# FOMC induced Sentiment Shock and the Stock Market Returns: the Social Media Channel

*By* Zhenning Zhao

Draft: December 22, 2022

*The rapid growth of social media networks over the past decade has facilitated the spread of sentiment shocks online, amplifying their impact on the stock market. This paper provides direct causal evidence that sentiment shocks induced from Federal Open Market Committee (FOMC) announcements can affect stock market returns through social media channels. By using historical Twitter data surrounding FOMC announcements, I combine machine learning and lexical methods to measure the sentiment changes induced by these announcements. By comparing the differences between stocks with high and low exposure to Twitter (measured by the number of tweets posted about a specific stock in the past week), I demonstrate that stocks with higher Twitter exposure are more affected by sentiment shocks, providing evidence for the existence of a social media transmission channel.*

There has been a long-standing interest in understanding the extent to which public sentiment can directly affect the economy. From a theoretical perspective, it is believed that changes in public sentiment, particularly those that are not tied to economic fundamentals, can have a significant impact on the economy through their effects on consumer demand and business investment. Specifically, when people have a positive perception of the economy, they may be more likely to make purchases and engage in economic activity, which can increase demand and drive economic growth. Similarly, businesses may be more inclined to invest in new projects and expand their operations when they have a positive outlook on the economy, further contributing to economic growth. Therefore, it is generally accepted that sentiment shocks, or sudden changes in public sentiment, can have a direct impact on the economy.

However, in recent years, the proliferation of social media networks has led to a rapid increase in the flow of sentiment online. This has given rise to the concept of the indirect sentiment impact, which suggests that changes in sentiment caused by events can influence others and have consequences in the financial market. This is especially relevant in today's interconnected world, where news and events can spread quickly through social media and other online channels. As entrepreneur Jay Baer once said, "Content is fire and social media is gasoline." This analogy highlights the potential for social media to amplify the impact of sentiment on the financial market.

Despite the appeal of the indirect effect hypothesis, empirical identification of this phenomenon is challenging for several reasons. First, it is difficult to accurately measure changes in sentiment that go beyond fundamental changes, particularly for the general public and at high frequencies. The most widely used survey databases for tracking economic expectations in the US, such as the Michigan University Consumer Survey, Survey of Professional Forecasters (SPF) dataset, and Gallup Survey, have low frequencies or only provide data at a daily frequency. Second, measurement error is a significant issue when attempting to gauge public sentiment, as it involves trying to understand the behaviors and expectations of human beings. Finally, even if these issues can be overcome, it is still difficult to establish strong causal evidence of the transmission of sentiment through social media channels due to the presence of reverse causality and simultaneity issues.

To address the challenges mentioned above, I have chosen to use Twitter data to analyze public sentiment about the US economy. There are several advantages to this approach. Twitter is a real-time platform with a much higher frequency than other sources of data. Twitter has millions of users, so we can analyze a large number of tweets to get a sense of overall sentiment towards the economy. Many tweets are publicly available, making it easy for market participants to access and pay close attention to them. The popularity of Twitter also makes it more likely to be a medium for the flow of sentiment. For the analysis, I have used approximately 2 million tweets related to the Federal Open Market Committee (FOMC) meetings from 2009-2019, which have been cleaned to include only original tweets.

After obtain the social media text data from twitter, I first leverage a ULMFiT machine learning model to classify tweets based on their geographical location in order to identify those related to the US economy. ULMFiT is one of the most advanced state of the art machine learning model specifically good at text mining. The model performs well, achieving an accuracy score of 82% in predicting whether a tweet was posted within the US. After training the model, it is applied to the entire dataset to label all tweets with their predicted geographical location. The results show that more than half of the raw text data is not related to the US economy, and the number of tweets related to the US economy varies over time.

Next, I applied a lexical method and the SentiWordNet resource to assign a sentiment score to each tweet in the dataset. The SentiWordNet database assigns sentiment scores to words based on their usage and context in the text. The sentiment score for a tweet is calculated by summing the positive scores of the words in the tweet minus the negative scores. The distribution of sentiment scores is approximately normal, with a high concentration at zero. Positive tweets are defined as those with a sentiment score above 0.05, and negative tweets are defined as those with a score below -0.05. Then the Twitter sentiment index is calculated as the fraction of positive tweets minus the fraction of negative tweets for a given time period, providing a measure of public sentiment about the US economy. The Twitter sentiment index is then compared to the Michigan survey consumer sentiment index, showing a strong positive correlation with a correlation of 0.74. This initial analysis using monthly sampling frequency confirms that the Twitter-based measurement accurately captures public sentiment about the US economy.

To measure the impact of Twitter sentiment on stock market returns, I use a high-frequency identification method to detect changes in sentiment induced by FOMC announcements and a linear regression to isolate the variations in public sentiment caused solely by the tone of the FOMC announcement and remove the effect of yield curve shocks. I first confirms that FOMC announcements affect Twitter user behavior by comparing the number of tweets on announcement dates with the average number of tweets on non-announcement dates. The changes in the Twitter sentiment index is then calculated for various time periods around FOMC meetings and compared to similar changes for normal days to identify any changes in sentiment. The content of the FOMC announcements is also analyzed using the same sentiment analysis procedure as the tweets to assign a sentiment score. A regression is then run to determine the effect of FOMC announcements and monetary policy shocks on changes in public sentiment, controlling for yield curve shocks. The results show that comparing to the normal days, a relative neutral FOMC announcement has a statistically significant positive effect on changes in public sentiment, with relatively positive announcements leading to a further increase in sentiment. As a preliminary analysis I also examine the relationship between intraday returns of the S&P 500 index and the identified sentiment shock while controlling for yield curve shocks. The focus is on the coefficient that shows the impact of the sentiment shock induced by the FOMC on intraday stock market returns. Results show that a positive sentiment shock related to the US economy results in an immediate increase in stock market returns that last for about 90 minutes.

This study aims to examine the casual relationship between sentiment on Twitter and stock market returns, and to determine whether the social media channel plays a role in this relationship. However, the result described above only speaks to the correlation but cannot definitively determine the causality, as there are multiple potential explanations for the observed correlation. To solve the identification problem, I propose using stock level data to compare the response of stocks with high exposure to Twitter to the response of stocks with low exposure to Twitter. By doing so, we hope to identify the causal effect of Twitter sentiment on stock market returns and determine the role of the social media channel in this relationship.

To address the problem of identifying the direction of causality, I create a counterfactual scenario in which only the causal link between Twitter sentiment and stock market returns is broken, while all other relationships remain unchanged. I do this by using stock level data and comparing the response of stocks with high exposure to Twitter to the response of stocks with low exposure to Twitter. To compare the real-world scenario, where the relationship between Twitter sentiment and stock market returns is intact, to the counterfactual scenario, where this relationship is broken, I use stock-level data and compare the response of stocks with high and low exposure to Twitter. Specifically, I will measure the exposure of a particular stock to Twitter by counting the number of tweets about the stock that have been posted in the past 7 days. Stocks with high exposure to Twitter will represent the real-world scenario, while stocks with low exposure will represent the counterfactual scenario. The results of the analysis show that Twitter sentiment has a statistically significant and economically meaningful impact on stock market returns. This impact is stronger for stocks with higher exposure to Twitter, suggesting that the social media channel does exist and that market participants do pay attention to Twitter sentiment when making investment decisions. This is the main contribution of this paper.

Lastly, I conduct several placebo tests to validate the main finding. One test involved using stock return data from the week before the FOMC announcement week and keeping the shocks the same as in the baseline regression. The results of this test showed that there was no significant difference in stock returns between stocks with high and low social media exposure for the counterfactual data, indicating that the baseline regression was not simply picking up a weekly pattern. Another test involved using data from the last Wednesday to Friday of the FOMC announcement week and creating several counterfactual shocks by moving the true shocks that would have occurred during the FOMC announcement dates to different dates. The results of this test showed that the parallel trend assumption held true, as none of the coefficients of the placebo regressions were significantly positive. The results of these tests supported the conclusion that sentiment shocks are amplified through the social media channel for stocks with high exposure to Twitter. The tests also indicate the presence of dynamic effects, with a positive difference in daily returns between high and low exposure stocks observed for two days after the event.

In conclusion, the analysis in this paper found that stocks with high exposure to Twitter tend to have a stronger response to sentiment shocks. This amplification effect of the social media suggests a new mechanism by which the Federal Reserve (FED) can influence the economy through its language choices in announcements. This finding highlights the importance of considering the role of social media in the transmission of economic news and the potential impact on financial markets. This emphasizes the need for the FED to carefully consider the words it uses in announcements, as they may have a greater impact on financial markets than previously thought.

The following sections are organized as follows: The next section provides a literature review. In section 2, the details of how to measure the sentiment of Twitter texts are explained. Section 3 discusses the specifics of FOMC-induced sentiment changes. The main focus of the paper, the identification of the social media channel, is discussed in section 4. The paper concludes in section 5.

## I. Literature Review

Before discussing the details of this paper, it is important to understand how it fits into the existing literature. There have been numerous studies in the fields of computer science and finance that have attempted to use Twitter text data to predict stock market returns, such as Rao and Srivastava (2012), Makrehchi, Shah and Liao (2013), Si et al. (2013), Ranco et al. (2015), and Oliveira, Cortez and Areal (2017). However, many of these early studies focused on the information of a specific stock rather than the aggregate economy. One reason for this lack of focus on the aggregate economy is that Twitter only made its API available to academic researchers for historical data starting in January 2021. Before this time, researchers could only access data from the most recent year. More recently, with the availability of a greater amount of data, there have been a number of studies that have attempted to use Twitter data to predict the aggregate economy and its relationship to inflation expectations and monetary policy, such as Gu and Kurov (2020), Lüdering and Tillmann (2020), Angelico et al. (2022), Hamraoui and Boubaker (2022), and Masciandaro, Romelli and Rubera (2022). One notable study in this field is Baker et al. (2021), which examined economic uncertainty but not sentiment. This paper aims to fill this gap in the literature and specifically focuses on identifying the causality of the social media transmission channel, which is its main contribution.

This paper focuses on the impact of sentiment shocks on the economy, building on previous research that has used dynamic stochastic general equilibrium (DSGE) models to examine sentiment shocks that are unrelated to economic fundamentals, for example Barsky and Sims (2012), Milani (2017) and Angeletos, Collard and Dellas (2018). In contrast to these studies, this paper specifically investigates sentiment shocks induced by the FOMC. This paper also differs from previous empirical studies on sentiment shocks like Levchenko and Pandalai-Nayar (2020) by using a new dataset containing Twitter text to examine this topic from a different angle.

The construction of sentiment shocks induced by the Federal Open Market Committee (FOMC) is based on the idea that there are multiple channels through which the FED can influence the economy, beyond the yield curve. Previous research has used high-frequency methods to identify yield curve shocks caused by traditional monetary policy, forward guidance, and large-scale asset purchases in Gürkaynak, Sack and Swanson (2005) and Swanson (2021). However, it has been recognized that the yield curve is not the only way that the FED can impact the economy. Boehm and Kroner (2021) used the method from Gürkaynak, Kisacikoğlu and Wright (2020) to create non-yield curve shocks from FOMC meetings, and Handlan (2020) applied machine learning techniques to study the text shocks of FOMC announcements. This paper is related to these studies in that it investigates the FOMC-induced sentiment shock as a component of non-yield curve shocks. However, it differs in its identification of a novel social media channel for analyzing this phenomenon.

## II. Measuring the Sentiment of Tweets

### *A. Twitter Text Data*

We begin with raw text data from Twitter, a popular micro-blogging platform that allows users to easily share their opinions and feelings. As such, it is a useful tool for the FED to transmit messages to the public. The large amount of text data available on Twitter provides a new way to construct public sentiment about the US economy. One major advantage of using Twitter data is its high frequency, which allows us to examine sentiment changes at a more granular, intraday level than other survey data such as the quarterly SPF survey and monthly Michigan survey.

However, Twitter data also has some limitations. Baker et al. (2021) found that Twitter users tend to be younger and that many tweets are generated by bots for advertising or campaigning purposes, both of which can decrease the credibility of the data. To address these issues, I used the same methods as Baker et al. (2021) to download tweets containing key words related to the economy.[1] Starting in 2006, Twitter was a small company with relatively few users, so the data from this period is of low quality. Therefore, the sampling period for this paper is from January 2009 to June 2019. To further improve the quality of the data, I only included original tweets without hyperlinks, retweets, replies, or duplicated text. For model training, I focused on the week of FOMC meetings, which includes two days before and after the meetings. The total number of tweets used in the main analysis is approximately 2 million, and the data includes the raw text of the tweets, the time they were posted, the number of comments, likes, and retweets, and the number of followers of the accounts. The details of the text cleaning procedure can be found in the appendix.

[1]To search for tweets related to the economy, I used the following key words: "economic", "economical", "economically", "economies", "economist", "economists", and "economy". An interesting note is that the word "economics" had to be excluded from the list because most of the tweets containing it were complaints from college students studying economics.

### *B. Prediction of the Tweet Location*

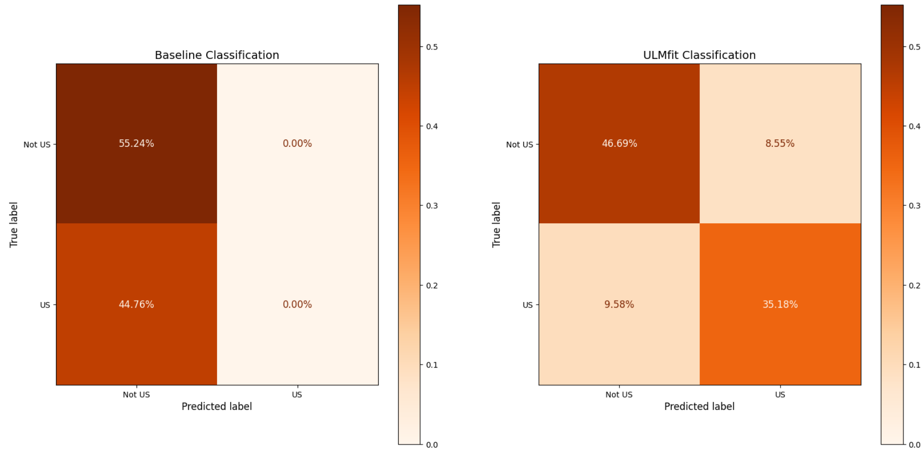
Before we begin to construct a sentiment measurement using Twitter data, we need to consider the fact that not all tweets in the dataset pertain to the US economy rather than the economies of other countries. To determine which tweets in our dataset are related to the US economy, we need to identify those that were posted within the US. Following the methods of Baker et al. (2021) and Handlan (2020), we can make the assumption that only tweets posted within the US is related to the US economy, and use a ULMFiT machine learning model to classify tweets based on their geographical origin. ULMFiT (short for "Universal Language Model fine tuning") is a state-of-the-art machine learning model developed by the team at Fastai that excels at classification tasks. The neural network model takes cleaned text as input and predicts the geographical location of the tweets. Further details about the ULMFiT model[2] can be found in the appendix.

The training dataset and geographical labels for the tweets are provided by Twitter's location services, which are turned off by default and must be manually activated by users. As a result, only a small portion of the tweets in our dataset include location information. Out of the approximately 142,000 tweets with location tags, I split the dataset into a training and test set using a 90-10% ratio. I trained the model on the training set and evaluated its performance on the test set.

The trained ULMFiT model performed well, achieving an out-of-sample accuracy score of approximately 82%, as shown in the confusion matrix in Figure 1. This means that it can accurately predict whether a tweet was posted within the US territory. In comparison, a dummy classifier that simply assigns all tweets to the most common group (i.e., those not posted within the US) has an accuracy of 55%, indicating that the ULMFiT model significantly improves predictive power by 27%.

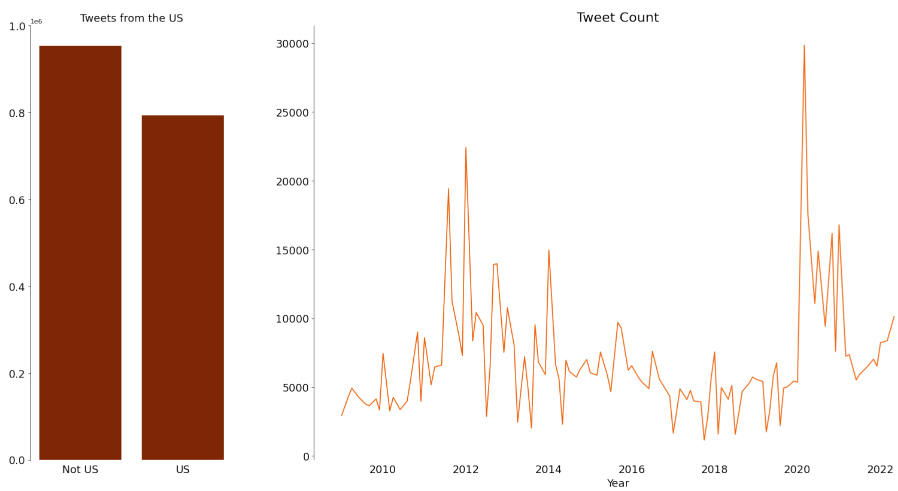
After training the model, I applied it to the entire dataset to obtain predicted geographical labels for all the tweets. Figure 2 provides a brief summary of the number of tweets that were posted within the US. The left panel shows that more than half of the raw text data is not related to the US economy. The right panel also reveals that the number of tweets varies over time, with stable levels before late 2011 and after mid-2014, and a spike in tweets related to COVID-19.

[2]The detail of the ULMfit model can be found in Howard and Ruder (2018).




*Note:* This figure contrasts the out-of-sample prediction accuracy of two models using confusion matrices. The left matrix is for the baseline dummy classifier, which simply assigns all tweets to the most common group, while the right matrix is for the trained ULMFiT model.

Figure 1. Confusion Matrix Comparison




*Note:* This figure illustrates the number of tweets predicted by the trained ULMFiT model. The left panel shows the number of tweets predicted as originating from the US, while the right panel displays the number of US tweets over time, with a focus on the tweets surrounding FOMC meetings.

Figure 2. Count of Tweets

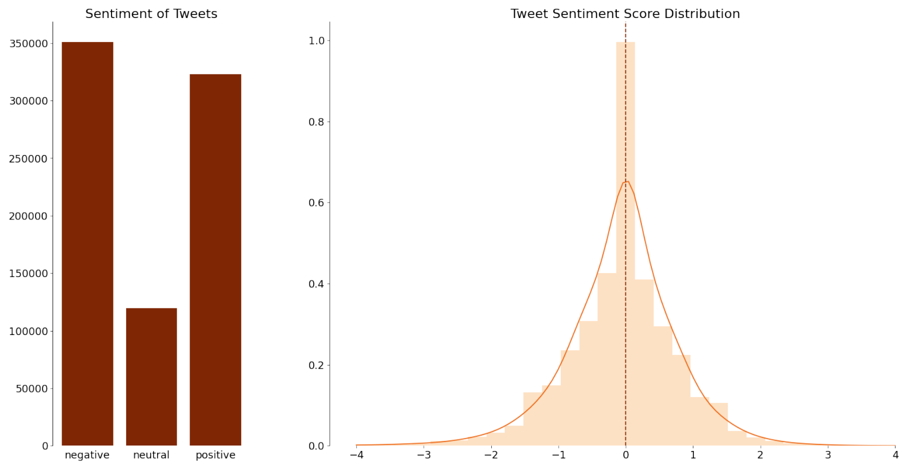
### C. Measuring Twitter Sentiment

Next, we will measure the sentiment of the Twitter text. Since we don't have a pre-labeled dataset for Twitter sentiment as in Shapiro, Sudhof and Wilson (2022), we cannot train a machine learning model to obtain the sentiment of the tweet. Therefore, we will use a more traditional method known as the lexical method. This method assigns a sentiment score to each tweet using SentiWordNet, a lexical resource for natural language processing. SentiWordNet assigns a sentiment score to words based on their usage and context in text. It is a database of words and their associations with multiple senses, synonyms, and antonyms, as well as their associated polarity and part-of-speech tags. Compared to other lexical resources commonly used in text sentiment analysis, SentiWordNet has a high coverage of over 150,000 words, making it suitable for opinion mining. For each synset of WordNet, three sentiment scores are assigned: positivity, negativity, and objectivity. After removing the stopping words and obtaining the part-of-speech tags, we can calculate the latent sentiment score for a tweet using the following formula:

$$\text{Sentiment Score} = \sum_{w \in \text{word of tweet}} \text{Positive Score}_w - \text{Negative Score}_w \tag{1}$$

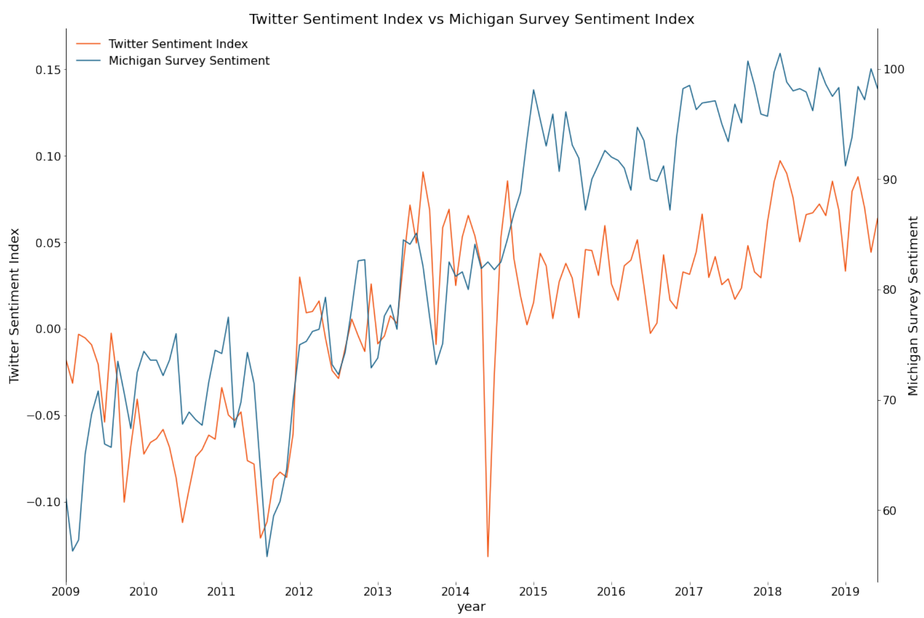
Figure 3 shows the distribution of latent sentiment scores for each tweet. The distribution is approximately normal with a high concentration at zero, indicating that many tweets either have a neutral tone or use words that do not appear in the corpus. To filter out noise and sensitivity to specific words and text length, we define positive tweets as those with a latent score above 0.05 and negative tweets as those with a score below -0.05. The Twitter sentiment index is then calculated as the fraction of positive tweets minus the fraction of negative tweets for a given time period, providing an aggregate measurement of public sentiment on the US economy.

Before we delve into the details of high-frequency identification of FOMC-induced sentiment changes, we must first confirm that our Twitter data can accurately capture public sentiment about the US economy. As a sanity check, we can calculate the monthly Twitter sentiment index and compare it to the Michigan survey consumer sentiment index. The time series data is plotted in Figure 4. The figure shows a strong positive correlation between the two indices,




*Note:* This figure presents the distribution of latent sentiment scores among the tweets. The dashed line indicates the median of the data, while the orange bar displays a histogram of the data. The orange line represents the distribution of the kernel density function obtained using a gaussian kernel with a band width of 0.2.

FIGURE 3. DISTRIBUTION OF THE LATENT TWEET SENTIMENT




*Note:* The figure displays the Michigan survey sentiment index and the Twitter sentiment index on the same plot. To directly compare with the monthly survey data, the Twitter sentiment index is computed on a monthly basis.

FIGURE 4. SENTIMENT INDEX TIME SERIES COMPARISON

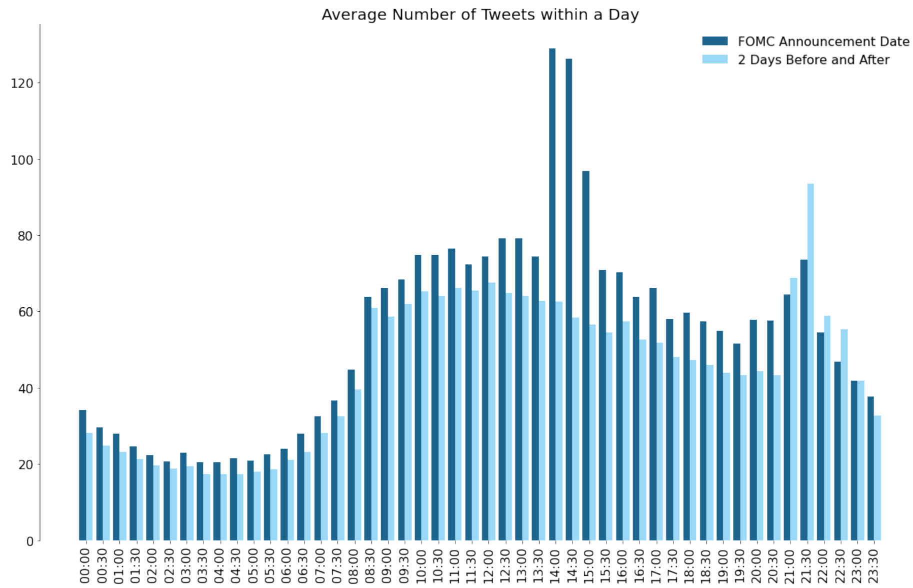
with a correlation of 0.74 and a p-value close to zero. The results of our preliminary analysis, which uses monthly sampling frequency, are useful for confirming that our Twitter-based measurement accurately captures public sentiment about the US economy. Although our main analysis in the next section uses a higher sampling frequency, this initial result provides valuable reassurance.

## III. FOMC Induced Sentiment Changes

With the measurement of the Twitter sentiment index for any given time period, we can use the high frequency identification method to detect changes in sentiment induced by the FOMC announcements. We can then use a linear regression to isolate the variations in public sentiment caused solely by the tone of the FOMC announcement and remove the effect of yield curve shocks.

The FOMC announcements are released at 2:00 PM Eastern Time after January 2013 and at 2:15 PM Eastern Time before that. To avoid endogeneity resulting from irregular FOMC meetings, we follow the approach of Boehm and Kroner (2021) and only include regular FOMC meetings in our analysis. By analyzing changes in sentiment on Twitter, we can gain insight into market reactions and investor sentiment following the release of the FOMC announcements. We first confirm that the FOMC announcement will affect Twitter user behavior by comparing the number of tweets on announcement dates with the average number of tweets on non-announcement dates. To do this, we plot the number of tweets in each 30-minute window for both groups in Figure 5. To ensure that other economic news does not confound our results, we use the average number of tweets from two days before and after the FOMC announcements as a reference for non-announcement dates. As we can see from the data, the number of tweets is lowest at night, begins to increase in the morning, and then decreases in the afternoon. There is a small peak at night when people are off work and may use Twitter for entertainment. The key observation is that there are significantly more tweets on FOMC announcement dates than on non-announcement dates immediately after the announcement is made. The increase in tweets lasts for about 90 minutes before returning to normal levels. By comparing the number of tweets on announcement dates with the average from non-announcement dates, we can more accurately identify any changes in Twitter user behavior that may be attributed to the FOMC announcement.

To construct the sentiment changes around FOMC meetings, we can calculate




*Note:* The figure compares the number of the tweets posted within each 90 minute window in 24 hours for the FOMC announcement dates and the average of 2 days before and after the FOMC announcements.

Figure 5. Daily Count of Tweets

the Twitter sentiment index between the time of the meeting announcement and x hours after the announcement, and compare it to the sentiment index from 9:30 AM (market opening time) to the time of the announcement. We can vary the time window x from 30 minutes to 90 minutes. We can also calculate the same sentiment changes for normal days to provide a comparison.

The sentiment changes around FOMC meetings cannot arise out of nowhere and must be driven by some underlying factor. To understand what is driving these changes, I first downloaded the FOMC announcement text from the FED's website. I then applied the same sentiment analysis procedure that I used for the tweets to the announcement text in order to assign a sentiment score. I defined a FOMC announcement as relatively positive if its latent sentiment score was higher than the median of all announcements in the sample period, and as relatively neutral or negative if its score was lower than the median. This allowed me to compare the sentiment of the announcement text to the changes in sentiment observed in the tweets. By doing this, I can determine whether the changes in sentiment are related to the content of the FOMC announcement or if they are due to some other factor that have been well studied by the literature,

Table 1—regression of Twitter Sentiment Changes on the FOMC Announcements

| | Tweets before 14:30 | Tweets before 15:00 | Tweets before 15:30 | Tweets before 16:00 |
|---|---|---|---|---|
| Is FOMC Date | 0.199*** | 0.189*** | 0.151*** | 0.128*** |
| | (0.028) | (0.023) | (0.022) | (0.021) |
| Is Positive | 0.048* | 0.040* | 0.040* | 0.041** |
| | (0.029) | (0.024) | (0.022) | (0.021) |
| FFR Shock | 0.027 | 0.070 | 0.074 | 0.065 |
| | (0.104) | (0.078) | (0.085) | (0.088) |
| FGF Shock | -0.044** | -0.043*** | -0.040*** | -0.043*** |
| | (0.018) | (0.014) | (0.014) | (0.014) |
| LSAP Shock | 0.009 | 0.004 | -0.008 | -0.012 |
| | (0.014) | (0.013) | (0.014) | (0.015) |
| Obs | 420 | 420 | 420 | 420 |
| R Squared | 0.283 | 0.388 | 0.345 | 0.319 |

*Note:* This table displays the regression result of the basic regression in equation 2. The robust standard errors are shown in parentheses. The 4 columns denotes 4 separate regressions where the definition of the Twitter sentiment changes are defined using different time intervals. The coefficients of the yearly linear trend and the quarterly effect are omitted. ***: p<0.01, **: p<0.05, *: p<0.1.

such as a yield curve shock.

We run the following regression to determine the effect of FOMC announcements and monetary policy shocks on the public sentiment changes:

$$\Delta TSI_t = \gamma_0 + \gamma_1 1(\text{FOMC Date})_t + \gamma_2 1(\text{is Positive})_t + \gamma_3 X_t + v_t \tag{2}$$

where $\Delta TSI_t$ is the change in Twitter sentiment on date $t$, $1(\text{FOMC Date})_t$ is a dummy variable that is equal to 1 on FOMC announcement dates and 0 on all other dates, and $1(\text{is Positive})_t$ is a dummy variable that is equal to 1 if the FOMC announcement on date $t$ is relatively positive and 0 if it is relatively neutral or negative. This variable is also equal to 0 on normal days. $X_t$ mostly represents the control variables for the yield curve channel, which includes traditional monetary policy shock, forward guidance shock, and large-scale asset purchasing shock. These shocks are taken from Swanson (2021) and are equal to zero when the FED is not making any announcements. The control variables also include a quarterly effect and a linear yearly trend. The regression result is shown in table 1.

Several key observations can be made from the results in Table 1. Firstly, the forward guidance shock is negatively correlated with changes in aggregate sentiment, which aligns with the expectation that a contractionary monetary policy shock would dampen people's perception of the overall economy. This finding is also consistent with previous research such as Lewis, Makridis and

Mertens (2019). Second, the effects of traditional monetary policy and large-scale asset purchases on sentiment changes are not statistically significant. One potential reason for this is that the sample period under examination coincides with a time when the FED was at the zero lower bound and primarily relied on other tools to stimulate the economy. Additionally, the infrequent use of large-scale asset purchases may have contributed to the insignificance of the coefficient.

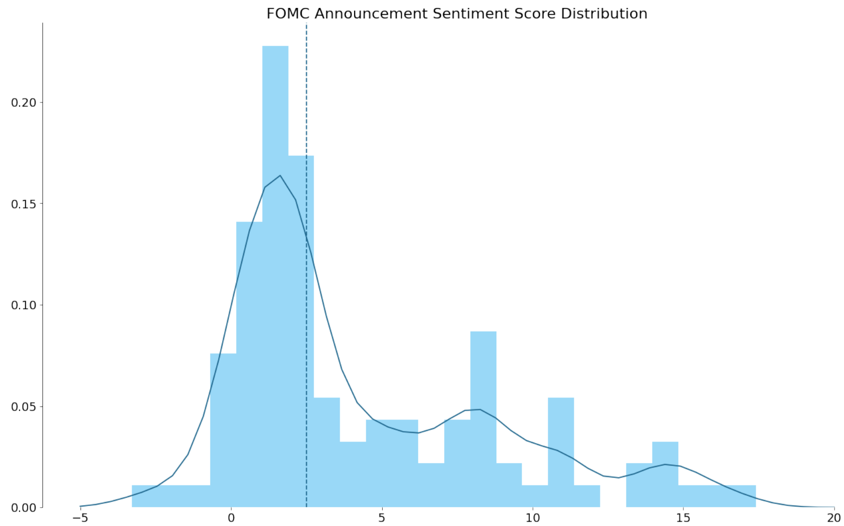
The final and most crucial point to consider is that, compared to normal days, the public sentiment tends to become more positive when the FED makes an announcement, and even more positive when the announcement is relatively positive in tone. It's important to note that this result controls for yield curve shocks. While it may seem natural that the tone of the FED's announcement would impact public sentiment in the same direction, the finding that the mere act of making an announcement leads to more positive sentiment is somewhat puzzling. One possible explanation is that the FED's use of relatively neutral or negative language is not entirely negative. This can be seen by examining the distribution of the latent sentiment score for FED announcements in Figure 6. While there are a few announcements with genuinely negative sentiment, the latent score for most announcements characterized as neutral or slightly negative is actually quite positive. As a result, the public tends to respond positively to these announcements, though perhaps not as strongly as they do to genuinely positive announcements.

Using this regression, we can quantify the changes in public sentiment that are specifically induced by the FOMC announcement and independent of yield curve shocks. Specifically, on normal days, the shock will be zero. On days when the FOMC makes an announcement, the shock is set to $\gamma_1$ for relatively neutral or negative announcements and $\gamma_1 + \gamma_2$ for relatively positive announcements. This definition ensures that the shock reflects the sentiment changes solely related to the FOMC meetings and is not influenced by information about interest rates or the future path of interest rates.

As a preliminary analysis, we examine the relationship between intraday returns of the S&P 500 index and the identified sentiment shock, while controlling for yield curve shocks. The specific regression equation we use is:

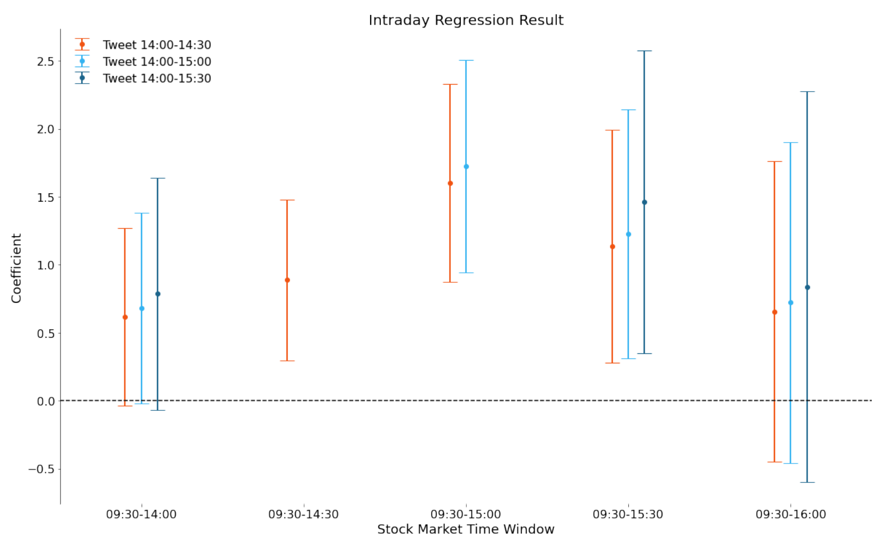
$$RET_{t_0,t_1+\Delta t} = \beta_{0,\Delta t} + \beta_{1,\Delta t}\Delta S_{t_0,t_1} + \beta_{2,\Delta t}X_t + \epsilon_t \tag{3}$$

where $RET_{t_0,t_1+\Delta t}$ represents the returns of the S&P 500 index within a single day




*Note:* The figure presents the distribution of latent sentiment scores for FOMC announcements between January 2009 and June 2019. These sentiment scores are calculated using a similar process as the one used to obtain sentiment scores from Twitter text. The dashed line represents the median of the distribution, while the blue bar illustrates the histogram and the blue solid line shows the kernel density function of the distribution, using a gaussian kernel with a bandwidth of 1.

FIGURE 6. DISTRIBUTION OF THE LATENT SENTIMENT SCORE OF THE FOMC ANNOUNCEMENT




*Note:* The plot displays the regression coefficient $\beta_1$. The vertical segment denotes the 90% confidence intervals. The x-axis illustrates the stock market return from 9:30 AM to a specific time point following the announcements. The plot is presented in different colors, each representing a regression in which the definition of the FOMC-induced sentiment shock is defined using a distinct time interval.

FIGURE 7. INDEX LEVEL REGRESSION COEFFICIENT

from time $t_0$ to time $t_1 + \Delta_t$. $\Delta S_{t_0,t_1}$ denotes the sentiment shock identified from time $t_0$ to time $t_1$. It's important to note that $t_0$ is 9:30 AM, the opening time of the stock market, while $t_1$ and $t_1 + \Delta t$ vary from 2:30 PM to 4:00 PM, the closing time of the stock market. It's also worth noting that $t_1 \leq t_1 + \Delta t$ by definition. In addition, the regression controls for the three monetary policy shocks identified in Swanson (2020), as well as a linear yearly trend and quarterly effects. The primary focus of the regression is the coefficient $\beta_{1,\Delta t}$, which shows the impact of the sentiment shock induced by the FOMC on intraday stock market returns.

As illustrated in Figure 7, a positive sentiment shock related to the US economy results in an immediate increase in stock market returns, which is contemporaneous with the sentiment shock. This increase is short-lived and returns to its baseline level within approximately 90 minutes, coinciding with the duration of the sentiment shock. The coincidence of these events makes it challenging to determine whether the changes in public sentiment are the cause of the changes in stock market returns. In the following section, we will focus on addressing this issue.

## IV. Evidence of the Social Media Channel

In this section, I will delve into the identification of the social media channel in the transmission of sentiment shocks induced by the FOMC. I will also present the results of placebo tests, which serve to enhance the credibility of the main regression results.

### *A. Identification of the Social Media Channel*

The positive sentiment changes observed in response to the FOMC announcement coincide with an increase in stock market returns. However, this does not necessarily prove the existence of a causative relationship or the existence of the indirect channel through social media depicted in the first hypothesis in Figure 1. Remember that the social media channel depicts a story that stock traders pay attention to the text on twitter and adjust their positions accordingly.

It is worth noting that other explanations for the correlation between sentiment changes and stock market returns are possible. For example, it is possible that the FED announcement has a direct impact on the public sentiment, leading to immediate reactions from stock traders. At the same time, people's sentiment about the economy may be reflected in tweets posted online, resulting in the coincidence of sentiment changes and stock market returns without the traders paying close attention to these tweets. This scenario aligns with the second hypothesis in figure 1.

Alternatively, it is also possible that the causality may be reversed, with stock market returns reacting quickly to FOMC announcement tones and inducing changes in sentiment among the general public. This corresponds to the third hypothesis in figure 1.

In reality, it is difficult to determine whether the social media channel plays a role in the relationship between FOMC announcements, sentiment changes, and stock market returns, as all three channels may potentially exist.

To illustrate the process of solving the identification problem, let us consider two hypothetical scenarios: the "real" world in which market participants have access to Twitter, and the "counterfactual" world in which market participants are denied this access but econometricians are not. In the counterfactual world, we eliminate the influence of Twitter on stock market trading. If only hypothesis 1 is true and breaking this connection would affect stock market returns, we would expect to see a reduced response to the FOMC induced sentiment changes in the counterfactual world compared to the real world. On the other hand, in hypothesis 2 and 3, there is no such influence from Twitter sentiment to the stock market to begin with, so there is nothing to disrupt. Therefore, the stock market's response should be similar in both worlds. This is the key idea of the paper. The distinctive difference between the real world and the counterfactual world in the case of hypothesis 1 allows us to identify the presence of the social media channel. It is worth noting that in reality, all three channels may be present, but this will not hinder our ability to identify the first channel as long as we observe a discrepancy in the stock market's response to FOMC-induced sentiment changes between the real and counterfactual worlds.

To solve the identification problem, we need to create the counterfactual scenario in which only the causal link between Twitter sentiment and stock market returns is broken, while all other relationships remain unchanged. One way to do this is by using stock level data and comparing the response of stocks with high exposure to Twitter to the response of stocks with low exposure to Twitter. We can assume that different stocks have varying levels of exposure to Twitter, and if we can directly observe this exposure, we can use it to construct a counterfactual scenario. Specifically, stocks with high exposure to Twitter will represent the real world, where the relationship between Twitter sentiment and stock market returns is intact, while stocks with low exposure will represent the counterfactual world, where this relationship is broken. By comparing the response of these two groups of stocks to sentiment shocks induced by the FOMC, we can solve the identification problem and determine the causal effect of Twitter sentiment on stock market returns.

The question now becomes how to measure the exposure of different stocks to Twitter. To do this, we will use a proxy of the number of tweets about a specific stock in the past 7 days as a measure of the exposure of that stock to Twitter. The assumption here is that people who post tweets about a specific stock are likely to be potential traders of that stock, and that after posting a tweet about a specific stock on Twitter, they may also browse other economic news. As a result, they may be more affected by sentiment changes induced by the FOMC on Twitter. If more people are tweeting about one stock than another, then the exposure of this stock is higher than the others. We search for the number of tweets related to the stock tickers[3] and count the number of tweets for each stock in the S&P 500 stock index. In total, we search for 805 stocks from January 2009 to June 2019. To enhance the robustness of our analysis and to consider the relative exposure and changes over time, we define a stock's exposure to Twitter as high if the number of tweets is higher than the median among all stocks for each week.

There are two important points to note. First, the proxy uses a different set of Twitter data from that used to calculate Twitter sentiment changes. When calculating Twitter sentiment changes, we search for key words related to the economy and its variations, which relate to tweets about the aggregate economy. However, the proxy for stock-level exposure to Twitter depends on tweets about a specific stock, rather than the aggregate economy. This helps to break the one-directional causality from Twitter sentiment changes to stock market returns. One might argue that the exposure to Twitter may also break the link from stock market returns to Twitter sentiment changes, rendering the identification strategy useless. However, this is only true if sentiment changes are also measured at the stock level. Since they are actually measured at the aggregate level, there is less

[3] Twitter uses a dollar sign ($) to label the stock tickers.

concern about a broken reverse linkage.

Second, the data used to calculate the proxy is lagged by one week. We utilize lagged data to break the simultaneity between stock returns and exposure measurement. It is possible that when the FOMC releases an unexpected sentiment-altering announcement, stocks that react most to the shock will generate the most discussion on social media. In this case, using the number of tweets posted after the FOMC announcement would not be a reliable indicator of a stock's exposure to Twitter. By contrast, using lagged data as a proxy helps to mitigate this issue. It is unlikely that people will discuss events that have not yet occurred. If a firm has been discussed extensively in the recent past, individuals who were previously focused on the stock may consider trading it in response to a shock in the economy when the aggregate shock occurs. Therefore, the number of recent tweets can be seen as a natural indicator of a firm's exposure to Twitter.

With the well-defined measurement of stock exposure to twitter, the following regression is used to analyze its effect on the daily returns of stock $i$ on date $t$ ($\Delta y_{i,t}$):

$$
\begin{aligned}
\Delta y_{i,t} =& \beta_0 + \beta_1 \Delta S_t + \beta_2 \Delta S_t \times 1(\text{High Expo})_{i,t} \\
&+ \beta_3 1(\text{High Expo})_{i,t} + \beta_4 X_{i,t} + \epsilon_{i,t}
\end{aligned}
\tag{4}
$$

where $\Delta S_t$ represents the FOMC induced sentiment changes, and the dummy variable $1(\text{High Expo})i, t$ indicates whether the stock has a relatively high exposure to Twitter. The controls for the regression $(Xi, t)$ include the three yield curve shocks identified in Swanson (2021), as well as the interaction term between the firm's exposure to Twitter and the yield curve shocks. This controls for the potential effect of the transmission of yield curve shocks through the social media channel on the transmission of sentiment shocks. The regression uses panel data at the stock level, with daily data covering the FOMC announcement weeks. It also includes firm fixed effect and yearly and quarterly fixed effects, and the standard errors are clustered at the 2-digit SIC code level. The result of the regression is presented in column 1 of Table 2.

The primary focus of the regression is the coefficient $\beta_2$, which is in front of the interaction term. The null hypothesis is that $\beta_2 = 0$, which means that the stock returns of firms with high or low exposures to Twitter will react to sentiment changes in the same way. On the other hand, if $\beta_2 > 0$, it indicates that firms

Table 2—Main Econometrical Model

| | (1) | (2) | (3) |
|---|---|---|---|
| Sentiment Shock | 0.001 | 0.045*** | 0.027*** |
| | (0.001) | (0.007) | (0.007) |
| Sentiment Shock × High Exposure | 0.004* | 0.007*** | 0.005** |
| | (0.002) | (0.002) | (0.002) |
| FFR Shock × High Exposure | -0.001 | -0.001 | 0.002 |
| | (0.002) | (0.001) | (0.002) |
| FGF Shock× High Exposure | 0.001 | 0.001 | 0.000 |
| | (0.001) | (0.001) | (0.000) |
| LSAP Shock× High Exposure | 0.000 | 0.000 | 0.001 |
| | (0.001) | (0.001) | (0.000) |
| Yield Curve Shocks × Firm Level Controls | N | N | Y |
| Sentiment Shocks × Firm Level Controls | N | Y | Y |
| Firm Level Controls | Y | Y | Y |
| Observations | 244,878 | 244,878 | 244,878 |
| R-squared | 0.007 | 0.009 | 0.011 |

*Note:* This table displays the key regression result of the basic regression in equation 4. The clustered standard errors are shown in parentheses. The dependent variable in the three regressions is the stock returns. The coefficients of the firm fixed effect, yearly effect, the quarterly effect and the firm level controls are omitted. ***: p<0.01, **: p<0.05, *: p<0.1.

with higher exposure to Twitter will have a greater response to the sentiment shocks induced by FOMC. As shown in column 1 of Table 3, the coefficient $\beta_2$ is significantly positive, indicating that there is a difference between the two groups of stocks and providing strong evidence that the social media channel can amplify the effect of monetary policy shocks.

It is possible that a firm's exposure to Twitter could be correlated with other observable features of the firm, such as the firm size. For example, Tesla is a larger company than other electric car manufacturers, so it may naturally get more exposure on Twitter and receive more attention. If we do not take this into account, we may not be fully convinced that firms with high exposures to Twitter are more responsive to sentiment shocks, and therefore we cannot conclude that the social media amplification channel exists. To address this concern, we include direct controls for firm-level observables that may be correlated with the firm's exposure to Twitter. We also include an interaction term between the firm-level controls and the shocks, allowing us to compare stocks with similar observable features in the two groups. The firm-level controls follow Duchin, Ozbas and Sensoy (2010) and are obtained from the linked dataset of CRSP and Compustat. The data is taken from the most recent quarter's financial report. The regression results are presented in column 2 of Table 2. As a robustness check, the interaction term between the yield curve shocks and the firm-level controls is also included in the analysis, and the results are shown in column 3 of Table 2.

The results of the analysis indicate that, after controlling for sentiment shocks, the coefficient $\beta_2$ remains significantly positive, indicating that the original relationship between the variables is robust. Additionally, controlling for firm-level controls revealed that the interaction terms between yield curve shocks and social media exposure are not significant and nearly zero. This suggests that social media only seems to have an effect on sentiment changes caused by FOMC meetings, rather than on yield curve shocks, which are information shocks that can fundamentally alter people's understanding of the future economy. This result makes sense because it suggests that while people's understanding of the economy may not be easily influenced by social media, their sentiment can be easily affected by others on social media.

### *B. Placebo Test*

In order to make sure that the result that the stocks that have high exposure to Twitter response more to the sentiment shocks is not accidentally a coincidence, I run a few placebo test to reassure this result.

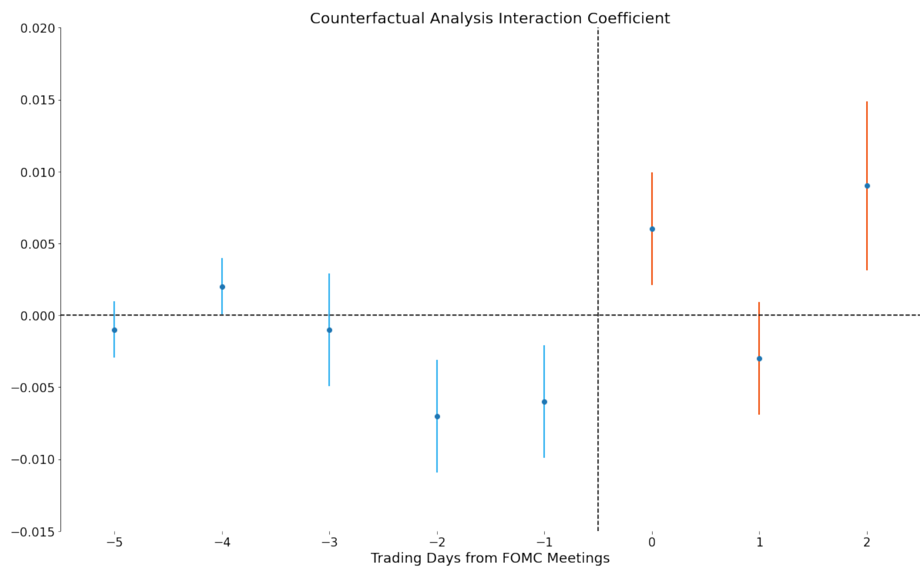
The first thing to note is that most FOMC announcements occur on Wednesdays. This raises the possibility that weekly patterns in the stock market could be driving the results of the regression. To address this concern, we conducted a placebo test using stock return data from the week before the FOMC announcement week and kept the shocks the same as in the baseline regression. This created a counterfactual dataset in which any effects of sentiment shocks on stock returns should not differ between stocks with high and low social media exposure. If we still observed a significantly positive coefficient in this test, it would suggest that the pattern seen in the baseline regression is present every Wednesday and therefore our main conclusion would be less convincing. However, as shown in Table 3, the counterfactual regression revealed that there was no significant difference in stock returns between stocks with high and low social media exposure, indicating that our baseline regression is not trivially picking up a weekly pattern and is capturing the amplification of FOMC-induced sentiment shocks through the social media channel.

The second thing to consider in our analysis, which is similar to a difference-in-

Table 3—Placebo Test Result

| | (1) | (2) | (3) |
|---|---|---|---|
| Sentiment Shock | -0.001 | -0.026*** | 0.026** |
| | (0.001) | (0.009) | (0.011) |
| Sentiment Shock × High Exposure | -0.002 | -0.006* | 0.003 |
| | (0.002) | (0.003) | (0.002) |
| FFR Shock × High Exposure | 0.006** | 0.005** | -0.005* |
| | (0.003) | (0.002) | (0.003) |
| FGF Shock × High Exposure | 0.000 | 0.000 | -0.000 |
| | (0.000) | (0.000) | (0.000) |
| LSAP Shock × High Exposure | 0.001 | 0.001 | -0.001** |
| | (0.001) | (0.001) | (0.001) |
| Interaction of Interest Rate Shocks with the Firm Level Controls | N | N | Y |
| Interaction of Sentiment Shock with the Firm Level Controls | N | Y | Y |
| Firm Level Controls | Y | Y | Y |
| Observations | 243,663 | 243,663 | 243,663 |
| R-squared | 0.007 | 0.007 | 0.011 |

*Note:* This table displays the first placebo test results of the basic regression. The regression is still equation 4, but with counterfactual data one week ago. The clustered standard errors are shown in parentheses. The dependent variable in the three regressions is the stock return. The coefficients of the firm fixed effect, yearly effect, the quarterly effect and the firm level controls are omitted. ***: p<0.01, **: p<0.05, *: p<0.1.




*Note:* The figure presents the coefficients in front pf the interaction between the counterfactual sentiment shocks and the dummy variable indicating whether the firm have a high exposure to twitter. The x-axis denotes how many days away from the FOMC meetings. For example, -1 denotes it is one trading day before the FOMC announcement meetings. The segment shows the 90% confidence intervals of the regressions. Each point denotes one counterfactual regression.

Figure 8. Distribution of the Latent Sentiment Score of the FOMC Announcement

difference approach, is the need to test for the parallel assumption. To do this, we used a placebo test and analyzed data from the last Wednesday to Friday of the FOMC announcement week. We created several counterfactual shocks by moving the true shocks that would have occurred during the FOMC announcement dates ($t$) to different dates ($t'$). If the parallel trend assumption holds true, the counterfactual shocks and corresponding regressions should not show any significant positive difference between the high exposure firms and low exposure firms. As Figure 8 shows, this was the case before the true announcement date, as none of the coefficients of the placebo regressions were significantly positive. The direction of the interaction coefficient even suggests that the main effect identified in the baseline analysis would be even larger. Figure 9 also reveals dynamic effects by showing a positive difference in daily returns between the high exposure stocks and low exposure stocks for t+2 after moving the counterfactual shocks to dates after the event. It's important to note that the dependent variables in the regressions are daily returns, meaning that the prices of the stocks will differ during the announcement days and for a further two days after the event.

In conclusion, our analysis found that stocks with high exposure to Twitter tend to respond more to sentiment shocks. This result was validated through several placebo tests, which showed that it was not simply a coincidence and that the parallel assumption held true. The tests also indicated the presence of dynamic effects, with a positive difference in daily returns between high and low exposure stocks observed for two days after the event. These findings support the idea that sentiment shocks are amplified through the social media channel for stocks with high exposure to Twitter.

## V. Conclusion

The study analyzed the relationship between public sentiment about the US economy, as measured by the Twitter sentiment index, and stock market returns. To measure the impact of Twitter sentiment on stock market returns, the study used a high-frequency identification method to detect changes in sentiment induced by FOMC announcements and a linear regression to isolate the variations in public sentiment caused solely by the tone of the FOMC announcement and remove the effect of yield curve shocks. The study found that a relatively neutral FOMC announcement had a statistically significant positive effect on changes in public sentiment, and that a positive sentiment shock related to the US economy resulted in an immediate increase in stock market returns that lasted for about 90 minutes. This suggests that there may be a causal relationship between public sentiment about the US economy and stock market returns, at least in the short term.

To investigate the causal relationship between Twitter sentiment and stock market returns, this study proposes using stock-level data to compare the daily returns of stocks with high exposure to Twitter to those with low exposure. The results show that Twitter sentiment has a statistically significant and economically meaningful impact on stock market returns, with a stronger impact for stocks with higher exposure to Twitter. This suggests that market participants do consider Twitter sentiment when making investment decisions, and that the social media channel plays a role in the relationship between Twitter sentiment and stock market returns.

One potential direction for further research is to more accurately identify the sentiment change induced by economic news. For example, this can be done by applying the methods used in Boehm and Kroner (2021) and Gürkaynak, Kisacikoğlu and Wright (2020) to produce a FOMC non-yield sentiment shock, and then studying the effect of this shock on asset returns. As noted in Baker et al. (2021), tweet data can be useful in constructing high-frequency indicators and in relating tweet-based measures to financial market responses. This study represents a novel contribution in this direction, but further research could improve upon the identification of FOMC-induced sentiment change.

## REFERENCES


**Angeletos, George-Marios, Fabrice Collard, and Harris Dellas.** 2018. "Quantifying Confidence." *Econometrica*, 86(5): 1689–1726. I

**Angelico, Cristina, Juri Marcucci, Marcello Miccoli, and Filippo Quarta.** 2022. "Can We Measure Inflation Expectations Using Twitter?" *Journal of Econometrics*, 228(2): 259–277. I

**Baker, Scott R, Nicholas Bloom, Steven J Davis, and Thomas Renault.** 2021. "Twitter-Derived Measures of Economic Uncertainty." 14. I, II.A, II.B, V

**Barsky, Robert B., and Eric R. Sims.** 2012. "Information, Animal Spirits, and the Meaning of Innovations in Consumer Confidence." *American Economic Review*, 102(4): 1343–1377. I

**Boehm, Christoph, and Niklas Kroner.** 2021. "Beyond the Yield Curve: Understanding the Effect of FOMC Announcements on the Stock Market." I, III, V

**Duchin, Ran, Oguzhan Ozbas, and Berk A. Sensoy.** 2010. "Costly External Finance, Corporate Investment, and the Subprime Mortgage Credit Crisis." *Journal of Financial Economics*, 97(3): 418–435. IV.A

**Gu, Chen, and Alexander Kurov.** 2020. "Informational Role of Social Media: Evidence from Twitter Sentiment." *Journal of Banking & Finance*, 121: 105969. I

**Gürkaynak, Refet S., Brian Sack, and Eric Swanson.** 2005. "The Sensitivity of Long-Term Interest Rates to Economic News: Evidence and Implications for Macroeconomic Models." *American Economic Review*, 95(1): 425–436. I

**Gürkaynak, Refet S., Burçin Kisacikoğlu, and Jonathan H. Wright.** 2020. "Missing Events in Event Studies: Identifying the Effects of Partially Measured News Surprises." *American Economic Review*, 110(12): 3871–3912. I, V

**Hamraoui, Imen, and Adel Boubaker.** 2022. "Impact of Twitter Sentiment on Stock Price Returns." *Social Network Analysis and Mining*, 12(1): 28. I

**Handlan, Amy.** 2020. "Text Shocks and Monetary Surprises:." 65. I, II.B

**Howard, Jeremy, and Sebastian Ruder.** 2018. "Universal Language Model Fine-tuning for Text Classification." 2, B

**Levchenko, Andrei A, and Nitya Pandalai-Nayar.** 2020. "Tfp, News, and "Sentiments": The International Transmission of Business Cycles." *Journal of the European Economic Association*, 18(1): 302–341. I

**Lewis, Daniel J., Christos Makridis, and Karel Mertens.** 2019. "Do Monetary Policy Announcements Shift Household Expectations?" Social Science Research Network SSRN Scholarly Paper 3449467, Rochester, NY. III

**Lüdering, Jochen, and Peter Tillmann.** 2020. "Monetary Policy on Twitter and Asset Prices: Evidence from Computational Text Analysis." *The North American Journal of Economics and Finance*, 51: 100875. I

**Makrehchi, Masoud, Sameena Shah, and Wenhui Liao.** 2013. "Stock Prediction Using Event-Based Sentiment Analysis." Vol. 1, 337–342. I

**Masciandaro, Donato, Davide Romelli, and Gaia Rubera.** 2022. "Monetary Policy and Financial Markets: Evidence From Twitter Traffic." Social Science Research Network SSRN Scholarly Paper 3823185, Rochester, NY. I

**Milani, Fabio.** 2017. "Sentiment and the U.S. Business Cycle." *Journal of Economic Dynamics and Control*, 82: 289–311. I

**Oliveira, Nuno, Paulo Cortez, and Nelson Areal.** 2017. "The Impact of Microblogging Data for Stock Market Prediction: Using Twitter to Predict Returns, Volatility, Trading Volume and Survey Sentiment Indices." *Expert Systems with Applications*, 73: 125–144. I

**Ranco, Gabriele, Darko Aleksovski, Guido Caldarelli, Miha Grčar, and Igor Mozetič.** 2015. "The Effects of Twitter Sentiment on Stock Price Returns." *PLOS ONE*, 10(9): e0138441. I

**Rao, Tushar, and Saket Srivastava.** 2012. "Analyzing Stock Market Movements Using Twitter Sentiment Analysis." 119–123. Washington, DC, USA:IEEE Computer Society. I

**Shapiro, Adam Hale, Moritz Sudhof, and Daniel J. Wilson.** 2022. "Measuring News Sentiment." *Journal of Econometrics*, 228(2): 221–243. II.C

**Si, Jianfeng, Arjun Mukherjee, Bing Liu, Qing Li, Huayi Li, and Xiaotie Deng.** 2013. "Exploiting Topic Based Twitter Sentiment for Stock Prediction." 24–29. Sofia, Bulgaria:Association for Computational Linguistics. I

**Swanson, Eric T.** 2021. "Measuring the Effects of Federal Reserve Forward Guidance and Asset Purchases on Financial Markets." *Journal of Monetary Economics*, 118: 32–53. I, III, IV.A

## Appendix A: Data Cleaning Process

Before using tweet text data in a machine learning model or determining the sentiment score of a tweet, it is important to clean the text. This process involves several steps to ensure that the data is relevant and accurate for the task at hand.

The first step in text cleaning is to convert the time label from Universal Time (UTC) to the desired time zone. This is important if the tweets are being analyzed within a specific geographical region and the time zone of the tweets needs to be adjusted accordingly.

Next, it is important to remove any elements that may add noise to the text and are not relevant for sentiment analysis. This includes emojis, hyperlinks, user names (denoted by the symbol "@"), hashtags (denoted by the symbol "#"), and stock or asset tickers (denoted by the symbol "$").

Handling numbers is also an important step in the text cleaning process. This involves formatting numbers in a way that is consistent with the rest of the text. For example, changing money signs to indicate amounts and replacing years and months with tags can help the model better understand the context of the numbers in the text.

Abbreviations and slangs can be confusing for a machine learning model, as they often have multiple meanings or are specific to certain regions or groups. Expanding abbreviations and replacing slangs with their true meanings can help the model better understand the text.

To address spelling errors and properly segment words that may have been split incorrectly, it is important to run the text through an auto-correction and word segmentation process.

Lemmatization is the process of reducing words to their base form, which can help the model generalize better by reducing the number of unique word forms it needs to consider. Converting the words to lowercase can also help the model better identify similar words, as "inflation" and "Inflation" would be treated as the same word after being converted to lowercase.

Finally, it may be useful to filter out tweets that are extremely short, as they may not contain enough information for the model to accurately analyze. For example, excluding tweets with fewer than five words can help ensure that the model has enough context to make accurate predictions.

Overall, text cleaning is an important step in the process of using tweet text data for sentiment analysis. By following these steps, I can ensure that the data is relevant and accurate, which can lead to more accurate and reliable results.

## Appendix B: Summary of the ULMFiT Model

In order to predict the geographical origins of tweets, I used a ULMFiT machine learning model. ULMFiT, or Universal Language Model Fine-tuning, is a method developed by researchers at fastai that involves fine-tuning a pre-trained language model rather than training a model from scratch. This approach is efficient and effective, making it well-suited for our task. The model was trained using Python and the fastai package, and the training process was accelerated using CUDA and GPU. The training process consisted of three steps:

1. Language model pre-training: The first step in the ULMFiT process is to pre-train a language model on a large, general-purpose dataset. In this case, fastai used text data from Wikipedia for pre-training. During this stage, the language model is trained to predict the next word in a sequence of words, and the model's predictions are optimized by minimizing the cross-entropy loss.

2. Language model fine-tuning: After pre-training, the language model is fine-tuned on the task-specific dataset, which in this case is the cleaned Twitter text. The model is still trained to predict the next word in a sequence of words.

3. Classifier fine-tuning: In the final stage of training, the model is fine-tuned to perform the classification task. The input to the model is the raw text of the whole tweet, and the output is the geographical label in the training dataset.

To ensure optimal performance, the learning rate was chosen using the suggested optimal learning rate, and the model was gradually unfrozen during training. This technique involves first freezing the parameters of the last layer of the model, then fine-tuning only those parameters before gradually unfreezing additional layers. This process helps to train a model with better prediction accuracy. To learn more about the ULMFiT model, readers can refer to the paper Howard and Ruder (2018).

## Appendix C: Firm Level Control Definition

In the extended key regression, we control for firm-level features to ensure that they do not interfere with the analysis of the firm's exposure to Twitter. The main variables included in the regression, along with their definitions, are listed below:

Cash Ratio = current cash (cheq) / Total Asset (atq)

Short Term Debt Ratio = Short Term Debt (dlcq) / Total Asset (atq)

Long Term Debt Ratio = Long Term Debt (dlttq) / Total Asset (atq)

Cash Flow Ratio = Operating Income Before Depreciation (oibdpq) / Total Asset (atq)

Tobin' Q = (Total Asset (atq) + Number of Common Shares Outstanding (cshoq) * Close Price (prccq) - Ordinary Equity (ceqq) - Deferred Taxes (txdbq)) / (0.9*Total Asset (atq) + 0.1*(Total Asset (atq) + Number of Common Shares Outstanding (cshoq) * Close Price (prccq) - Ordinary Equity (ceqq) - Deferred Taxes (txdbq)))

Selling, General and Administrative Expense Ratio = SGA (xsgaq) / Sales (saleq)

Reserch and Development Cost Ratio = R&D Cost (xrdq) / Total Asset (atq)

Net Working Capital Ratio = (Current Assets (actq) - Current Liabilities(lctq) - Cash and Short-Term Investments(cheq)) / Total Asset (atq)

Inventory Ratio = Inventory (invtq) / Total Asset (atq)

Tangibility = Gross Property, Plant and Equipment (ppegtq) / Total Asset (atq)

Return of Asset = Income Before Extraordinary Items (ib) / Total Asset (atq)

Firm Size = log(Sales (saleq))

Investment Ratio = Capital Expenditures (capx) / Total Asset (atq)

All the firm level controls are winsorized by quarter at 1 and 99 percentiles.