference in daily closing stock prices between day t and day $t-1$, following the logarithmic formula depicted in equation 1¹.

$$R_t = \ln(1 + r_t) = \ln\left(\frac{P_t}{P_{t-1}}\right) \quad (1)$$

Where R_t is the daily log stock return on day t , r_t is the simple daily stock return on day t , P_t is the stock price on day t and P_{t-1} is the stock price on the previous day. The log returns are used over the simple stock returns due its statistical properties being more favorable for empirical analysis. The log returns for example tend to approach normality more closely and offer symmetry for both positive and negative returns. The use of log returns over simple returns is therefore assumed to

¹ Campbell et al. (1997) provide a similar equation (1.4.5) and further explain and detail its implications.

enhance the robustness of the empirical analysis. When computing the log returns from the simple returns, +1 is added to the simple return to compute the growth factor and account for the possibility of returns being negative. When computing the log returns directly from the stock price, as used in this thesis, this addition is not necessary since it is already ‘inside’ the price formula and price cannot (usually) be negative, as is also detailed by Campbell et al. (1997).

To explore the possible presence of lagged market reactions or lagged investor perception, the 1-day lead log return is used to find the effect of key independent variables on the following day’s stock returns. The formula looks similar to equation 1, with the main difference being the inclusion the following day’s stock return over the previous day’s stock return as is portrayed in the formula of equation 2. Where R_{t+1} is the daily log stock return of the following day, r_{t+1} is the simple daily stock return of the following day and P_{t+1} is the stock price of the following day.

$$R_{t+1} = \ln(r_{t+1}) = \ln\left(\frac{P_{t+1}}{P_t}\right) \quad (2)$$

Additionally, the use of a Poisson (PPML) regression as a robustness check requires the dependent variable to be exclusively positive. However, since stock returns can be negative as well, each value is shifted by adding a constant. This constant consists of the minimum log return value of (-)2.703, increased by 0.01. An identical process is followed for the 1-day lead log return values used in the Poisson regression. This process is further detailed in the model estimation & regression methods section.

3.2.2 *Independent Variables*

To enhance the interpretability of the interaction effects the continuous independent variables used in this thesis are centered, meaning that each value is subtracted by the mean of the variable. Doing this allows me to interpret the interaction terms as the effect when the regarding variable remains equal to the mean, opposed to remaining equal to zero which is highly unlikely considering the control variables include market value & trade volume. On top of the control variables, multiple fixed effects variables are included in the model to control for unobserved variations in firm, time and industry sector. These fixed effects variables include firm ticker, date and industry sector.

Due to time and resource constraints, I am using a sentiment prediction model which is different to the Naïve Bayes learning model often used in existing literature by for instance

Antweiler & Frank (2004) and Leung & Ton (2015). In this thesis I implement a word matching model aiming to match the individual words of the messages to words included in multiple existing dictionaries with sentiment valuations assigned to those words. The main dictionary used in this matching model is the master dictionary constructed by Loughran & McDonald (2024), who compiled their dictionary starting in 2011 and contains a wide variety of both common words and words often used in finance all with matching sentiment scores. Due to the inclusion of a large quantity of words used in the finance field, I extended the dictionary with the AFINN dictionary created by Nielsen (2011). The AFINN dictionary is not directed at financial language, but more general or common language suitable for short-style messages like the ones found on Twitter and, more appropriate for this thesis, StockTwits. On top of that, I found when examining the messages in the dataset that a significant quantity of messages contained one or multiple emojis. To not lose the information regarding the sentiment of the message implemented in these emojis, I further extended the dictionary to include the emoji sentiment dictionary created by Novak et al. (2015) as well. Before being able to match the words of the messages to the ones in the dictionary, the messages have to be cleaned by for instance removing capitalization, punctuation, stop words and unnecessary white space. After cleaning the messages are tokenized, meaning that each word in the cleaned messages is treated as an individual unit, opposed to having a long string message, to make it easier to match individual words to the dictionary. Additionally, the emojis present in the messages are tokenized in a similar way. After having cleaned and tokenized the content of the messages, each individual word of the messages is attempted to be matched to the full dictionary and is assigned either a positive, negative or neutral sentiment in accordance with the values of the dictionary. Finally, all sentiments of the words in individual messages are aggregated to assign a single sentiment of either positive, negative or neutral to each message, similar sentiment assignments are used by Antweiler & Frank (2004).

The construction of the MAGA association variables implements the same word matching model as outlined for the message sentiments. However, there is no dictionary directed at MAGA associations publicly available leading to me manually creating such a dictionary with the assistance of generative AI. By examining news articles, social media messages and political announcements, a dictionary consisting of a total of 300 terms assigned as either pro or anti MAGA is constructed. These terms range from often used slogans like ‘Make America Great Again’ to high profile individuals like ‘Elon Musk’ & ‘Kamala Harris’ and derogatory terms like ‘Sleepy

Joe'. After matching the cleaned and tokenized words of the messages to the newly constructed MAGA dictionary an aggregated MAGA association is assigned to each message of either positive, negative or neutral. For both sentiment and MAGA variables the neutral category is left out of the regression models to serve as the reference category. Figure 1 displays example messages which have been assigned either a positive or negative sentiment and MAGA association.

FIGURE 1. EXAMPLE MESSAGES: SENTIMENTS & MAGA ASSOCIATIONS.

Positive Sentiment (*Ticker: ACAD*)

“This was a good report. They are firing on all cylinders! They are still generating almost \$1B in revenue, profitable have almost \$500M in cash. I am buyer after hours. This one opens higher after the call.”

Negative Sentiment (*Ticker: ATOM*)

“How much do you want to bet that he hides behind NDA's. Ya'll should ask for an independent third party accounting firm to substantiate the validity of those NDA's. Losses rise yet again this quarter.”

Positive MAGA Association (*Ticker: ACAD*)

“Trump 2024! Make the weak minded Stocktwits users strong again! You all know who I'm referring too!”

Negative MAGA Association (*Ticker: BCAB*)

“It has not started any war but this time it will destroy the world economy and start the third world War.. He's out of his mind now.. They went from someone with senile dementia, to a crazy person out of their mind.. What a great country.”

Notes: This figure presents example messages from the dataset, showing a clear example for each positive and negative message sentiments and MAGA associations. The ticker indicates the Russell 2000 firm at which the StockTwits message was directed and posted on its individual StockTwits page.

Additionally, to test the robustness of the method used to assign MAGA associations an alternative method of assigning is carried out. This alternative method assigns MAGA associations based on industry sectors which are considered to be positively related to MAGA. These industries have been identified by news outlets like Bloomberg (Lucey, 2025) and the Financial Times (Murphy et al., 2024). The industry sectors assigned to be positively related to MAGA are the energy, finance, industrials, technology and consumer discretionary sectors. A dummy variable is created which assigns value 1 to all observations including a MAGA positive sector and value 0 otherwise.

Finally, the independent variables used in the heterogeneity analyses include a category variable of the industry sectors and four time-spike dummies, which indicate a 5-day period around the spikes visible in average daily sentiment and MAGA association portrayed in figure 2. A 5-day period is used to include possible spillover effects and market delays, since not all investors might be able to perceive the effects of these spikes immediately and therefore the market might not be able to react immediately. Regarding the two MAGA spikes, the first spike occurs around October 28th 2024 which is a little over a week prior to the 2024 US election day of November 5th. The second spike occurs around March 7th 2025, which is a few days after the import tariffs on Canada, Mexico and China imposed by the Trump administration go into effect on March 4th and are postponed again on March 6th. The closeness of the spikes to these politically charged dates, could provide evidence for an effect such political announcements has on the MAGA related sentiment of the public, observable in the StockTwits messages.

3.3 Model Estimation & Regression Methods

This section will detail the model used to test the hypotheses described in the hypothesis development section, by first detailing the key variables used for each separate hypothesis before combining them into a full model used for the primary and robustness regressions. Finally, the model used for the alternative regressions will be examined as well as the lead and heterogeneity models.

Starting with H1, which focusses on the effect message sentiment has on stock returns. As is the case for both hypotheses in this thesis, the dependent variable is the daily log stock return. With the key independent variables being the dummy variables for message sentiment. Where both positive and negative sentiments are added to model and the neutral sentiment dummy is left out as the reference category. Similarly, for H2 dummy variables are included to find the main effects MAGA associations have on stock returns. Again, the neutral category is left out of the model to act as the reference category. Additionally, to explore if the main relationships of the hypotheses are affected by firm characteristics and the number of daily messages, which are included as control variables, interactions are added to the model between the key variables; message sentiments & MAGA associations and the control variables; message count, market value, trade volume & dividend yield. Similar interactions are added between the key variables and between the key variables and sector categories, used for the heterogeneity analysis. Finally, multiple fixed effects

are included in the model to account for unobserved heterogeneity regarding firm, time and industry sector. The complete scientific formula for this model is reported in equation 3.

$$\begin{aligned}
 R_{it} & & (3) \\
 &= \beta_1 \times Sent_{pos,it} + \beta_2 \times Sent_{neg,it} + \beta_3 \times MAGA_{pos,it} \\
 &+ \beta_4 \times MAGA_{neg,it} + \beta_5 \times Mes_C_{it} + \beta_6 \times MV_{it} + \beta_7 \times TV_{it} \\
 &+ \beta_8 \times Div_Yield_{it} + \beta_9 \times Interactions_{it} + \gamma_i + \delta_t + \theta_s + \varepsilon_{it}
 \end{aligned}$$

Where R_{it} denotes the daily log return for firm i at time t . The variables $Sent_{pos,it}$ and $MAGA_{neg,it}$ represent the positive and negative categories regarding message sentiments and MAGA associations. Mes_C_{it} indicates the daily message count and MV_{it} , TV_{it} and Div_yield_{it} represent the control variables for the firm characteristics; market value, trade volume and dividend yield respectively. $Interactions_{it}$ denotes the multiple interaction terms of the key independent variables and the control variables as well as the sector categories. γ_i , δ_t and θ_s represent the fixed effects included in the model, which are respectively firm, time and industry sector. Finally, ε_{it} denotes the error term of the model.

The regression method used to estimate the model depicted in equation 3 is an OLS regression. However due to the high dimensionality of the model and the inclusion of a large number of interaction terms, a FEOLS² (fixed effects ordinary least squares) regression, which is an OLS method capable of estimating a large quantity of interactions efficiently, is conducted to make use of its additional power over the standard OLS regression. The complete OLS assumption tests and graphs are reported in Appendix A, with the main takeaways being that these assumption tests are conducted on a simplified model excluding the multiple interaction terms and fixed effects and that violations are observed for the assumptions of homoskedasticity and no autocorrelation. To account for these violations the OLS model is adjusted to use heteroskedasticity-robust standard errors, clustered by firm.

To test the robustness of the main findings and their sensitivity to model estimations, I also ran the primary model using a panel (PLM)³ and Poisson (PPML)⁴ regression. The panel model resembles the OLS model, with the main differences being the way it handles the fixed effects.

² The FEOLS regression conducted in this thesis is part of the 'fixest' R package.

³ The PLM regression is part of the 'plm' R package.

⁴ The FEPOIS (fixed effects Poisson) regression is part of the 'fixest' R package.

Where the OLS and Poisson models estimate the variation both within and between firms and the panel regression only accounts for within firm variation. Additionally, the panel method is unable to include multiple observations per firm on the same date. Since the original dataset contains message level data, the same dates can have multiple observations for individual firms. To account for this constraint, I aggregated the sentiments and MAGA associations to only 1 value for each date. Therefore, reducing the number of observations in the original dataset from 69,708 to 19,780. To still account for the violations in the assumptions of homoskedasticity and no autocorrelation the panel and Poisson regression methods use robust standard errors, clustered by firm as well. A Poisson regression is mainly aimed to be used on a positive count dependent variable, however as argued by Silva & Tenreyro (2006) it can also be used on a positive continuous dependent variable. Where the log stock returns used in the OLS and panel regressions can have negative values, the return values are shifted to be exclusively positive and thus suitable for a Poisson regression.

Further testing the robustness of the method used to assign MAGA associations and sensitivity of its results to a different method, an alternative method of assigning MAGA associations, outlined in the variable development section, is used in a separate regression. Since the new MAGA dummy variable is based on industry sector and on firm level, the fixed effects of firm and industry sector are removed from the model due to them being multicollinear to the new MAGA variable. Other than that, the only difference to the primary model is the replacement of the positive and negative MAGA variables with the new MAGA dummy variable.

The model and regressions used in the exploratory analysis regarding the 1-day lead log return mainly uses the same model as depicted in equation 3 and the panel & Poisson regressions above. The only difference being the dependent variable, which will change from the daily log return in the main model, to the 1-day lead daily log return. Similar to the main model robust standard errors, clustered by firm are used to account for the violations in OLS assumptions.

To explore possible differences in effect between different industry sectors and specific time periods, a heterogeneity analysis is carried out in this thesis as well. Regarding the differences between industry sectors, the sectors have been added as categorical variables in the primary regression model. Testing the heterogeneity for specific times however, is done by first identifying differences in sentiment and MAGA associations through a visual inspection of the daily averages over time. After creating dummy variables of these spikes, the interactions between these spike dummies and key independent variables are added to the existing primary model. Leading to the

Interactions_{it} denotation of equation 3 being extended to also include the time-spike and key variable interactions. The regression methods used for this extended model are equal to the OLS, panel and Poisson regressions used to test the main hypotheses and therefore also use robust standard errors, clustered by firm to account for the violations in the assumptions of homoskedasticity and no autocorrelation.

4 Results

This section reports the empirical results of the analysis outlined in the methodology section. Starting with the summary statistics to provide an overview of the key variables used in the regressions. After that the results of the primary regression model are reported, followed by the robustness regression results and the alternative analyses, including the exploratory and heterogeneity specifications.

4.1 Summary Statistics

Table 1 reports the summary statistics of all the key variables used in the primary regression model. Both average daily log returns and lead log returns are close to 0 with a standard deviation of respectively 0.212 and 0.082. Such a high volatility (SD) could indicate the presence of outliers in the dataset, which is already accounted for by the high number of observations and the use of robust and clustered standard errors at the firm level. The mean of the positive sentiment, being 0.268, indicates that around 27% of the messages included in the dataset possess a positive sentiment. For the negative sentiment this value is 21%, meaning that the remaining 52% of the messages possess a neutral sentiment. With regard to the MAGA associations, these values are 3.8% for positive associations and 3% for negative associations. Indicating that only approximately 7% of the messages in the dataset can be attributed either a positive or negative MAGA association. Although this value is seemingly low, it is to be expected since the StockTwits platform is not directly focused on politics and the presence of a political movement in its messages is therefore expected to be of a secondary nature. The minimum and maximum value of the sentiment and MAGA variables are all 0 and 1 respectively, since these are dummy variables which can only contain either of those values. On average each firm in the dataset contains approximately 36 daily messages, with a minimum of 1 and a maximum of 449 messages. Indicating a wide variation in

the number of messages between firms and dates. A similar variation is observable in the control variables being the market value, trade volume and dividend yield.

TABLE 1. SUMMARY STATISTICS.

Variable	Mean	SD	Min	Max
<i>Log Return</i>	-0.007	0.212	-2.703	0.693
<i>Log Return (1-Day Lead)</i>	-0.002	0.082	-2.703	0.693
<i>Sentiment_{pos}</i>	0.268	0.443	0.000	1.000
<i>Sentiment_{neg}</i>	0.210	0.407	0.000	1.000
<i>MAGA_{pos}</i>	0.038	0.190	0.000	1.000
<i>MAGA_{neg}</i>	0.030	0.172	0.000	1.000
<i>Message Count</i>	36.152	70.710	1.000	449.000
<i>Market Value</i>	1504.955	2142.174	2.100	14298.940
<i>Trade Volume</i>	17236.766	19099.756	1.000	99038.000
<i>Dividend Yield</i>	0.752	2.335	0.000	22.640

Notes: This table reports the summary statistics of the key and control variables used in the analyses. With the statistics reported in the columns representing the mean, standard deviation, minimum and maximum values respectively. The sample consists of 69,708 message level observations over the period from 26/06/2023 to 07/03/2025. The log returns are computed as the log difference of the daily closing prices of firms listed on the Russell 2000 index. Sentiment and MAGA variables are constructed using a dictionary-based word matching analysis of StockTwits message content.

4.2 Primary Regression Results

The second column of table 2 reports the results of the primary OLS regression model. Overall the primary model reports an adjusted R^2 of 0.6618, which is seemingly high for a daily return model, but could be explained by the inclusion of firm, time and sector fixed effects. The variables are all included in the same primary model but the variables are grouped by their corresponding hypothesis in panel A and B, the interaction terms between the key variables are depicted in panel C and the control variables are reported in panel D to increase the readability of table 2. Starting with H1, a negative sentiment of messages is negatively and significantly (-0.02905) correlated with the daily stock returns at a 1% significance level. On the other hand, a positive sentiment (-0.00559) is not significantly correlated with the stock returns. Therefore, only partial support is observed regarding H1, that the sentiment of messages significantly affects the stock returns of Russell 2000 firms. These results thus also partially align with Antweiler & Frank (2004), who found in their study that the sentiment of messages have a significant effect on the stock returns of DJIA firms. On top of that Antweiler & Frank (2004) found that the economic impact of this effect is limited, a similar result is observed in this thesis since the significant coefficient for negative sentiments is only approximately (-)3%. Similar to Antweiler & Frank

(2004), Leung & Ton (2015) found a significant effect in their study aimed at the impact of message sentiment and message volume on stock returns of Australian small-cap firms. This thesis on the other hand offers a more nuanced view, indicating that only negative sentiments possess a significant effect on daily stock returns. These current findings could also be linked to the study of Sabherwal et al. (2011), who found evidence for a pump and dump pattern related to messages posted on internet stock market message boards. While their study is mostly aimed at positive buy messages, the findings of this thesis indicate a significant effect of negative sentiments. Potentially indicating that negative messages possess a similar effect on investor behavior through fear for example. Similarly, Tumarkin (2002) found a significant and positive relationship between strong positive sentiments and an abnormal increase in stock returns in their study. The significant interactions between the sentiments and market value, indicate that the negative effect of negative sentiments on daily stock returns diminishes as the market value of a firm increases ($2.69e-06$). Which could mean that larger firms are potentially less vulnerable to the negative effects of negative messages than smaller firms are. For positive sentiments the interaction effect ($-1.00e-06$) appears to be reversed. The diminishing effect of negative sentiments on stock returns for larger firms aligns with the findings of Leung & Ton (2015), who again found significant effects of message sentiment on the stock returns of smaller-cap firms. A higher trade volume ($-3.90e-07$) on the other hand significantly amplifies the negative effect of negative sentiments on daily stock returns, while diminishing the insignificant negative effect of positive sentiments ($3.24e-07$). Potentially indicating that firms with a higher trading volume are more vulnerable to investor sentiments reflected in the messages. The interaction between the sentiments and the number of daily messages, portrays a similar effect. With the message count (-0.00011) significantly amplifying the already negative effect of negative messages. Indicating that firms with a greater message activity are more sensitive to the negative sentiments portrayed in the messages.

Continuing with H2, a significant correlation is observed between positive MAGA associations and daily stock returns (0.05199) at a 1% significance level. On the other hand, messages negatively associated with MAGA (0.02563) do not possess a significant correlation. These findings possibly indicate that investors slightly favor firms which are in some way associated or believed to be associated with the MAGA movement. Overall these results, due to only the positive association being significant, again offer only partial support for H2 that Russell 2000 firms which are associated with the MAGA movement possess significantly different stock

returns. Comparing the results of this thesis to existing literature shows that the results partially align with Born et al. (2017), Burggraf et al. (2020) & Wagner et al. (2018), who all found in their respective studies that the figurehead of the MAGA movement, Donald J. Trump, significantly impacted stock prices or returns through his presidency and tweets. The results of this thesis however, indicate that even though the MAGA movement without direct involvement from its figurehead possesses a significant impact on daily stock returns, the economic impact is limited to approximately 5%. The interactions between a positive association with the MAGA movement and both market value ($5.46e-06$) and message count (0.00050) show marginal significance at a 10% level, indicating that the positive effect of positive MAGA associations is slightly amplified for larger firms and firms possessing a greater message activity. For firms with a higher trade volume ($-1.51e-06$) on the other hand, this positive effect appears to be slightly diminished.

Looking at the interactions between the key variables of sentiment and MAGA association, the only significant correlation is observed between the positive sentiment and negative MAGA association (-0.01025) variables. Indicating that the already negative effect of positive sentiments (-0.00559) is further amplified by the presence of negative MAGA associations observable in the messages. Which could mean that investors and financial markets discount the value of positive messages more if they are related to firms opposing the MAGA movement.

Overall, H1 is only partially supported by the results since only messages with a negative sentiment are significantly correlated with the daily stock returns at a 1% level. Similarly, H2 is only partially supported by the results as well, since only a positive association with the MAGA movement is significantly correlated with the daily stock returns, at a 1% significance level. Meaning that neither of the two hypotheses can be accepted, but indications are present in the results to partially support both hypotheses of this thesis. Additionally, the economic impact of the effects is limited. For instance, the effect of negative sentiments on stock returns is approximately -3%, meaning that on dates on which there is a negative sentiment present the stock returns are expected to be approximately 3% lower than on dates with no negative sentiments present.

TABLE 2. MAIN REGRESSION & ROBUSTNESS REGRESSION RESULTS.

Variable	OLS	PLM	PPML
Panel A: H1 Message Sentiment			
$Sentiment_{pos}$	-0.00559 (0.01917)	0.01529 (0.01760)	-0.00147 (0.00774)
$Sentiment_{neg}$	-0.02905*** (0.00862)	-0.09593*** (0.02775)	-0.01102*** (0.00303)
$Sentiment_{pos} \times Message\ Count$	-0.00003 (0.00002)	0.00096** (0.00049)	-0.00001 (9.61e-06)
$Sentiment_{neg} \times Message\ Count$	-0.00011*** (0.00004)	-0.00270*** (0.00084)	-0.00005*** (0.00002)
$Sentiment_{pos} \times Market\ Value$	-1.00e-06* (5.96e-07)	-2.84e-06*** (8.43e-07)	-3.55e-07 (2.22e-07)
$Sentiment_{neg} \times Market\ Value$	2.69e-06*** (8.84e-07)	1.22e-06 (1.38e-06)	9.61e-07*** (3.56e-07)
$Sentiment_{pos} \times Trade\ Volume$	3.24e-07*** (9.18e-08)	7.39e-07*** (2.23e-07)	1.19e-07*** (3.44e-08)
$Sentiment_{neg} \times Trade\ Volume$	-3.90e-07*** (1.21e-07)	-4.30e-09 (4.25e-07)	-1.27e-07** (5.00e-08)
Panel B: H2 MAGA Association			
$MAGA_{pos}$	0.05199*** (0.01438)	0.03319 (0.02130)	0.01860*** (0.00552)
$MAGA_{neg}$	0.02563 (0.02714)	-0.02557 (0.03298)	0.00944 (0.01033)
$MAGA_{pos} \times Message\ Count$	0.00050* (0.00030)	0.00054 (0.00056)	0.00022* (0.00013)
$MAGA_{neg} \times Message\ Count$	0.00013 (0.00011)	-0.00124 (0.00086)	0.00006 (0.00005)
$MAGA_{pos} \times Market\ Value$	5.46e-06* (3.25e-06)	3.40e-06* (1.92e-06)	2.25e-06* (1.30e-06)
$MAGA_{neg} \times Market\ Value$	-8.55e-07 (2.23e-06)	3.31e-07 (2.39e-06)	-2.54e-07 (8.59e-07)
$MAGA_{pos} \times Trade\ Volume$	-1.51e-06** (7.59e-07)	-9.77e-07** (4.94e-07)	-6.26e-07** (3.07e-07)
$MAGA_{neg} \times Trade\ Volume$	-6.64e-07** (3.00e-07)	1.33e-07 (7.50e-07)	-2.72e-07** (1.24e-07)
Panel C: Cross-Variable Interactions			
$Sentiment_{pos} \times MAGA_{pos}$	-0.00290 (0.00579)	-0.01218 (0.00778)	-0.00096 (0.00218)
$Sentiment_{neg} \times MAGA_{pos}$	0.01478 (0.00904)	0.00412 (0.00780)	0.00623* (0.00362)
$Sentiment_{pos} \times MAGA_{neg}$	-0.01025** (0.00478)	-0.00052 (0.00495)	-0.00382** (0.00182)
$Sentiment_{neg} \times MAGA_{neg}$	-0.00095 (0.00818)	-0.00703 (0.00563)	-0.00032 (0.00317)
Panel D: Control Variables			
<i>Message Count</i>	-0.00059 (0.00046)	0.00049 (0.00043)	-0.00025 (0.00019)
<i>Market Value</i>	0.00002*** (5.41e-06)	4.17e-06*** (1.22e-06)	7.97e-06*** (2.03e-06)
<i>Trade Volume</i>	1.83e-06*** (6.79e-07)	6.28e-08 (1.80e-07)	7.30e-07*** (2.80e-07)
<i>Dividend Yield</i>	0.00135 (0.00202)	-0.00080 (0.00058)	0.00056 (0.00083)
Firm Fixed Effects	Yes	Yes	Yes
Time Fixed Effects	Yes	Yes	Yes
Sector Fixed Effects	Yes	Yes	Yes
Observations	69708	19780	69708
Adjusted R^2	0.6618	0.0527	-0.0017

Notes: This table presents coefficient estimates and standard errors in brackets from the primary fixed effects OLS (FEOLS), panel (PLM) and Poisson (PPML) regressions. The dependent variable is the daily log stock returns of firms listed on the Russell 2000 index. To account for the possible negative values of the dependent variable it has been shifted to be exclusively positive in value for the Poisson regression only. The independent variables and their interactions are grouped in panels, but are all included in the same full model. With panel A and B corresponding with the key variables of Hypothesis 1 and 2 respectively. Panel C presents the interactions between the key variables and panel D presents the control variables. All regressions control for firm, time and industry sector fixed effects, indicated by 'Yes' in the fixed effects section. Heteroskedasticity-robust standard errors, clustered at firm level are used by all regressions to account for violations in the OLS assumptions of homoskedasticity and no autocorrelation. ***, ** and * indicate statistical significance at the 1%, 5% and 10% levels respectively.

4.3 Robustness Checks

Besides the primary regression results, table 2 also reports the results of the panel and Poisson regressions used as robustness checks in the third and fourth column respectively. Before continuing to the regression results its apparent that the number of observations in the panel regression (19,780), is much lower than the observations of the OLS and Poisson regressions, this is caused by the structure of the panel regression not supporting multiple observations with the same date and firm. Therefore, the observations have been aggregated to a single observation per

date for each firm reducing the total number of observations. The adjusted R^2 of the panel (0.0527) and Poisson (-0.0017) regressions, also report much lower values than the primary OLS regression. This is possibly due to the within R^2 and adjusted pseudo R^2 , used instead of the regular adjusted R^2 in respectively the panel and Poisson regression models. In the case of the fixed-effects panel regression model the fixed effects are not included in the variance explanation, but only in the estimation of the model. The panel regression also only considers the within firm variation opposed to the OLS regression, which also considers the between firm variation (Wooldridge, 2010).

Examining the results of the panel regression, it is observed that overall the signs of the coefficients remain similar to those in the primary OLS regression. With the only main differences in signs being observed in the negative MAGA association variable and interactions. However, none of these coefficients are significant in any of the two regressions. In terms of the magnitude of the coefficients the biggest difference is observed for the negative sentiment (-0.09593), which reports a greater negative effect in the panel regression. The significance of the negative sentiment holds across both regressions at a 1% significance level, enhancing its robustness. The significance of the positive MAGA association however, does not seem to hold for the panel regression. Similar changes in significance are observable in the interactions between negative sentiment and both market value and trade volume, as well as the interaction between positive MAGA association and message count. These changes in significance could indicate that these interactions and effects are sensitive to model specifications. This sensitivity appears to be mostly present for the panel regression, since the Poisson regression shows very similar results to the OLS regression. All signs are equal between the OLS and Poisson regressions, with the only difference in significance observed in the interaction between negative sentiment and positive MAGA association. Which is only marginally significant at a 10% level in the Poisson regression. Overall, the model seems to hold across different regression methods, especially the Poisson regression indicates very similar results to the OLS regression. With the main difference between regressions being the insignificant effect of positive MAGA association in the panel regression, opposed to the other two regressions.

To further examine the robustness of the effects of the MAGA associations, an alternative method of assigning MAGA associations is used in a separate model. The regression results of this alternative method are reported in table 3. At first, it's important to note that since the new variable regarding MAGA association is measured at industry sector level, the fixed effects regarding both firm and sector have been dropped from the model opposed to the primary model due to

multicollinearity with the new MAGA variable. The exclusion of these fixed effects could be the reason for the lower R^2 values regarding the OLS and panel regressions compared to the primary model. Another apparent difference to the primary model is the significance of positive sentiment (0.00540). This significance could be caused by the exclusion of the fixed effects, leading to positive sentiment potentially absorbing some of the variation otherwise accounted for by the excluded fixed effects. The most important difference observed in the results of the alternative method is the insignificance of the alternative MAGA variable (0.03418). Multiple possible reasons for this change in significance can be thought of; the method of assigning MAGA association could be less/more precise compared to the primary method or it could be caused by the variable absorbing the variation otherwise accounted for by the fixed effects. When looking at the different regression methods, very similar results to primary model are found. Where the panel regression appears to show some sign and significance changes and the Poisson regression is again very similar to the OLS regression. Overall, it's safe to say that the effects of MAGA association variables are sensitive to specific methods of assigning MAGA associations and therefore lead to differences in results between the methods and models.

TABLE 3. ALTERNATIVE MAGA SPECIFICATION REGRESSION RESULTS.

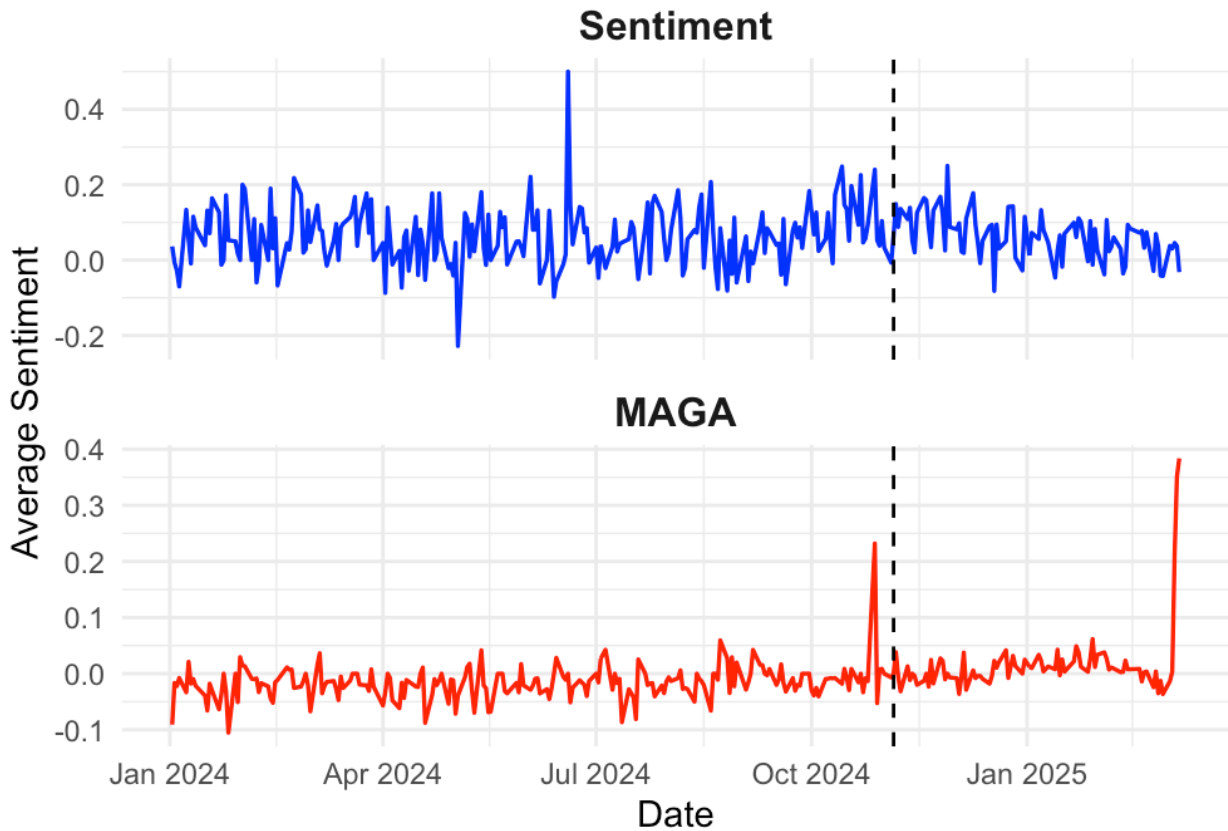
Variable	OLS	PLM	PPML
Panel A: H1 Message Sentiment			
$Sentiment_{pos}$	0.00540** (0.00255)	0.01308 (0.01225)	0.00213* (0.00119)
$Sentiment_{neg}$	-0.01026** (0.00414)	-0.08133*** (0.02710)	-0.00374** (0.00180)
$Sentiment_{pos} \times Message\ Count$	-0.00002 (0.00003)	-9.06e-07 (0.00039)	-7.74e-06 (0.00001)
$Sentiment_{neg} \times Message\ Count$	-0.00011* (0.00007)	-0.00233** (0.00091)	-0.00005 (0.00003)
$Sentiment_{pos} \times Market\ Value$	-1.25e-06* (7.16e-07)	-3.06e-06*** (8.88e-07)	-4.35e-07 (2.66e-07)
$Sentiment_{neg} \times Market\ Value$	3.56e-06*** (1.09e-06)	2.59e-06* (1.41e-06)	1.32e-06*** (4.27e-07)
$Sentiment_{pos} \times Trade\ Volume$	3.07e-07** (1.31e-07)	8.88e-07*** (2.36e-07)	1.08e-07** (4.93e-08)
$Sentiment_{neg} \times Trade\ Volume$	-4.07e-07** (1.76e-07)	-3.03e-07 (4.56e-07)	-1.37e-07* (7.76e-08)
Panel B: H2 Alternative MAGA Alignment			
$MAGA_{Alt}$	0.03418 (0.03023)	0.04079 (0.02757)	0.01495 (0.01520)
$MAGA_{Alt} \times Message\ Count$	0.00223 (0.00168)	0.00140 (0.00092)	0.00095 (0.00085)
$MAGA_{Alt} \times Market\ Value$	8.96e-06 (7.79e-06)	1.74e-06 (1.29e-06)	3.75e-06 (3.49e-06)
$MAGA_{Alt} \times Trade\ Volume$	-2.81e-06 (2.48e-06)	-8.15e-07 (5.07e-07)	-1.23e-06 (1.22e-06)
Panel C: Cross-Variable Interactions			
$Sentiment_{pos} \times MAGA_{Alt}$	-0.00366 (0.00293)	-0.00367 (0.00318)	-0.00149 (0.00126)
$Sentiment_{neg} \times MAGA_{Alt}$	-0.00632 (0.00443)	-0.00323 (0.00280)	-0.00232 (0.00182)
Panel D: Control Variables			
$Message\ Count$	-0.00202 (0.00168)	-0.00015 (0.00059)	-0.00088 (0.00086)
$Market\ Value$	-7.86e-06 (6.46e-06)	7.35e-07 (9.78e-07)	-3.33e-06 (3.08e-06)
$Trade\ Volume$	3.14e-06 (2.32e-06)	3.52e-07 (3.53e-07)	1.34e-06 (1.16e-06)
$Dividend\ Yield$	0.00086 (0.00154)	0.00018 (0.00016)	0.00041 (0.00073)
Firm Fixed Effects	No	No	No
Time Fixed Effects	Yes	Yes	Yes
Sector Fixed Effects	No	No	No
Observations	69708	19780	69708
Adjusted R^2	0.6432	0.0404	0.0006

Notes: This table presents coefficient estimates and standard errors in brackets from the fixed effects OLS (FEOLS), panel (PLM) and Poisson (PPML) regressions using an alternative method of assigning MAGA associations. The alternative method assigns firms to be positively/negatively associated with the MAGA movement, based on their corresponding industry sector. The dependent variable is the daily log stock returns of firms listed on the Russell 2000 index. To account for the possible negative values of the dependent variable it has been shifted to be exclusively positive in value for the Poisson regression only. The independent variables and their interactions are grouped in panels, but are all included in the same full model. With panel A and B corresponding with the key variables of Hypothesis 1 and 2 respectively. Panel C presents the interactions between the key variables and panel D presents the control variables. All regressions control for time fixed effects, indicated by 'Yes' in the fixed effects section. Firm and industry sector fixed effects are not controlled for in these regressions, because the alternative MAGA variable is measured at sector level and therefore multicollinear with the firm and sector fixed effects variables, indicated by 'No' in the fixed effects section. Heteroskedasticity-robust standard errors, clustered at firm level are used by all regressions to account for violations in the OLS assumptions of homoskedasticity and no autocorrelation. ***, ** and * indicate statistical significance at the 1%, 5% and 10% levels respectively.

4.4 Heterogeneity Analysis: Time & Industry Sector

Figure 2 illustrates the daily averages of both sentiment and MAGA association over time, where the dashed vertical line represents the US election day of November 5th 2024. Examining the sentiment, the average sentiment remains relatively stable over time, but a slight downtrend is visible close to the dashed vertical line. On top of that 2 spikes in sentiment are observable, the first being a negative spike and the second a positive spike. Indicating that for a short period of time the sentiment deviates from the average sentiment both negatively and positively. Regarding the MAGA association a similar stable trend over time is observable, with two positive spikes in

MAGA association, indicating a more positive attitude towards the MAGA movement for a short period of time. The first spike happens just a few days prior to the US election day, while the second happens a few days after the import tariffs on Canada, Mexico and China imposed by the Trump administration go into effect on March 4th and are postponed again on March 6th. Suggesting that political events like the election day and tariff implementation might inspire some political discourse observable in both sentiment and MAGA association displayed in the StockTwits messages. Table B1 of Appendix B reports the regression results of the four time-spike dummies and their interactions with the key variables of the primary model. Focusing on the time-spike interaction panel C of table B1, the only significant effect in the OLS regression is observed in the interaction of negative sentiment and the second MAGA spike or the MAGA spike possibly related to the tariff announcements (0.01378). The interaction is only marginally significant at a 10% level in the OLS regression and significant at a 5% level in the Poisson regression, however is insignificant in the panel regression. Indicating that a time of high positive MAGA associations, diminishes the negative effect of negative sentiments on stock returns. This effect however is sensitive to the regression method as significance increases or decreases dependent on the regression method. The only other significant interaction is found in the panel regression regarding positive MAGA association and the negative sentiment spike (0.05536). Indicating a possible amplification of the positive effect of positive MAGA associations on stock returns in times of exceptionally negative sentiments, which is highly dependent on the regression method used however. In general, since the great majority of the time-spike interactions are found to be insignificant, no true heterogeneity is found based on the results in times with exceptionally positive or negative sentiments and MAGA associations.

FIGURE 2. DAILY AVERAGE SENTIMENT & MAGA ASSOCIATION OVER TIME.

Notes: This figure presents the daily average sentiment and MAGA association based on StockTwits messages only. The dashed vertical line represents the 2024 US election day of November 5th 2024. Both the average sentiment and MAGA association show a relatively stable trend over time with multiple significant spikes in average. With only the first sentiment spike being negative. Four dummy variables are constructed to represent each of the four spikes, consisting of a 5-day period around the spike date. The regression results of the interactions between these spike dummies and the key variables of the hypotheses are presented in table B1 of Appendix B.

To explore the heterogeneity of effects on stock returns between different industry sectors, the results of the interactions between separate industry sectors and the key variables of the primary OLS model are reported in table 4. The industry sector interaction results from table 4, stem directly from the primary OLS regression model used in table 2, but are reported in a separate table to increase the readability of both tables. Starting with the positive sentiment interactions, no significant interactions are found for any of the different industry sectors. Negative sentiment on the other hand does have multiple significant interactions, with the energy (0.02574) and real estate sectors (0.03626) being the strongest at a 1% significance level. All significant sectors have a diminishing effect on the negative impact of negative sentiments on stock returns and thus potentially indicate a contradicting market reaction to negative investor sentiments. In contrast to positive sentiment, the positive MAGA association exhibits highly significant interactions for all

industry sectors except the consumer staples sector. The interaction coefficients are consistently negative, leading to these sectors diminishing the positive impact positive MAGA association has on stock returns. Similarly, all coefficients for the negative MAGA association are negative, with only the consumer staples sector (-0.06278) being significant at a 5% level, again diminishing the positive effect negative MAGA association has on stock returns. Possibly suggesting that the consumer staples sector is less sensitive to political alignment than the other industry sectors are.

TABLE 4. HETEROGENEITY ANALYSIS: INDUSTRY SECTOR RESULTS.

Industry Sector	<i>Sentiment_{pos}</i>	<i>Sentiment_{neg}</i>	<i>MAGA_{pos}</i>	<i>MAGA_{neg}</i>
<i>Consumer Discretionary</i>	0.00709 (0.01926)	0.00909 (0.00930)	-0.04826*** (0.01738)	-0.00786 (0.03106)
<i>Consumer Staples</i>	0.00946 (0.01912)	-0.00112 (0.01234)	-0.03341 (0.02023)	-0.06278** (0.02819)
<i>Energy</i>	0.01190 (0.01963)	0.02574*** (0.00958)	-0.06345*** (0.01976)	-0.02483 (0.02876)
<i>Finance</i>	0.00410 (0.01961)	0.01329 (0.01041)	-0.05549*** (0.02102)	-0.01625 (0.02887)
<i>Health Care</i>	0.00728 (0.01924)	0.01317 (0.00928)	-0.04763*** (0.01604)	-0.01274 (0.02763)
<i>Industrials</i>	-0.00392 (0.02018)	0.02097** (0.00934)	-0.05414*** (0.01780)	-0.00865 (0.02832)
<i>Real Estate</i>	0.00854 (0.01930)	0.03626*** (0.01250)	-0.05892*** (0.01555)	-0.01819 (0.02803)
<i>Technology</i>	0.00884 (0.01937)	0.01624 (0.01049)	-0.06215*** (0.01999)	-0.01369 (0.02882)
<i>Telecommunications</i>	0.00509 (0.01944)	0.02917** (0.01157)	-0.04571*** (0.01653)	-0.01902 (0.03047)
<i>Utilities</i>	0.00603 (0.02418)	0.01704* (0.01005)	-0.06764*** (0.01781)	-0.03570 (0.02978)
Observations				69708
Adjusted R^2				0.662

Notes: This table presents the coefficient estimations and standard errors in brackets from the industry sector interactions with the key independent variables of the primary fixed effects OLS (FEOLS) regression. The model used for these estimates is the exact same as the regressions presented in table 2, but the sector interactions are reported in a separate table to increase the readability of both tables. The dependent variable is the daily log stock returns of firms listed on the Russell 2000 index. The Basic Materials sector is not included in the table, since it acts as the reference category. The columns of the table represent the key independent variables of the two hypotheses. Since the exact primary model is used these estimations are controlled for firm, industry sector and time fixed effects and use heteroskedasticity-robust standard errors, clustered at firm level to account for violations in the OLS assumptions of homoskedasticity and no autocorrelation. ***, ** and * indicate statistical significance at the 1%, 5% and 10% levels respectively.

Overall, the limited number of significant interactions regarding the time-spike dummies suggests that there is little evidence of heterogeneity present in the effect of message sentiments and MAGA associations in times of exceptionally positive or negative average sentiment or MAGA association. In contrast, the industry sector results indicate that sectoral differences do influence the impacts of message sentiment and MAGA association on stock returns, in particular for negative sentiment and positive MAGA association. Confirming the presence of heterogeneity across different industry sectors in the relationships considered in this thesis.

4.5 Exploratory Analysis: 1-Day Lead Stock Returns

To explore the possible presence of lagged market reactions or lagged investor perception of message sentiments, the results of the lead regression model are reported in table 5. The adjusted R^2 values (0.2581, -0.0286 & -0.0064) of the 1-day lead model drop significantly compared to the primary model, possibly indicating that the primary model offers a better fit than the 1-day lead

model. Examining the key variables multiple divergences to the primary model are observed in the lead model. Not only do the signs of the coefficients of the message sentiments change from negative to positive, the positive sentiment (0.02375) also exhibits a strong significance in the lead model where the negative sentiment is strongly significant in the primary model. Possibly suggesting a lagged market reaction to positive messages and a more direct reaction to negative messages. Similar changes in sign are observed in the interactions of negative sentiments and both market value and trade volume, as well as positive sentiments and message count. Additionally, all these interactions are insignificant in the lead model where they are significant in the primary model. Focusing on the MAGA associations, the positive association (-0.00890) also changes its sign and loses its significance in the lead model. While the negative association (0.04544) in contrast, is found to be significant in the lead model only at a 5% level. The interaction terms of the MAGA associations mostly report similar signs to the primary model, with the only difference being the interaction of negative MAGA association and market value ($1.74e-08$). The interaction terms remain mostly similar in their significance levels to the primary model as well, with the most apparent difference being the reduction to insignificance of the positive MAGA association and market value interaction ($1.03e-06$). The interactions between the key variables in panel C follow the same trend and report a change in sign for the interaction between negative sentiment and negative MAGA association (0.00902). Additionally, this effect reports to be significant at a 5% level in the lead model, where the interaction between positive sentiment and negative MAGA association (-0.00007) is significant in the primary model. The robustness regressions of the lead model exhibit a similar trend to the robustness regressions of the primary model. Where mostly the same coefficient signs are reported for the panel and Poisson regressions, but the panel regression does portray some differences in significance levels opposed to the other regressions.

Overall, the multiple changes in sign and significance levels between the primary and lead models indicate the presence of variations in the examined effects. These variations are for one, a possible immediate effect of negative sentiments on stock returns, while positive sentiments display a lagged effect on stock returns. Secondly the changes in sign between models could indicate an immediate overreaction to for example positive MAGA messages followed by a correction of the stock returns visible in the lead model effects.

TABLE 5. 1-DAY LEAD REGRESSION RESULTS.

Variable	OLS	PLM	PPML
Panel A: H1 Message Sentiment			
$Sentiment_{pos}$	0.02375*** (0.00762)	0.04440*** (0.01713)	0.00879*** (0.00284)
$Sentiment_{neg}$	0.01708 (0.01929)	0.03437 (0.02312)	0.00630 (0.00714)
$Sentiment_{pos} \times Message\ Count$	3.70e-06 (8.18e-06)	0.00032 (0.00047)	1.43e-06 (3.03e-06)
$Sentiment_{neg} \times Message\ Count$	-8.26e-08 (9.87e-06)	0.00008 (0.00047)	2.55e-08 (3.67e-06)
$Sentiment_{pos} \times Market\ Value$	-3.27e-07 (2.23e-07)	1.80e-07 (5.83e-07)	-1.21e-07 (8.24e-08)
$Sentiment_{neg} \times Market\ Value$	-5.55e-07 (3.38e-07)	-1.47e-08 (6.36e-07)	-2.06e-07 (1.25e-07)
$Sentiment_{pos} \times Trade\ Volume$	2.19e-08 (3.67e-08)	7.30e-08 (1.29e-07)	7.89e-09 (1.35e-08)
$Sentiment_{neg} \times Trade\ Volume$	4.77e-08 (4.32e-08)	3.16e-07** (1.61e-07)	1.76e-08 (1.61e-08)
Panel B: H2 MAGA Association			
$MAGA_{pos}$	-0.00890 (0.02614)	0.01602 (0.01062)	-0.00336 (0.00964)
$MAGA_{neg}$	0.04544** (0.02077)	0.05719*** (0.01948)	0.01663** (0.00765)
$MAGA_{pos} \times Message\ Count$	0.00016* (0.00008)	0.00025 (0.00019)	0.00006* (0.00003)
$MAGA_{neg} \times Message\ Count$	0.00008 (0.00005)	0.00030 (0.00020)	0.00003 (0.00002)
$MAGA_{pos} \times Market\ Value$	1.03e-06 (7.91e-07)	1.94e-07 (1.15e-06)	3.71e-07 (2.90e-07)
$MAGA_{neg} \times Market\ Value$	1.74e-08 (8.49e-07)	1.01e-07 (8.33e-07)	8.25e-09 (3.15e-07)
$MAGA_{pos} \times Trade\ Volume$	-6.37e-07*** (1.93e-07)	-2.79e-07 (2.31e-07)	-2.32e-07*** (7.17e-08)
$MAGA_{neg} \times Trade\ Volume$	-2.90e-07** (1.29e-07)	-1.95e-08 (2.12e-07)	-1.07e-07** (4.82e-08)
Panel C: Cross-Variable Interactions			
$Sentiment_{pos} \times MAGA_{pos}$	-0.00074 (0.00398)	-0.00122 (0.00593)	-0.00027 (0.00147)
$Sentiment_{neg} \times MAGA_{pos}$	0.00020 (0.00357)	0.00434 (0.00720)	0.00008 (0.00132)
$Sentiment_{pos} \times MAGA_{neg}$	-0.00007 (0.00306)	-0.00126 (0.00395)	-0.00002 (0.00113)
$Sentiment_{neg} \times MAGA_{neg}$	0.00902** (0.00347)	0.00588 (0.00478)	0.00332*** (0.00128)
Panel D: Control Variables			
$Message\ Count$	-0.00015 (0.00010)	-0.00050 (0.00050)	-0.00006 (0.00004)
$Market\ Value$	-9.10e-06*** (2.71e-06)	-6.25e-06*** (1.44e-06)	-3.35e-06*** (9.96e-07)
$Trade\ Volume$	4.18e-07** (2.01e-07)	-4.33e-08 (1.24e-07)	1.55e-07** (7.48e-08)
$Dividend\ Yield$	0.00263*** (0.00075)	0.00162*** (0.00032)	0.00098*** (0.00028)
Firm Fixed Effects	Yes	Yes	Yes
Time Fixed Effects	Yes	Yes	Yes
Sector Fixed Effects	Yes	Yes	Yes
Observations	70402	19780	70402
Adjusted R^2	0.2581	-0.0286	-0.0064

Notes: This table presents coefficient estimates and standard errors in brackets from the 1-day fixed effects OLS (FEOLS), panel (PLM) and Poisson (PPML) regressions acting as an exploratory analysis. The dependent variable is the 1-day lead daily log stock returns of firms listed on the Russell 2000 index. To account for the possible negative values of the dependent variable it has been shifted to be exclusively positive in value for the Poisson regression only. The independent variables and their interactions are grouped in panels, but are all included in the same full model. With panel A and B corresponding with the key variables of Hypothesis 1 and 2 respectively. Panel C presents the interactions between the key variables and panel D presents the control variables. All regressions control for firm, time and industry sector fixed effects, indicated by 'Yes' in the fixed effects section. Heteroskedasticity-robust standard errors, clustered at firm level are used by all regressions to account for violations in the OLS assumptions of homoskedasticity and no autocorrelation. ***, ** and * indicate statistical significance at the 1%, 5% and 10% levels respectively.

5 Discussion & Conclusion

This thesis contributes to the existing empirical literature by providing a more nuanced view to the studies of for example Antweiler & Frank (2004) and Leung & Ton (2015) by reporting evidence for an immediate effect of negative message sentiments and a lagged effect of positive message sentiments on stock returns. Additionally, this thesis provides evidence for a possible effect of political movements, measured on internet message boards, on stock returns. Providing a

possible starting point for future research to improve upon and further explore its potential effects on financial markets.

While this thesis provides potentially valuable insights into the effects message sentiments and MAGA associations have on stock returns, it's important to acknowledge some limitations which might affect the robustness and generalizability of these findings. Firstly where existing literature often uses a learning model, like the Naïve Bayes model (Antweiler & Frank, 2004; Leung & Ton, 2015), to automatically assign sentiments and potentially MAGA or other political movement associations to individual messages, this thesis implements a more simplified word matching model. This word matching model is fully dependent on the dictionaries used and the sentiment values assigned to the words in the dictionaries. Additionally, the model is unable to interpret tones of messages or more nuanced writing styles like sarcasm and jokes, due to the model only being able to match exact words. A Naïve Bayes learning model could be more appropriate to interpret such messages correctly, dependent on the manually assigned training set of messages used for the learning model. Secondly the restricted scope of the dataset considered in this thesis, limited to 181 Russell 2000 firms over an approximately two-year time period, could provide difficulties in generalizing the findings to different time periods and indices. Thirdly the StockTwits messages included in this thesis do not offer detailed time-specifications, but only indicate the date on which the messages were posted. This could lead to messages, that were posted after the financial markets have closed, still contributing to the effects on closing stock returns measured on that same market day. Potentially impacting the effects on stock returns, while the true effects of these after-hours messages might be measured in the following day's stock returns. Finally, the potential presence of reverse causality is an important limitation. Where this thesis is aimed at the effects message board contents have on stock returns, the true causal effect could prove to be reversed, where the stock returns might affect the message board content.

Future research could improve on this thesis by increasing its scope to multiple indices and wider time periods, for example focusing on the MAGA movement during president Trumps first candidacy & administration as well as the storming of the Capitol. Additionally, the study of different political movements, like the left-wing Extinction Rebellion movement, could provide insightful evidence for the possible generalizable effects political movements might have on financial markets.

6 References

- Al-Maadid, A., Caporale, G. M., Spagnolo, F., & Spagnolo, N. (2020). The impact of business and political news on the GCC stock markets. *Research in International Business and Finance*, 52, 101102. <https://doi.org/10.1016/j.ribaf.2019.101102>
- Antweiler, W., & Frank, M. Z. (2004). Is All That Talk Just Noise? The Information Content of Internet Stock Message Boards. *The Journal of Finance*, 59(3), 1259–1294. <https://doi.org/10.1111/j.1540-6261.2004.00662.x>
- Baker, M., & Wurgler, J. (2007). Investor Sentiment in the Stock Market. *Journal of Economic Perspectives*, 21(2), 129–151. <https://doi.org/10.1257/jep.21.2.129>
- Bollen, J., Mao, H., & Zeng, X. (2011). Twitter mood predicts the stock market. *Journal of Computational Science*, 2(1), 1–8. <https://doi.org/10.1016/j.jocs.2010.12.007>
- Born, J. A., Myers, D. H., & Clark, W. J. (2017). Trump tweets and the efficient Market Hypothesis. *Algorithmic Finance*, 6(3–4), 103–109. <https://doi.org/10.3233/AF-170211>
- Burggraf, T., Fendel, R., & Huynh, T. L. D. (2020). Political news and stock prices: Evidence from Trump’s trade war. *Applied Economics Letters*, 27(18), 1485–1488. <https://doi.org/10.1080/13504851.2019.1690626>
- Campbell, J. Y., Lo, A. W., & MacKinlay, A. C. (1997). *The Econometrics of Financial Markets*. Princeton University Press.
- Greenwood, R., & Shleifer, A. (2014). Expectations of Returns and Expected Returns. *Review of Financial Studies*, 27(3), 714–746. <https://doi.org/10.1093/rfs/hht082>
- Karime, S., & Sayilir, Ö. (2019). Political news and stock market reactions: Evidence from Turkey over the period 2008–2017. *International Journal of Management and Economics*, 55(2), 83–98. <https://doi.org/10.2478/ijme-2019-0013>
- Leung, H., & Ton, T. (2015). The impact of internet stock message boards on cross-sectional returns of small-capitalization stocks. *Journal of Banking & Finance*, 55, 37–55. <https://doi.org/10.1016/j.jbankfin.2015.01.009>
- Lucey, C. (2025, April 5). Trump Promised a Manufacturing Boom, But Industries Aren’t Sure. *Bloomberg.Com*. <https://www.bloomberg.com/news/articles/2025-04-05/trump-promised-a-manufacturing-boom-industries-are-not-so-sure>
- Murphy, H., Gandel, S., & Temple-West, P. (2024, November 17). Maga Inc: How Donald Trump’s allies stand to gain from ‘anti-woke’ resurgence. *Financial Times*. <https://www.ft.com/content/daae53d0-3825-4d83-b8fd-a4478ccf3050>

- Novak, P. K., Smailović, J., Sluban, B., & Mozetič, I. (2015). Sentiment of Emojis. *PLOS ONE*, *10*(12), e0144296. <https://doi.org/10.1371/journal.pone.0144296>
- Sabherwal, S., Sarkar, S. K., & Zhang, Y. (2011). Do Internet Stock Message Boards Influence Trading? Evidence from Heavily Discussed Stocks with No Fundamental News. *Journal of Business Finance & Accounting*, *38*(9–10), 1209–1237. <https://doi.org/10.1111/j.1468-5957.2011.02258.x>
- Shiller, R. J. (2017). Narrative Economics. *American Economic Review*, *107*(4), 967–1004. <https://doi.org/10.1257/aer.107.4.967>
- Shiller, R. J. (2020). Narrative Economics: How Stories Go Viral and Drive Major Economic Events. Princeton University Press 1-408. <https://doi.org/10.2307/j.ctvc4j8j>
- Silva, J. M. C. S., & Tenreyro, S. (2006). The Log of Gravity. *Review of Economics & Statistics*, *88*(4), 641–658. <https://doi.org/10.1162/rest.88.4.641>
- Tumarkin, R. (2002). Internet Message Board Activity and Market Efficiency: A Case Study of the Internet Service Sector using RagingBull.Com. *Financial Markets, Institutions & Instruments*, *11*(4), 313–335. <https://doi.org/10.1111/1468-0416.11403>
- Wagner, A. F., Zeckhauser, R. J., & Ziegler, A. (2018). Company stock price reactions to the 2016 election shock: Trump, taxes, and trade. *Journal of Financial Economics*, *130*(2), 428–451. <https://doi.org/10.1016/j.jfineco.2018.06.013>
- Wooldridge, J. M. (2010). *Econometric Analysis of Cross Section and Panel Data*. MIT Press. <http://ebookcentral.proquest.com/lib/ubnru-ebooks/detail.action?docID=3339196>

7 Appendices

7.1 Appendix A: OLS Assumptions

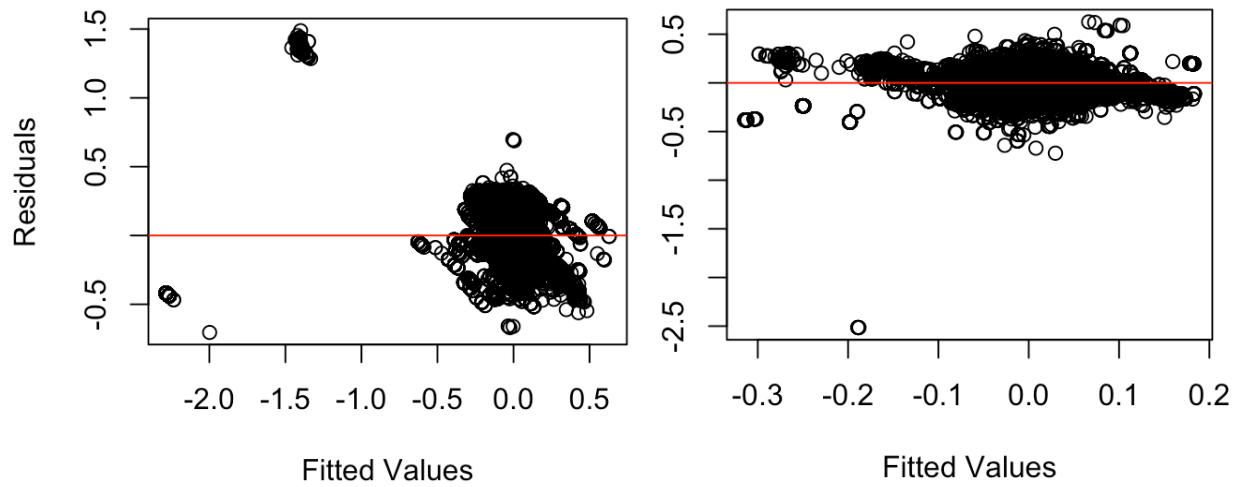
This appendix reports various figures and tables reporting the OLS assumption tests carried out in this thesis. The OLS assumptions tested in this thesis include: homoskedasticity, linearity, normally distributed residuals, no autocorrelation and no multicollinearity. Where figure A1 visualizes the assumptions of homoskedasticity and linearity, figure A2 reports the assumption of normally distributed residuals and table A1 reports the results of the formal statistical tests for the assumptions of homoskedasticity, no autocorrelation and no multicollinearity. Note that these assumptions are tested on a simplified model, excluding the various interaction terms since the formal statistical tests are not able to estimate such complex models.

Figure A1 indicates that for both the primary (left) and lead (right) model there is a presence of heteroskedasticity, since the plotted values vary in distance to the red 0-line and therefore a violation in the assumption of homoskedasticity is potentially present. The linearity assumption appears to be mostly met by the primary model since most values are randomly distributed around the 0-0 point, with some presence of (non-random) clustering. The lead model appears to be mostly linear, since the values seem to be randomly scattered around 0-0 while its values are somewhat heavier on the negative side.

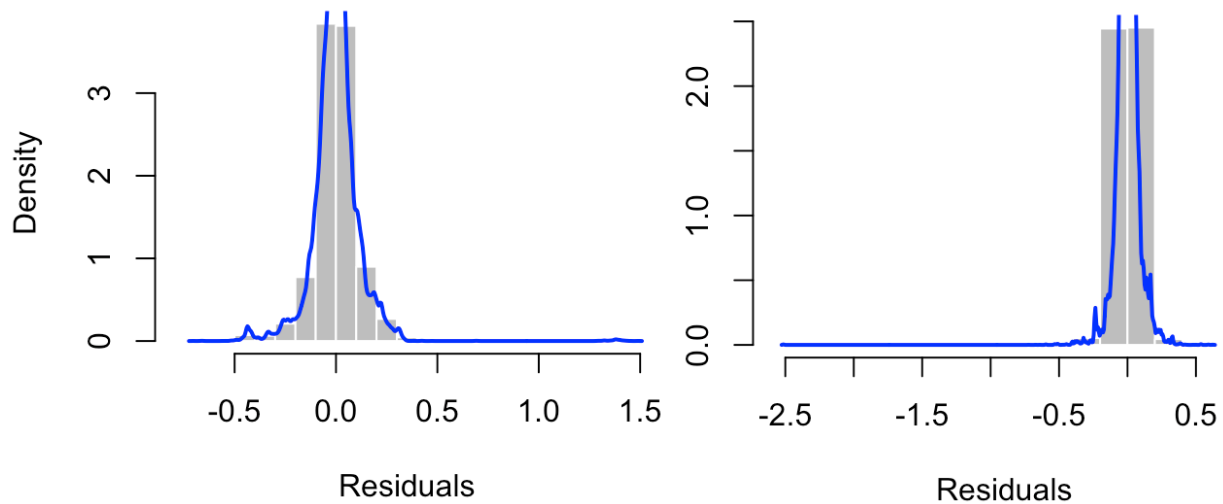
Figure A2 indicates that both models mostly meet the assumption of normally distributed residuals, with the distributions approaching normality by illustrating a rough bell-shape with fat-tails. For the primary model this tail is right sided, while being left sided for the lead model.

Finally, table A1 shows the results of the formal statistical tests of multiple OLS assumptions. The Breusch-Pagan and Durbin-Watson tests both indicate a violation in their respective assumptions, indicated with their p-values being highly significant. The VIF test for multicollinearity on the other hand does not report an indication of assumption violation, since the max VIF value for both models being 1.47 which is well below the thresholds of 5 or 10 and therefore not indicating the presence of multicollinearity.

Due to the violations in the homoskedasticity and no autocorrelation assumptions all regressions in this thesis implement heteroskedasticity-robust standard errors, clustered by firm to account for the violations.

FIGURE A1. RESIDUAL VS FITTED VALUE PLOT: PRIMARY & LEAD MODELS.

Notes: Residual vs fitted value plots provide a visual to inspect the OLS assumptions of homoskedasticity and linearity. Since the plotted values for both the primary (left) and lead (right) models appear to vary in distance to the red 0-line it is an indication for the presence of some heteroskedasticity in both models. A model appears to be linear when the values are randomly scattered around the 0-0 point. The primary model thus appears to be mostly linear, with some presence of (non-random) clustering. The lead model appears to be mostly scattered around the 0-0 point, indicating the lead model is mostly linear.

FIGURE A2. RESIDUAL DISTRIBUTION PLOT: PRIMARY & LEAD MODELS.

Notes: The residual distribution plots are aimed at the OLS assumption of normally distributed residuals. The residuals are deemed to approach normality when they are evenly distributed around 0. The primary model (left) appears to be mostly symmetrically distributed around 0 with a tail to the right. The lead model (left) shows heavy density around 0, but also a tail to the left. Indicating that both models might have residuals which appear to approach normality.

TABLE A1. FORMAL OLS ASSUMPTION TEST RESULTS.

Test	Primary Model		Lead Model	
	Test Value	p-Value	Test Value	p-Value
Breusch-Pagan (Homoscedasticity)	7701.98	< 0.001	331.47	< 0.001
Durbin-Watson (Autocorrelation)	0.07	< 0.001	0.38	< 0.001
Max VIF (Multicollinearity)	1.47		1.47	

Notes: The left most column indicates which statistical test is used and which OLS assumption it refers to; the test value indicates the strength and direction of the tests results. The OLS assumptions of homoskedasticity and no-autocorrelation are rejected at a threshold of a p-value < 0.05. The Breusch-Pagan test shows exceptionally high test values indicating the presence of heteroskedasticity in the models, which is formally confirmed by the very low p-values of < 0.001. The Durbin-Watson tests indicates the presence of positive autocorrelation due the test values being < 1. Again, being confirmed by the low p-values of < 0.001. The VIF tests reports a max value of 1.47 in both models for all the variables included. A VIF value of < 5 indicates no multicollinearity in the variable, indicating multicollinearity is not a concern in both models. To account for the violations in homoskedasticity and no-autocorrelation, heteroskedasticity-robust standard errors, clustered by firm are used all regressions.

7.2 Appendix B: Heterogeneity Analysis

TABLE B1. HETEROGENEITY ANALYSIS: TIME-SPIKE RESULTS.

Variable	OLS	PLM	PPML
Panel A: H1 Message Sentiment			
$Sentiment_{pos}$	-0.00685 (0.02047)	0.01336 (0.01725)	-0.00235 (0.00857)
$Sentiment_{neg}$	-0.03151*** (0.00816)	-0.09321*** (0.02560)	-0.01208*** (0.00293)
$Sentiment_{pos} \times Message\ Count$	-0.00003 (0.00003)	0.00098** (0.00049)	-0.00001 (0.00001)
$Sentiment_{neg} \times Message\ Count$	-0.00015*** (0.00005)	-0.00263*** (0.00076)	-0.00006*** (0.00002)
$Sentiment_{pos} \times Market\ Value$	-9.95e-07* (5.95e-07)	-2.78e-06*** (8.58e-07)	-3.55e-07 (2.22e-07)
$Sentiment_{neg} \times Market\ Value$	2.59e-06*** (9.10e-07)	1.33e-06 (1.32e-06)	9.38e-07*** (3.53e-07)
$Sentiment_{pos} \times Trade\ Volume$	3.26e-07*** (9.24e-08)	7.24e-07*** (2.26e-07)	1.19e-07*** (3.43e-08)
$Sentiment_{neg} \times Trade\ Volume$	-3.29e-07** (1.37e-07)	-4.27e-08 (3.93e-07)	-1.12e-07** (5.44e-08)
Panel B: H2 MAGA Association			
$MAGA_{pos}$	0.03252 (0.04090)	0.02472 (0.01965)	0.00655 (0.02291)
$MAGA_{neg}$	0.01977 (0.02852)	-0.02808 (0.03254)	0.00723 (0.01082)
$MAGA_{pos} \times Message\ Count$	0.00047 (0.00030)	0.00050 (0.00053)	0.00021 (0.00013)
$MAGA_{neg} \times Message\ Count$	0.00009 (0.00014)	-0.00122 (0.00078)	0.00005 (0.00006)
$MAGA_{pos} \times Market\ Value$	5.79e-06* (3.24e-06)	3.44e-06* (1.94e-06)	2.42e-06* (1.31e-06)
$MAGA_{neg} \times Market\ Value$	-1.11e-06 (2.26e-06)	2.94e-07 (2.30e-06)	-3.56e-07 (8.77e-07)
$MAGA_{pos} \times Trade\ Volume$	-1.59e-06** (7.50e-07)	-1.02e-06** (4.92e-07)	-6.59e-07** (3.05e-07)
$MAGA_{neg} \times Trade\ Volume$	-6.76e-07** (3.34e-07)	7.54e-08 (6.65e-07)	-2.74e-07** (1.34e-07)
Panel C: Time-Spike Interactions			
$Sentiment_{pos} \times Sentiment\ Spike_{pos}$	0.00125 (0.03472)	0.00890 (0.01181)	-0.00064 (0.01866)
$Sentiment_{pos} \times Sentiment\ Spike_{neg}$	-0.00729 (0.01141)	0.00770 (0.01067)	-0.00251 (0.00427)
$Sentiment_{pos} \times MAGA\ Spike_{Election}$	0.00539 (0.00856)	0.00059 (0.01083)	0.00223 (0.00316)
$Sentiment_{pos} \times MAGA\ Spike_{Tariff}$	-0.00199 (0.00823)	-0.00737 (0.01268)	-0.00055 (0.00305)
$Sentiment_{neg} \times Sentiment\ Spike_{pos}$	0.05645 (0.04301)	-0.03541 (0.02585)	0.02459 (0.02482)
$Sentiment_{neg} \times Sentiment\ Spike_{neg}$	-0.02854 (0.02167)	0.01334 (0.01581)	-0.01107 (0.00797)
$Sentiment_{neg} \times MAGA\ Spike_{Election}$	0.00766 (0.00834)	-0.00051 (0.01194)	0.00312 (0.00314)
$Sentiment_{neg} \times MAGA\ Spike_{Tariff}$	0.01378* (0.00715)	0.01813 (0.01444)	0.00567** (0.00275)
$MAGA_{pos} \times Sentiment\ Spike_{pos}$	0.18032 (0.28596)	0.04264 (0.06287)	0.10650 (0.15482)
$MAGA_{pos} \times Sentiment\ Spike_{neg}$	0.02544 (0.06201)	0.05536*** (0.01591)	0.01124 (0.02474)
$MAGA_{pos} \times MAGA\ Spike_{Election}$	-0.01532 (0.01472)	-0.00235 (0.01791)	-0.00509 (0.00551)
$MAGA_{pos} \times MAGA\ Spike_{Tariff}$	0.04193 (0.03261)	0.02146 (0.01640)	0.01484 (0.01240)
$MAGA_{neg} \times Sentiment\ Spike_{pos}$	0.02107 (0.10691)	-0.12065 (0.08857)	0.01042 (0.04957)
$MAGA_{neg} \times Sentiment\ Spike_{neg}$	0.01867 (0.02454)	-0.00356 (0.02599)	0.00660 (0.00923)
$MAGA_{neg} \times MAGA\ Spike_{Election}$	0.00231 (0.01934)	0.02340 (0.01917)	0.00108 (0.00713)
$MAGA_{neg} \times MAGA\ Spike_{Tariff}$	0.04197 (0.03127)	0.04453 (0.02892)	0.01580 (0.01150)
Panel D: Cross-Variable Interactions			
$Sentiment_{pos} \times MAGA_{pos}$	-0.00161 (0.00614)	-0.01149 (0.00750)	-0.00064 (0.00226)
$Sentiment_{pos} \times MAGA_{neg}$	-0.01023** (0.00476)	-0.00076 (0.00491)	-0.00381** (0.00183)
$Sentiment_{neg} \times MAGA_{pos}$	0.01445 (0.00900)	0.00288 (0.00752)	0.00551 (0.00343)
$Sentiment_{neg} \times MAGA_{neg}$	-0.00051 (0.00825)	-0.00771 (0.00559)	-0.00020 (0.00319)
Panel E: Control Variables			
<i>Message Count</i>	-0.00057 (0.00046)	0.00049 (0.00043)	-0.00025 (0.00019)
<i>Market Value</i>	0.00002*** (5.41e-06)	4.12e-06*** (1.21e-06)	7.99e-06*** (2.03e-06)
<i>Trade Volume</i>	1.82e-06*** (6.71e-07)	7.47e-08 (1.79e-07)	7.27e-07*** (2.78e-07)
<i>Dividend Yield</i>	0.00131 (0.00203)	-0.00078 (0.00057)	0.00055 (0.00084)
Firm Fixed Effects	Yes	Yes	Yes
Time Fixed Effects	Yes	Yes	Yes
Sector Fixed Effects	Yes	Yes	Yes
Observations	69708	19780	69708
Adjusted R^2	0.6621	0.0560	-0.0018

Notes: This table presents coefficient estimates and standard errors in brackets from time heterogeneity analysis fixed effects OLS (FEOLS), panel (PLM) and Poisson (PPML) regressions. The dependent variable is the daily log stock returns of firms listed on the Russell 2000 index. To account for the possible negative values of the dependent variable it has been shifted to be exclusively positive in value for the Poisson regression only. The independent variables and their interactions are grouped in panels, but are all included in the same full model. With panel A and B corresponding with the key variables of Hypothesis 1 and 2 respectively. Panel C presents the interactions of the spike dummies with the key variables, panel D presents the interactions between key variables and panel E presents the control variables. The spike dummies consist of a 5-day period around the date of the spike. All regressions control for firm, time and industry sector fixed effects, indicated by 'Yes' in the fixed effects section. Heteroskedasticity-robust standard errors, clustered at firm level are used by all regressions to account for violations in the OLS assumptions of homoskedasticity and no autocorrelation. ***, ** and * indicate statistical significance at the 1%, 5% and 10% levels respectively.

7.3 Appendix C: Documentation of Generative AI Usage

1. Statement of purpose & Used Tools Specification

This appendix reports a detailed overview of the use of generative AI in developing and writing this thesis. The generative AI tools were used to support in the coding, data analysis, visualization of results in for example tables or graphs and in the creation of a MAGA dictionary used to assign MAGA associations to the StockTwits messages. All outputs provided by generative AI were critically evaluated and modified by myself where necessary, to be aligned with the objectives of this thesis.

OpenAI's ChatGPT is the only generative AI tool used in the development and writing of this thesis, more specifically its 4o model.

2. Scope of Use

This section provides a clear breakdown of how and where generative AI was used in this thesis, specifying in which sections of this thesis the generative AI played a role.

2.1 Coding & Data Analysis

- Generative AI Tool: ChatGPT (Model 4o)
- Purpose: Generating the initial code of compiling the dataset by extracting the StockTwits messages for each of the Russell 2000 firms considered in the thesis. Additionally, to assist in compiling a MAGA dictionary used to assign MAGA associations to the StockTwits messages. Finally, to generate the initial code to analyze the dataset on message sentiments and MAGA associations, as well as to create tables to report the results of the analyses.
- My role: The AI outputs were reviewed, modified and tested for correctness and alignment with the hypotheses of this thesis.
- My reflection: The generative AI helped me to improve my coding in R, allowing me to potentially perform similar analyses and data extractions without the help of generative AI in the future. Additionally, the support of Generative AI allowed to me perform a more complex analysis with limited time and resources, mostly due to the complexity of initially extracting over 100,000 StockTwits messages and creating a model to assign sentiments and MAGA associations automatically to these messages without me needing to manually assign them.
- Applied in sections 3 & 4 and for all tables reported in the thesis and its appendices.

2.2 Data Visualization

- Generative AI Tool: ChatGPT (Model 4o)
-

- Purpose: Generating the initial R code to visualize the data from the dataset and its regression results in graphs and tables. Due to multiple errors arising in trying to report the regression results in clear and professional tables, generative AI assisted to provide fixes and solutions to these errors.
- My role: Where the generative AI provided the initial base R code for graphs and tables, I reviewed these codes and edited them to ensure accurate reporting and to include clear labels and notes. This to ensure the graphs and tables are aligned with the visual guidelines for the thesis.
- My reflection: Generative AI assisted me in saving already limited time, by providing a base code on which I could build and improve. This allowed me to spend additional time on the interpretation of results and actually writing the thesis over spending a lot of time on creating professional tables and graphs.
- Applied in sections 3 & 4 and appendices A & B, more specifically on the tables and graphs reported in this thesis.

2.3 Writing & Grammar Refinement

- Generative AI Tool: ChatGPT (Model 4o)
- Purpose: To improve the clarity, readability and use of professional language in the thesis, by providing suggestions for restructuring sentences.
- My role: Suggestions provided by generative AI were critically reviewed and edited before being included in the thesis. This to ensure alignment with the overall writing style of the thesis and thus not harm the fluency of the text. Additionally, no sections or sentences were directly copied from the generative AI's suggestions without being reviewed, edited and adjusted.
- My reflection: generative AI assisted in providing a clear sentence structure for more difficult or complex sentences of my thesis. As well as to be more concise and therefore increasing the readability and reducing the length of the thesis.
- Applied to sections 1, 2, 3 and 4.

3. Log of Generative AI Interaction

This section reports an overview of the interaction I had with generative AI for the three main purposes detailed above. The overview includes the input queries and the generative AI's output to these queries, as well as an explanation on how each output was used or edited. Note that since writing a suitable R code without errors which is aligned with the hypotheses, for example, is an iterative process not every single query is detailed below. Rather an indicative query for each of the three generative AI purposes is provided, to ensure this appendix and the thesis overall remains compact but clear and insightful. For the same reason extensive R code suggested by generative AI is indicated by (...) below.

3.1 Coding & Data Analysis

- **Query to ChatGPT:** “Write an R code to extract StockTwits messages for Russell 2000 firms”
- **Output from ChatGPT:** “To extract StockTwits messages for **Russell 2000 firms** in R, you’ll need to:

1. Load a list of **Russell 2000 tickers** (e.g., from a CSV file or vector).
2. Loop over each ticker and extract messages using the StockTwits API.
3. Optionally paginate to retrieve more messages per stock.
4. Combine and save all data.”

Additionally each step was then further detailed as follows: “Step 1: Prepare Ticker List. If you have a CSV file of Russell 2000 tickers load it like this: (...). Or define a small test vector manually: (...).

Step 2: Define Message Extraction Function. This function pulls messages: (...).

Step 3: Loop Over All Russell 2000 Tickers: (...).

Step 4: Save All Messages: $(final_df \leftarrow bind_rows(all_stocktwits_data)) write.csv(final_df, "stocktwits_russell_2000.csv", row.names = FALSE)$

Notes: - You can safely extract messages for approximately 150 tickers per hour with default rate limits.

- For large-scale date: run the script in batches or on a schedule.

- Consider storing intermediate results in RDS or CSV to avoid data loss.”

- **How it was used:** The generated code was used as a starting point of the iterative process of writing an R code without errors which does not exceed the extraction limits imposed by StockTwits’ API.

3.2 Data Visualization

- **Query to ChatGPT:** “Write an R code for a table in which I can report the regression results of my models. Using Latex in R.”
- **Output from ChatGPT:** “Here is a complete **R script** for estimating **three regression models** and exporting the results to a **LaTeX table**, along with a **step-by-step explanation** of how it works.
 - Step 1: Load Required Packages. 1. Fixest is used to run regression models (like OLS, fixed effects, etc.). 2. Modelsummary() helps format the model output into publication-style tables (especially for Latex). (...)
 - Step 2: Estimate Three Models. We use the feols function from fixest. (...)
 - Step 3: Coefficient Labels. 1. Coef_map helps you customize how variable names appear in the

table. 2. For example, `hp:wt` will be shown as *Horsepower x Weight*. (...)

- Step 4: Create Latex Table. 1. `Modelsummary()` collects the model results into a single LaTeX table. 2. `Statistics = "{std.error}"` means standard errors will appear in parentheses below estimates. 3. `Stars = TRUE` adds *, **, *** for significance at 0.1, 0.05, 0.01 levels. 4. `Gof_omit` removes clutter from the goodness-of-fit section (AIC, BIC, etc.). 5. `Output = "latex"` sends the LaTeX code to the R console. (...)

Step 5 (Optional): Save the Table to a File. 1. The `writeLines()` command saves the LaTeX table to a file called `regression_table.tex`. 2. You can then `/input{regression_table.tex}` in your LaTeX document. (...)"

- **How it was used:** The suggested code was reviewed and edited multiple times to fix occurring errors, again indicating an iterative process with the suggested code acting as a starting point.

3.3 Writing & Grammar Refinement

- **Query to ChatGPT:** *"Is the grammar used in the following sentence correct? This increasing number of new investors, potentially lacking knowledge and experience in financial markets, could be one of the reasons for the increasing popularity of and activity on internet stock market message boards."*
- **Output from ChatGPT:** *"Your sentence is mostly clear and grammatically sound, but it can be slightly improved for readability and flow. Here's a polished version: **"The increasing number of new investors, potentially lacking knowledge and experience in financial markets, could be one of the reasons for the growing popularity and activity on internet stock market message boards."***
Changes made: - Replaced "This increasing number" with "The increasing number" – "This" implies a reference to something specific just mentioned, which may not be the case here. - Changed "increasing popularity of and activity on" to "growing popularity and activity on" – this avoids the awkward split phrase and is more concise."
- **How it was used:** After having reviewed the suggested new sentence from ChatGPT and made sure the suggestion was correct and did not impact or change the intended message of the sentence, the new sentence was directly incorporated in the thesis.