What Can we Learn from Stock Market Tweets? *

Vahid Gholampour

Bucknell University

Abstract

I develop a finance dictionary to identify and classify tweets that express a directional prediction for the stock market index. The opinionated tweets are used to create a daily Twitter Sentiment Index (TSI) and Dispersion of Sentiment Index (DSI). The results indicate that over 932 trading days, (1) follower-weighted sentiment predicts the daily stock market returns and the effect is reversed in two trading days, (2) the TSI is positively correlated with weekly changes in net long position of investment managers and negatively correlated with those of hedge funds, and (3) disagreement among investors is positively related to return volatility and abnormal trading volume.

*I gratefully acknowledge financial support from the Bankard Fund for Political Economy and Radulovacki Fund. I like to thank Eric Young, Eric van Wincoop, Michael Gallmeyer, Richard Evans, Toshihiko Mukoyama and seminar participants at the University of Virginia for useful comments. James Elmendorf provided research assistance.

1 Introduction

I construct a daily measure of sentiment and disagreement by aggregating directional forecasts about the S&P500 index expressed on Twitter. I have used Twitter's search tools to download stock market tweets every half an hour since September 2013 and developed a dictionary based on the lexicon used by traders to identify and classify opinionated tweets. The daily measures of sentiment and disagreement are used to empirically test the predictions of some theories in the finance literature.

Sentiment theories, such as Delong et.al. (1990), usually rely on the assumption that a class of investors make rational portfolio decisions based on biased or inaccurate information. These theories predict that the equity prices temporarily rise beyond fundamentals during high sentiment periods and later fall. Consistent with the sentiment theory, a follower-weighted measure of sentiment that aggregates opinions posted before the market open predicts the daily return of the stock market and the effect is almost completely reversed over two days.

Dynamic models of heterogeneous beliefs, such as Basak (2005), Buraschi and Jiltsov (2006), Gallmeyer and Hollifield (2008) and Atmaz and Basak (2016), predict that an increase in disagreement among investors leads to higher return volatility. This paper shows that the daily dispersion of sentiment measured from the opinionated tweets is positively related to return volatility of the stock market index.

Several empirical and theoretical papers, such as Karpoff (1986), Banerjee and Kremer (2010), and Carlin et.al. (2014), highlight the role of disagreement as a motivation to trade and the positive relationship between disagreement and trading volume. It is shown in the paper that the daily disagreement index is positively related to the abnormal trading volume of the S&P500 index ETF.

The Twitter Sentiment Index (TSI) is validated, as a measure of sentiment, by calculating its correlation with a number of survey-based and market-based measures of sentiment. The TSI shows positive correlation with other measures of sentiment and negative correlation with the fear indicators, such as CBOE volatility index (VIX) and put-call ratio of short-term options on the S&P500 index. Moreover, comparing the TSI and Commitment of Traders (COT) data shows that the TSI is positively correlated with weekly changes in net long position of investment managers and negatively correlated with those of hedge funds.

In order to quantify the magnitude of potential gain from trading the stock market index using the Twitter opinions, I consider a trading strategy that takes a position on the index at market open based on the sentiment of opinionated tweets posted before the market open. The trading strategy provides an annualized Sharpe ratio of 1.37 over 932 trading days.

There are several studies that focus on the information content of social media messages but there are three important differences between this study and previous work. First, I use the number of followers of accounts to give different weight to the opinions posted by the accounts. Sentiment surveys, Google search sentiment and other message board measures of sentiment usually assign equal weight to individual opinions. This paper shows that the number of followers is a useful proxy for the quality of private signals. The second distinction is the text analysis methodology used to extract opinions from the tweets. I develop a finance dictionary that is designed to include the lexicon of traders and investors. Several studies, such as Tetlock (2007), Li (2008), Chen et.al. (2014) and Da et.al. (2015) count the number of words in different categories to extract opinions from a text. The papers show that counting the number of words is an effective method to extract the sentiment of a lengthy text such as an article or a newspaper column. Since tweets are limited to only 140 characters and traders often express their views by tweeting their recent trades, counting the number of positive or negative words might not be effective in extracting the opinions from the tweets. Third, the universe of tweets included in this study is limited to those that mention the S&P500 index or an ETF that tracks the index. Other studies, such as Bollen et.al. (2011) and Zhang et.al. (2011), explore the relationship between "Twitter mood" and the stock market returns by using a random sample of tweets that are not necessarily related to the stock market. Mao et.al.(2015) examines the return predictability of tweets that mention the words "bullish" or "bearish".

In recent years, Twitter has become a major source of information and an effective communication tool for investors and public companies. In April 2013, U.S. Securities and Exchange Commission (SEC) issued a press release stating that companies can use social media outlets such as Facebook and Twitter to announce key information.[1] Stock market participants use Twitter

[1] www.sec.gov/news/press-release/2013-2013-51htm

to share their information with others and receive real-time information about the stock market and individual companies. There are several anecdotal stories that highlight the role of Twitter in providing information to the market and significant effects on stock prices following the release of information on Twitter. For instance, Carl Icahn, an activist investor, tweeted about his large position in Apple on August 13th 2013. As the result, the stock surged by over four percent in a few seconds. Almost two years later on April 28th 2015, a data mining company obtained Twitter's quarterly earning and posted it on Twitter before the scheduled release time. Twitter's stock plummeted by twenty percent following the early release of its earning and trading was halted by NYSE. Given the widespread use of Twitter among stock market participants, it is a natural choice to study the link between expectations of future returns and asset prices.

Text analysis is an important element of extracting useful information from internet message boards and social media networks. Antweiler and Frank (2004) use Naïve Bayes as the main algorithm to classify messages. Tetlock (2007) and Da et.al. (2015) rely on the Harvard IV-4 Dictionary to count the number of words in different categories. Both methods are shown to be effective in practice given the data sources used in the studies. In this paper, a different approach is used to classify the tweets. A number of word combinations are defined as indicators for bullish, bearish and neutral tweets. If a tweet contains one of the bullish word combinations and does not mention any of the bearish or neutral word combinations, it is placed in the most bullish category and is associated with the numeric score of +1. The details of message classification is provided in section 2. Given that many traders express their opinion by announcing their current option or ETF positions and the Harvard dictionary is not structured for the vocabulary of traders, using a finance dictionary could be more effective than alternative methods in extracting investor sentiment from the tweets.

The TSI has two important advantages over investor sentiment surveys. First, surveys assign equal weight to individual opinions when aggregating the responses. Twitter data provides the number of followers of the account that posts a tweet. The number of followers can be used as an indication for the quality of private signals and be taken into account when aggregating the opinions. The regression results show that a follower-weighted sentiment index predicts the daily open-to-close return of the S&P500 index. The equal-weighted index, however, shows no predictive power for the daily stock market returns. An increase in the level of bullishness in Twitter messages prior to market open on a given day also predicts lower volatility and negative returns for volatility futures on that day. Second, survey participants have no incentive to express their true opinion because the individual opinions are not disclosed with survey results. Individuals who have a large number of followers have the incentive to tweet their true opinion to maintain their reputation. In addition, survey results are usually reported in weekly or monthly frequency while the Twitter sentiment can be measures in daily frequency.

2 Data

I start with tweets that mention the S&P500 index or an Exchange Traded Fund (ETF) that tracks the daily return of the index. For instance, tweets that mention the index or any of the ticker symbols "SPY", "SSO", "SDS", "UPRO", or "SPXU" are included in the dataset. These tickers represent the ETFs that correspond to unleveraged and leveraged long and short positions on the S&P500 index.[2] I have used Twitter's publicly available search tools to download the tweets since September 18th 2013. Every half an hour, a search query is sent to Twitter's server and tweets that mention at least one of the search words in their text are downloaded and stored. The Twitter data sample includes 4,073,428 tweets posted over 932 trading days between September 18th 2013 and May 31st 2017.

Since most of the tweets that mention the stock market index or its ETFs provide no opinion about future returns, opinionated tweets should be identified to construct a measure of sentiment and disagreement based on the expectations of future returns. I studied a large number of tweets and developed a dictionary based on the lexicon used by traders and investors. The dictionary was then used to identify the opinionated tweets in the sample. In order to make sure that the dictionary identifies and classifies the opinionated tweets correctly, I randomly selected 100 tweets and compared the results of automatic classification to those of manual classification and made changes to the dictionary to improve the automatic classification. This process was repeated multiple times until the dictionary correctly classified over 90 percent of the tweets. The main

[2] Tweets that mention at list one of the following are included in the dataset: sp500, s&p500,s&p 500, $spx, $spxu, $spxs, $spy, $sso, $sds, $upro, $ivv, $sh, $voo

idea behind developing a special purpose finance dictionary is to capture certain phrases or word combinations that have a different meaning outside of the financial markets. For instance, buying a put option on the stock market index is a sign of bearishness and a tweet that says "bought $SPY put 0.35" should be classified as negative but a naïve dictionary might classify the tweet as positive because "bought" and "$SPY" are mentioned in the same tweet. The text analysis methodology adopted in this paper captures the type of words that traders and investors use to describe their current position or outlook for the stock market. Opinionated tweets are identified using 433 word combinations. For instance, an optimistic trader could buy call option on the S&P index ETF (SPY) so a tweet that contains the words "bought", "spy", "call" in this order is placed in positive category. If a tweet contains the words "increase", "spy", "short" in this order is identified as negative because it indicates that a trader expects further drop in equity prices and is willing to increase the size of an existing short position. The words "will", "buy", "if" put a tweet in neutral category because they indicate a decision to buy the equity index conditional on some event. Table 1 provides examples of actual positive, neutral, and negative tweets. The placement of the tweets in the categories is independent of daily returns in the stock market and is determined only by the content of tweets. The dictionary finds 128,488 opinionated tweets over 932 trading days in the sample.[3] Figure 1 shows the distribution of the daily number of opinionated tweets. Figure 2 provides the distribution of the number of followers of 24,304 unique accounts that posted at least one opinionated tweet during the sample period.

A number of related studies, such as Tetlock (2007) and Da et.al. (2015), use the Harvard IV-4 Dictionary to carry out text analysis. The dictionary assigns a large number of words to 182 categories such as "Positive", "Strong", and "Weak". Each word is assigned to one or more categories in the Harvard dictionary and one could analyze a text by counting the number of words assigned to different categories. The text analysis methodology of this study is not based on the Harvard dictionary because traders usually express their opinion using words that convey a different message according to a general language dictionary such as the Harvard dictionary. For example, traders often use the word "Bullish" to express a positive outlook. The closest word to bullish in the Harvard dictionary is "Bull" which is placed in "Male" category.

[3] Multiple tweets from the same account during a day are counted as one opinion.

3 Twitter Sentiment and Stock Market Returns

In order to investigate the relationship between the expectations of future returns and stock market returns, we need to create a measure of sentiment by aggregate the predictions expressed in the opinionated tweets. In this section, I describe, in some detail, the methodology of constructing two versions of Twitter sentiment, examine short-term return predictability of the Twitter sentiment, test return reversal prediction of sentiment theories, quantify the potential gain from using the TSI in trading the stock market index and show the relationship between the Twitter sentiment and market variables related to investor sentiment.

3.1 Twitter Sentiment Index (TSI)

Directional predictions are usually expressed in the form of a direct prediction stated in a tweet or announcement of a recent trade. Opinionated tweets are identified using the finance dictionary developed in this study and are assigned to positive, neutral and negative categories. Tweets that express a bullish outlook are placed in the positive category. Neutral tweets indicate that a trader is indecisive or waiting for more information before taking a position in the stock market. Finally, bearish comments in a tweet put it in the negative category. A numerical score is assigned to each opinionated tweet. All positive tweets are coded as $+1$, neutral tweets as 0 and negative tweets as -1. Tweets that mention both positive and neutral word combinations are coded as $+0.5$ and those that mention negative and neutral combinations are coded as -0.5. The TSI for a given time period is the average of the numerical score of the opinionated tweets posted during that period. I construct follower-weighted and equal-weighted TSI for a variety of experiments in this study.

Follower-weighted TSI takes into account the differences between the number of followers of accounts that post opinionated tweets. If we assume that the number of followers of each account is the equilibrium value given by the information marketplace to the opinions posted by that account, weighting individual opinions based on their followers takes into account the differences between the quality of private signals. Although the number of followers might not always reflect the quality of information shared by an individual, it is a useful observable variable for giving more weight to opinions that are potentially viewed by a larger audience. The follower-weighted TSI (denoted by TSI_t^{fw}) is defined as follows:

$$TSI_t^{fw} = \frac{\sum_{i=1}^{n_t} w_i s_i}{\sum_{i=1}^{n_t} w_i},$$

where s_i is the numeric score associated to the tweet i and $s_i \in \{-1, -0.5, 0, +0.5, +1\}$, w_i is the number of followers of the account that posted tweet i and n_t is the number of opinionated tweets posted during the time period t. TSI_t^{fw} by construction gives higher weight to the opinion of individuals with more followers. Accounts that post tweets about stock market and have a large number of followers are typically controlled by professional traders or money managers. As the result, TSI_t^{fw} measures the optimism among professionals. However, TSI_t^{fw} is not dominated by few individuals as evidenced by the distribution of weight of tweets over the number of followers. Figure 3 shows that the opinion of accounts with more than 500,000 followers has an average weight of 0.14 in the daily TSI_t^{fw}.

Equal-weighted TSI assigns equal weight to the opinionated tweets and is a measure of optimism among those who express their outlook in a tweet.The equal-weighted TSI (denoted by TSI_t^{ew}) is a simple average of the numerical score of the opinionated tweets and is defined as follows:

$$TSI_t^{ew} = \frac{\sum_{i=1}^{n_t} s_i}{n_t},$$

By definition, the TSI can take any value between -1 and $+1$. If all the tweets on a given day express a negative outlook for the stock market, the TSI would be -1, which is the most bearish value for the sentiment. Conversely, the TSI takes the most bullish value of $+1$ if every tweet contains a positive outlook.[4] Figure 4 shows the accumulated TSI and the S&P500 index over the sample period.

[4]TSI is assumed to be zero on eight trading days that there is no opinionated tweets.

3.2 Return Predictability

The first question is whether the TSI can predict the stock market returns. I regress the daily open-to-close price changes of a liquid S&P500 ETF, SPDR NYSEARCA:SPY, on the Twitter sentiment and some control variables. In order to avoid the feedback problem caused by the influence of price changes on the sentiment of traders, I construct a pre-market TSI by using only the tweets posted between midnight and 9:30AM U.S. Eastern Standard Time. The pre-market TSI on a given day is created using the exact same method described in section 3.1 but its data universe is limited to tweets posted before the market open on that day. The regression results are reported in Table 2. The first column shows the regression result when the follower-weighted TSI is the explanatory variable and the second column reports the result when the equal-weighted TSI is the independent variable. The regression results highlight the role of assigned weights to individual opinions when aggregating the expectation of investors. The equal-weighted sentiment index does not capture the differences between the quality of private signals posted by the individuals and consequently shows no predictive power for the daily returns of the stock market index. The predictive power of the follower-weighted sentiment index for the daily returns indicates that the marketplace for information effectively identifies individuals with higher quality private signals and gives them more followers. According to the data, one standard deviation increase in the follower-weighted pre-market TSI predicts a return of 5.2 basis points for the S&P500 index[5]. The control variables are one day lagged CBOE volatility index (VIX), five days lagged returns of the S&P500 ETF, changes in ADS business conditions index and changes in Economic Policy Uncertainty (EPU) index.

In order to control for the effect of changes in macroeconomic activities on the stock market, a daily measure of macroeconomic conditions is included in the regressions. Aruoba, Diebold, and Scoitti (2009) use a number of macroeconomic variables of different frequencies and construct the ADS index as a daily measure of business conditions. The ADS index is obtained from the Federal Reserve Bank of Philadelphia and includes seasonally adjusted value of quarterly real GDP, industrial production, manufacturing and trade sales, personal income minus transfer payments, monthly payroll employment, and weekly jobless claims.

[5]one standard deviation change in the pre-market TSI corresponds to 0.432

Uncertainty about economic policy could influence investor sentiment and asset prices. Baker, Bloom, and Davis (2013) define three categories of words associated with Economic Policy Uncertainty (EPU) and count the number of U.S. newspaper articles that mention any of the words in these categories. Baker, Bloom, and Davis (2013) show that their measure of EPU offers a good proxy for uncertainty related to economic policy over time. The EPU index is the control variable that accounts for economic policy uncertainty in the regressions.

3.3 Robustness Checks

This section provides a number of robustness tests for return predictability of the follower-weighted TSI. Table 3 reports the regressions results where more control variables are included.

There are a number of market-based measures that are often used as a gauge for investor sentiment. Brown and Cliff (2004) show that investor sentiment measured by surveys is related to the market measures such as the ratio of advancing to declining stocks and the change in short interest. One might argue that the TSI is influenced by these market indicators and contains no additional information. To ensure that the index is not a reflection of lagged stock market indicators, some of these indicators are included as additional control variables in the first column of Table 3. More specifically, one day lagged the number of advancing stocks divided by the number of declining stocks in NYSE, ARMS index[6], the number of stocks at their 52 weeks high price divided by the number of stocks at their 52 weeks low price, the number of put options on the S&P500 index divided by the number of call options, and the difference between the interest rate of 10 years and 3 months U.S. Treasury bonds are included in the regression. According to the reported result in column (1), including the market indicators does little to change the effect of pre-market TSI on the daily returns.

Since the information flow is continuous and news could influence the prices through the futures market overnight, there is a concern that the TSI reflects the price changes of the futures market. In order to control for the effect of overnight price changes, the difference between the open price of SPY and its close price on the prior day is included as a control variable in the regression. The second column of Table 3 shows that including the overnight price changes does not alter the main

[6] ARMS index is the ratio of the number of advancing to declining stocks divided by their trading volume.

result.

Several empirical studies, such as Keim (1983) and Gultekin and Gultekin (1983), provide evidence for seasonality in the stock market returns. To ensure that the results are not driven by known seasonal effects, month of the year dummies are included in the regression and the results are reported in the third column of Table 3.

The news-based EPU index is included in the regressions to control for the effect of economic policy uncertainty that could influence the sentiment and the stock market returns. Since the twitter data used in this study is related to the time period that Federal Reserve was communicating and implementing its exit strategy from the unconventional monetary policy, one might still be concerned that the results are driven by a few sizable returns following unexpected announcements in the FOMC meetings. Column (4) of Table 3 shows that including a dummy variable for the FOMC rate decision days does not change the sign and significance of the pre-market TSI.

Given that major economic data such as GDP and unemployment are reported at 8:30am EST, the pre-market TSI and the stock market index could change due to surprises in the economic data on the data release dates. To ensure that the predictability of the TSI is not driven by a few data announcement days, an economic news day dummy is included in the regression. The dummy variable is one on the days that GDP, unemployment or inflation data is released and zero otherwise. The regression result is reported in column (5) of Table 3 and shows that the effect of pre-market TSI is not driven by economic news announcement days.

3.4 Short-term Return Reversal

As shown earlier, the pre-market TSI predicts the same day return of the stock market index. The question is whether the return predictability is due to popularity of certain individuals or the predictability indicates slow reaction of investors to existing information that popular accounts share before the market open. If investors trade based on the pre-market popular narrative, the relation between the pre-market TSI and the daily return should be reversed on future days. Sentiment theories, such as Delong et.al. (1990), predict that the equity prices temporarily rise during high sentiment periods and later fall. In a recent study, Da et.al. (2015) show that an increase in the daily volume of internet search for negative economic words is associated with negative contemporaneous return in the stock market but the effect is almost completely reversed over the next two days. They conclude that the internet search behavior of households in the U.S. is consistent with the return reversal prediction of sentiment theories. Following Da et.al. (2015), I test the return reversal prediction by estimating the coefficient of the pre-market TSI in the following regressions:

$$R_{t+j} = \alpha + \beta preTSI_t^{fw} + \gamma Z_t,$$

where R_{t+j} is the daily return of the S&P500 index ETF on day $t+j$ and Z_t is the vector of control variables that includes five lags of daily returns, the S&P500 volatility index (VIX), changes in ADS business conditions index, and changes in the economic policy uncertainty (EPU) index. Table 4 shows that the positive relation between the pre-market TSI and same day return is reversed in two trading days. This result is consistent with the prediction of sentiment theories and findings of Da et.al. (2015) about return reversal on the days that follow a high or low sentiment day.

3.5 Trading based on the TSI

In order to quantify the potential gain from trading based on the TSI, I measure the Sharpe ratio of a trading strategy that places daily bets on the stock market index depending on the sign and magnitude of the pre-market TSI. The trading strategy takes a long (short) position on the S&P500 index at market open if the pre-market sentiment index is positive (negative). The size of the position is proportional to the magnitude of the pre-market TSI. Every daily position is closed at the end of the day. Table 5 shows that the trading strategy provides an annualized Sharpe ratio of 1.37 over the sample period. The table also reports the Sharpe ratio of a simple long only strategy over the same period. The returns of the long only strategy are the daily open-to-close returns of the S&P500 index ETF.[7] The null hypothesis that the Sharpe ratios are equal can not be rejected but the Sharpe ratio of the TSI trading strategy is above that of the long only strategy in 73 percent of the random draws from the sample.[8] Comparing the Sharpe ratio of the trading

[7]The open price of the SPY ETF takes into account overnight gaps in the index. Stock market index could open significantly higher or lower than its prior close as a result of overnight news or economic announcements prior to market open.

[8]The probability is estimated using the bootstrap method. Random pair observations of the daily returns associated with the trade strategies are drawn and the corresponding Sharpe ratio is calculated.

strategies highlights the incremental benefit of using the pre-market TSI for trading the stock market index.

3.6 TSI and Volatility

Investor optimism is negatively related to the demand for protection against future volatility. In this subsection, the daily returns of volatility futures is regressed on the pre-market TSI and control variables. Table 6 shows that the follower-weighted sentiment of tweets posted before the market open predict the daily returns of the volatility futures ETN (VXX).[9] One standard deviation increase in the pre-market TSI predicts a loss of 25 basis point for the volatility futures ETN.

3.7 TSI and other measures of sentiment

TSI measures the sentiment of individuals who post their directional forecast of the S&P500 index on Twitter. In order to validate the TSI, as a measure of investor sentiment, it is useful to examine the correlation between the TSI and other measures of sentiment. If there is no correlation between the TSI and other measures, we can conclude that the TSI measures something unrelated to investor sentiment. High correlation, on the other hand, would suggest that individuals simply post an opinion based on other observable variables and the TSI provides no incremental information about expectations.

Table 7 shows the correlation between the TSI and alternative measures of sentiment. The opinions posted by accounts with more than 500 followers are used to construct an equal-weighted daily sentiment index and the correlations are provided in the first column. The second column shows the correlations when the sentiment index is constructed from the opinions of accounts with less than 500 followers. The alternative measures of sentiment are classified in three groups.

The first group includes measures that indicate changes in net position of speculators reported by the Commodity Futures Trading Commission (CFTC). Every week, CFTC reports open interest and changes in long and short positions in the futures and options market by aggregating

[9] VXX is the ticker symbol of an Exchange Traded Note (ETN) that tracks the price of 30 days forward volatility in the S&P500 index.

regulatory transaction data. Market participants often monitor the weekly changes to learn about expectations of others. The first two rows in Table 7 show the correlation between the TSI and changes in net long position of institutional asset managers and leverage funds respectively.[10] The TSI shows positive correlation with weekly changes in net long position of asset managers that include pension funds, insurance companies, mutual funds and investment managers whose clients are predominantly institutional. Conversely, there is a negative correlation between the TSI and changes in net long position of leverage funds that are typically hedge funds.

The second group includes a survey-based measure of sentiment. American Association of Individual Investors (AAII) conducts a sentiment survey and measures the percentage of individual investors who are bullish, bearish, and neutral on the stock market for the next six months. Every individual who subscribes to the AAII services can submit a vote in the weekly surveys. The proportion of investors in each group is measured based on the data received each week by Wednesday and the results are reported on Thursdays. The survey was first conducted in July 1987[11]. Positive and statistically significant correlation between the TSI and the bull-bear spread of the weekly AAII survey shows that they both pick up investor optimism using data from different sources.[12]

The third group includes a number of market based variables that are often used as measures of investor optimism. Market based measures are usually available in daily frequency but they could be influenced by other factors unrelated to sentiment (see Baker and Wurgler (2007)). The daily TSI is positively related to measures of market breath such as advance to decline ratio and the ratio of the number of stocks at their 52 weeks high price to those at their 52 weeks low price.[13] There is also a negative correlation between the TSI and measures of fear such as CBOE volatility index (VIX) and the put-call ratio of short-term options on the S&P500 index.

The correlations reported in Table 7 indicate that the TSI is consistent with what would be expected from a measure of sentiment. Modest correlations show that there is a common

[10] Since the commitment of traders data is available in weekly frequency, a weekly measure of TSI is created to compute its correlation with weekly changes in net long positions. The weekly TSI is a simple average of daily TSI over a week and is computed for 188 weeks in the sample.

[11] Source:www.aaii.com/sentimentsurvey

[12] There are other survey-based measures of sentiment but since they are reported in monthly frequency and the Twitter sample has only 45 months of data, I only consider AAII survey in the survey-based measures.

[13] The daily TSI is the average of the numerical score of opinionated tweets posted on a trading day.

component in the TSI and three types of sentiment proxies. Moreover, positive and statistically significant correlation between the TSI and changes in the ADS business conditions index suggests that economic news is part of the information set that forms expectations of investors about future returns. The second column of Table 7 shows that the correlations are generally weaker when the Twitter sentiment is constructed from the opinions of less popular accounts. This is another evidence that shows the importance of giving more weight to opinions posted by accounts with large number of followers.

4 Disagreement

Literature has shown that disagreement among investors is related to return volatility.[14] Moreover, disagreement is linked to trading volume as there is more potential for trading when disagreement rises. In this section, the opinionated tweets are used to construct a daily measure of disagreement and the relation between disagreement, volatility and abnormal trading volume is examined.

4.1 Dispersion of Sentiment Index (DSI)

In order to construct a daily measure of disagreement, I follow Antweiler and Frank (2004) and consider all the bullish and bearish tweets posted during a day. The daily Dispersion of Sentiment Index, denoted by DSI_t, is defined as the standard deviation of the numerical score of positive and negative tweets. Since the opinions are coded as $+1$ and -1, the standard deviation of the opinions can be expressed as a function of the average sentiment:

$$DSI_t = \sqrt{1 - TSI_t^2}.$$

The DSI could take any value between 0 and 1. Dispersion of sentiment will be zero if all the tweets posted on a trading day express the same view about the direction of the stock market index. The other extreme case is when the views expressed by tweets are equally split between positive and negative categories. In this case, dispersion of opinions will be 1. Disagreement is

[14] Basak (2005) and Buraschi and Jiltsov (2006)

meaningless if there is less than two opinionated tweets so the DSI is not calculated for 12 trading days with less than two opinionated tweets.

4.2 Disagreement and Volatility

The literature offers many ways to model volatility. Hansen and Lunde (2005) compare 330 ARCH models and find no evidence that sophisticated models can predict the conditional variance better than a simple GARCH(1,1) model. In this study, a GARCH(1,1) with additional independent variables is used to model the stock market volatility. More specifically, the volatility is given by:

$$\begin{aligned} r_t &= \sqrt{h_t}\epsilon_t \\ h_t &= \alpha_0 + \alpha_1 r_{t-1}^2 + \alpha_2 h_{t-1} + \big(\beta_1 DSI_t + \beta_2 logDOT_t\big) + \gamma Z_t, \end{aligned}$$

where r_t is the daily return of the S&P500 index ETF and h_t represents the return volatility. DSI_t is the daily measure of disagreement defined in section 4.1 and $logDOT_t$ is log Daily number of Opinionated Tweets that controls for the density of information arrival. Z_t is the vector of control variable that includes the changes in ADS business conditions index and the Economic Policy Uncertainty (EPU) index.

Table 8 shows that an increase in disagreement among investors predicts higher daily volatility. This finding is consistent with dynamic models of heterogeneous beliefs, such as Buraschi and Jiltsov (2006), that predict a positive correlation between dispersion of opinions and return volatility. The table also shows that improvement in business conditions measured by the ADS index lead to less volatility and high economic policy uncertainty measured by the EPU index is associated with more volatility.

The daily log number of opinionated tweets is used as a proxy for density of information arrival and is included to control for the effect of information arrival rate on the volatility.[15] As mentioned earlier in the subsection 2.1, the dataset includes the tweets that provide a forward looking opinion about the direction of stock market. Individuals usually express their opinion in a tweet by directly pointing out their outlook or announcing their current position. Opinionated tweets are typically

[15] See Andersen (1996)

posted when an individual acquires new information. For instance, a trader sends a bullish tweet following the public announcement of some economic data or based on a conclusion after analyzing technical indicators. Therefore, the number of opinionated tweets is considered a proxy for public and private information arrival rate. Positive and statistically significant coefficient of $logDOT_t$ is consistent with the model of Andersen (1996) that assumes a positive relation between the return volatility and information flow rate.

4.3 Disagreement and Abnormal Volume

Disagreement is often considered a motivating factor for trading in the financial markets. To examine the link between disagreement among investors and trading volume, I follow Cookson and Niessner (2016) and estimate the coefficients of the following model:

$$abVol_t = \alpha + \beta_1 DSI_t + \beta_2 abVol_{t-1} + \gamma Z_t,$$

where $abVol_t$ is the abnormal log trading volume of SPY ETF on day t and Z_t is the vector of control variables that includes weekday, month and year dummies to control for seasonal factors in abnormal trading volume. The abnormal trading volume on a given day t is the difference between the trading volume and its average over a six months period ending on $t - 20$. One day lagged abnormal trading volume is included to control for persistence of trading volume.

Table 9 shows the coefficient of disagreement estimated in the above regression. The first column indicates that disagreement is positively related to abnormal trading volume. Disagreement remains significant as additional dummy variables are included in the second column to control for potentially more than usual trading volume on FOMC decision days and employment data release days. In column (3), a dummy variable is included for option expiry days to account for substantially higher than usual trading volume on the days that options expire every month. The last column shows that the sign and significance of disagreement remains unchanged after including the option expiry days dummy.

5 Conclusion

Every day, large numbers of people go online and share their opinion about different topics including asset prices on Twitter. In this paper, the directional forecasts about the stock market are extracted from Twitter messages and a daily measure of investor sentiment is constructed by aggregating the opinionated tweets. In order to account for the heterogeneity in the quality of private signals, opinionated messages are weighted by the number of followers of the accounts. It is shown that the follower-weighted index predicts the daily returns of the S&P500 index and a trading strategy that uses the TSI as a trading signal provides an annualized Sharpe ratio of 1.37. Consistent with the sentiment theories, the effect is reversed on future days.

The TSI shows positive correlation with survey-based and market-based measures of sentiment and negative correlation with the fear gauges, such as VIX index and put-call ratio of short-term options on the S&P500 index. There are two important differences between the TSI and survey-based measures of sentiment. First, survey-based measures assign equal weight to the opinions but the TSI provides the possibility of giving different weight to the opinions based on their followers. Second, individuals have the incentive to express their true opinion on Twitter to maintain their reputation but there is no such incentive for participants in the sentiment surveys because the individual opinions are not disclosed. Moreover, most survey-based measures are reported in weekly or monthly frequency but the TSI could be constructed and reported in daily frequency from a large number of opinionated tweets posted every day.

In the paper, the opinionated tweets are also used to construct a daily measure of disagreement. Consistent with several theoretical models and empirical studies in the finance literature, the volatility of daily returns is positively related to disagreement among investors. It is also shown that the disagreement is associated with abnormal trading volume.

References

[1] Andersen, Torben G. (1996), Return volatility and trading volume: An information flow interpretation of stochastic volatility, *The Journal of Finance*, 51(1), 169-204.

[2] Antweiler, Werner, and Murray Z. Frank (2004), Is all that talk just noise? The information content of internet stock message boards, *The Journal of Finance*, 59(3), 1259-1294.

[3] Aruoba, S. Boraan, Francis X. Diebold, and Chiara Scotti (2009), Real-time measurement of business conditions, *Journal of Business & Economic Statistics*, 27:41727.

[4] Atmaz, Adam, and Suleyman Basak (2016), Belief dispersion in the stock market, *The Journal of Finance*, Forthcoming.

[5] Baker, Scott R., Nicholas Bloom, and Steven J. Davis (2013), Measuring economic policy uncertainty, Chicago Booth Research Paper No. 13-02.

[6] Baker, Malcolm, and Jeffrey Wurgler (2007), Investor sentiment in the stock market, *Journal of Economic Perspectives*, 21(2), 129-151.

[7] Banerjee, Snehal, and Ilan Kremer (2010), Disagreement and learning: Dynamic patterns of trade, *The Journal of Finance*, 65(4), 1269-1302.

[8] Basak, Suleyman (2005), Asset pricing with heterogeneous beliefs, *Journal of Banking and Finance*, 29(11), 2849-2881.

[9] Bollen, Johan, Huina Mao and Xuaojun Zeng (2011), Twitter mood predicts stock market *Journal of Computational Science*, 2, 1-8.

[10] Brown, Gregory W., and Michael T. Cliff (2004), Investor sentiment and the near-term stock market, *Journal of Empirical Finance*, 11(1), 1-27.

[11] Buraschi, Andrea, and Alexei Jiltsov (2006), Model uncertainty and option markets with heterogeneous beliefs, *The Journal of Finance*, 61(6), 2841-2897.

[12] Carlin, Bruce I., Francis A. Longstaff, and Kyle Matoba (2014), Disagreement and asset prices, *Journal of Financial Economics*, 114(2), 226-238.

[13] Chen, Hailiang, Prabuddha De, Yu (Jeffrey) Hu, Byoung-Hyoun Hwang (2014), Wisdom of crowds: The value of stock opinions transmitted through social media, *The Review of Financial Studies*, 27(5), 13671403.

[14] Cookson, J. Anthony, and Marina Niessner (2016), Why don't we agree? Evidence from a social network of investors, *Working paper, SSRN 2754086.*

[15] Da, Zhi, Joseph Engelberg, and Pengjie Gao (2015), The sum of all FEARS: Investor sentiment and asset prices, *Review of Financial Studies*, 28(1): 1-32

[16] Das, Sanjiv R., and Mike Y. Chen (2007), Yahoo! for Amazon: Sentiment extraction from small talk on the web, *Management Science*, 53(9), 1375-1388.

[17] De Long, J. Bradford, Andrei Shleifer, Lawrence H. Summers, and Robert J. Waldmann (1990), Noise trader risk in financial markets, *Journal of Political Economy*, 98(4), 703-38.

[18] Gallmeyer) Gallmeyer, Michael, and Burton Hollifield (2008), An examination of heterogeneous beliefs with a short-sale constraint in a dynamic economy, *Review of Finance*, 12(2), 323-364.

[19] Gultekin, Mustafa N., and N. Bulent Gultekin (1983), Stock market seasonality: International evidence, *Journal of Financial Economics*, 12(4), 469-481.

[20] Hansen, Peter R., and Asger Lunde (2005), A forecast comparison of volatility models: does anything beat a GARCH (1,1)? *Journal of applied econometrics*, 20(7), 873-889.

[21] Karpoff, J. M. (1986), A theory of trading volume, *The Journal of Finance*, 41(5), 1069-1087.

[22] Keim, Donald B. (1983), Size-related anomalies and stock return seasonality: Further empirical evidence, *Journal of Financial Economics*, 12, 13-32.

[23] Li, Feng (2008), Annual report readability, current earnings, and earnings persistence, *Journal of Accounting and Economics*, 45, 221247.

[24] Mao, Huina, Scott Counts and Johan Bollen (2015), Quantifying the effects of online bullishness on international financial markets, ECB Statistics Paper 9, July 2015.

[25] Newey, Whitney K., and Kenneth D. West (1987), A simple, positive semi-definite, heteroskedasticity and autocorrelation consistent covariance matrix, *Econometrica*, 55, 703708.

[26] Tetlock, Paul C. (2007), Giving content to investor sentiment: the role of media in the stock market, *The Journal of Finance*, 62(3), 1139-1168.

[27] Zhang, Xue, Hauke Fuehres and Peter A. Gloor (2011), Predicting stock market indicators through Twitter "I Hope it is not as bad as I fear", Procedia-Social and Behavioral Sciences, 26, 55-62.

Figure 1: Distribution of the Daily Number of Opinionated Tweets.

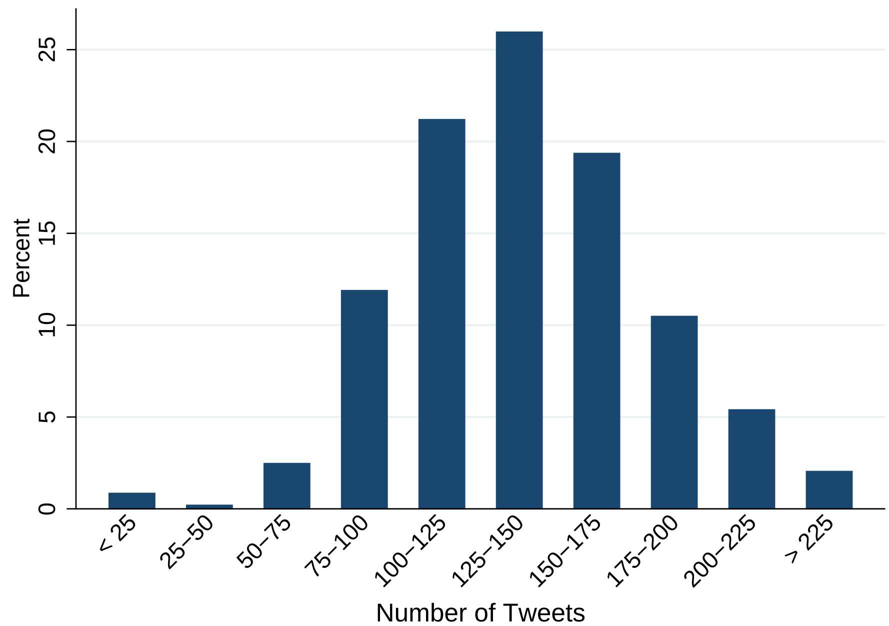

Figure 2: Distribution of the Number of Followers

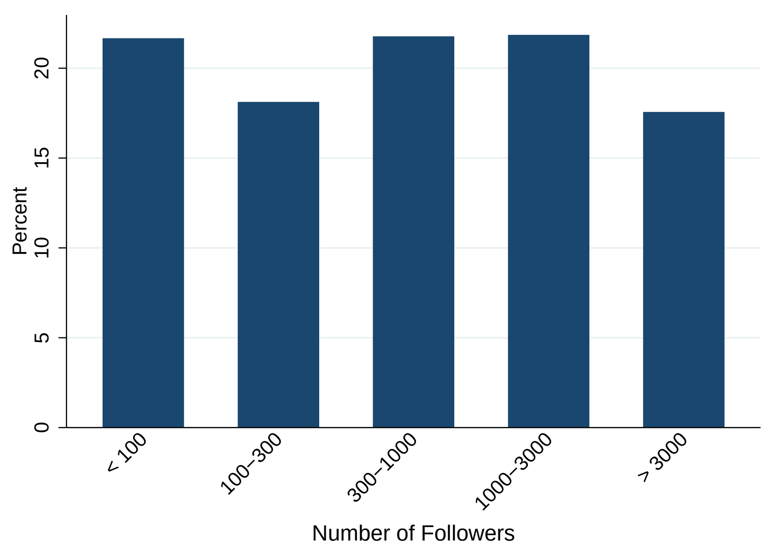

Figure 3: Average Daily Weight of Tweets in the Sentiment Index Over the Number of Followers.

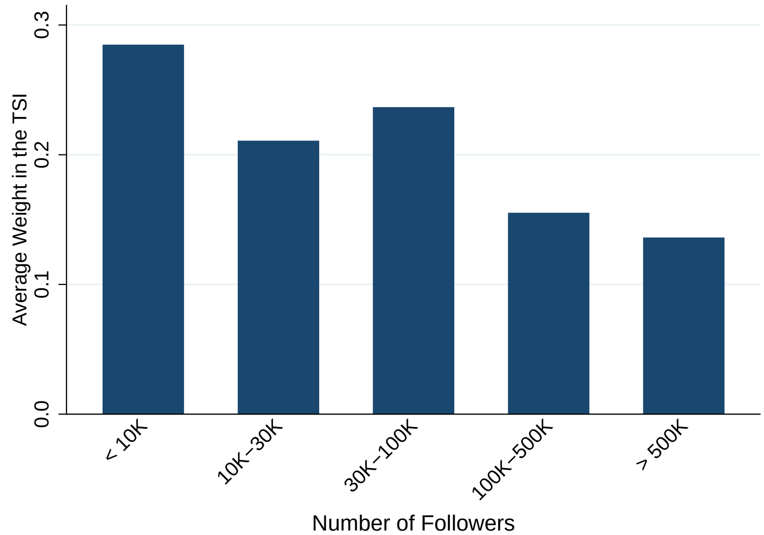

Figure 4: Twitter Sentiment Index (TSI) and S&P500 Index

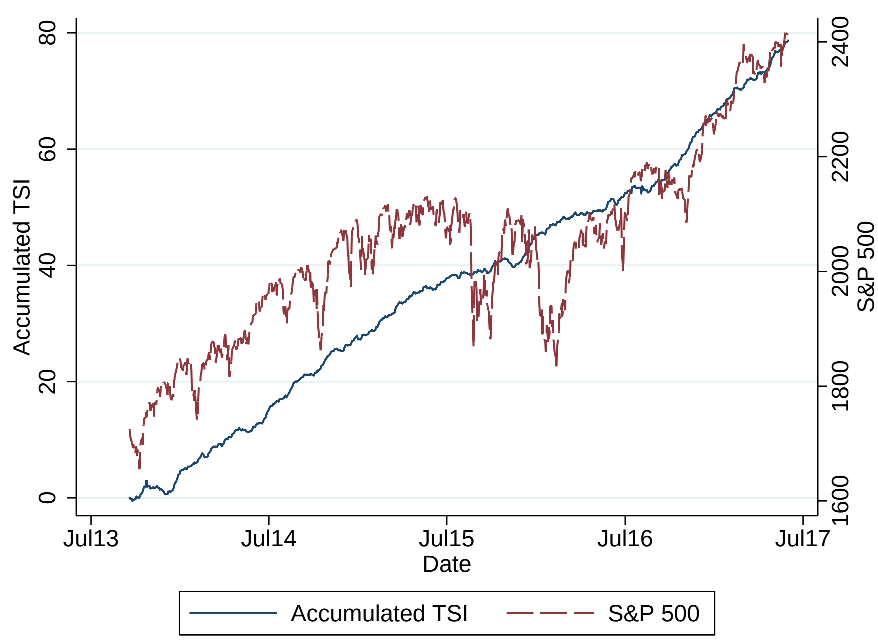

Table 1: Examples of Positive, Neutral and Negative Tweets

Score	Category	Text
+1	Positive	I am long $SPXL $TNA $TQQQ coming into today.
+1	Positive	Bought $SPY 227 calls .59
+1	Positive	Trade: SELL -1 $SPY PUTS: JUL15 200 0.91.
+1	Positive	added to some longs $SPX
+1	Positive	im out of my $spy short positions
+1	Positive	A Tight Range with a Slightly Bullish Bias Next Week (Says the OI) $SPX $SPY
+1	Positive	US stocks turn lower as crude tumbles; but could we see a rebound? $SPX chart bullish
0	Neutral	$SPX Waiting on housing number at top of the hour.
0	Neutral	$SPY watch 169.84 here
0	Neutral	$SPY also below 10sma - be patient - look for relative strength - sit on your hands if needed
0	Neutral	Staying flat for now. Good day $SPX $SPY
-1	Negative	Slightly bearish setup for the day on $SPY.
-1	Negative	Pressing my shorts with an objective to move 20% net short this morning on any future strength. $SPY
-1	Negative	Start the day 10% net short. $SPY
-1	Negative	Weekly S&P 500 ChartStorm - Bearish Beacons Building $SPY
-1	Negative	Bonds and stocks are now overbought.$TLT $SPY
-1	Negative	Still looking for more downside. $SPX

Table 2: Regression results of daily open-to-close stock market returns on the Pre-market sentiment, control variables, and lagged return of stock market index.

SPY is the daily open-to-close return of the ETF that tracks daily returns of the S&P500 index, $PreTSI$ is the pre-market Twitter sentiment index that is constructed using the tweets posted before the market open, VIX is the CBOE volatility index, ΔEPU is changes in the economic policy uncertainty index, and ΔADS is changes in the business conditions index from the Federal Reserve Bank of Philadelphia

	Follower Weighted	**Equal Weighted**
	SPY_t	SPY_t
$PreTSI_t$	0.121**	-0.030
	(0.048)	(0.089)
ΔADS_t	2.822	2.794
	(2.258)	(2.268)
ΔEPU_t	-0.014	-0.017
	(0.044)	(0.044)
VIX_{t-1}	0.004	0.004
	(0.006)	(0.006)
Observations	932	932
Adjusted R-squared	0.014	0.007

Standard errors are in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Table 3: Robustness Checks: Regression results of daily open-to-close stock market returns on the Pre-market sentiment, control variables of Table 2 and some additional control variables.

(1) One day lagged market indicators such as the number of advancing stocks divided by the number of declining stocks in NYSE, ARMS index, the number of stocks at their 52 weeks high price divided by the number of stocks at their 52 weeks low price, the number of put options on the S&P500 index divided by the number of call options, and the difference between the interest rate of 10 years and 3 months U.S. Treasury bonds are included as control variables.
(2) Overnight price changes of the stock market index is included to control for the effect of overnight news on sentiment and stock prices in futures market.
(3) Month dummy is included to control for seasonal effects.
(4) FOMC dummy is included to control for the potentially significant reaction of the stock market to the release of FOMC statements. It takes the value of 1 on the dates of FOMC statement release and 0 otherwise.
(5) Macro news dummy is included to control for the effect of GDP, unemployment, and inflation announcements on the tweets and stock market returns.

	(1) SPY	(2) SPY	(3) SPY	(4) SPY	(5) SPY
PreTSI	0.1189**	0.1268***	0.1383***	0.1216**	0.1247**
	(0.0478)	(0.0481)	(0.0479)	(0.0477)	(0.0477)
Control Variable(s)	Market indicators	Overnight price changes	Month dummy	FOMC dummy	Macro news dummy
Observations	932	932	932	932	932
Adjusted R-squared	0.020	0.015	0.025	0.014	0.017

*** p<0.01, ** p<0.05, * p<0.1

Table 4: Regression results of the S&P500 daily return on the pre-market TSI, control variables, and five days lagged returns of the S&P500 index.

	Ret_t	Ret_{t+1}	Ret_{t+2}	Ret_{t+3}	Ret_{t+4}	Ret_{t+5}
$PreTSI_t^{fw}$	0.115**	0.069	-0.123**	0.050	-0.035	-0.055
	(0.047)	(0.062)	(0.062)	(0.062)	(0.062)	(0.062)
VIX_t	-0.032***	0.016**	0.024***	0.022***	0.024***	0.016**
	(0.006)	(0.008)	(0.008)	(0.008)	(0.008)	(0.008)
ΔADS	2.687	1.887	2.187	2.062	2.484	2.787
	(2.226)	(2.901)	(2.889)	(2.890)	(2.894)	(2.902)
ΔEPU	-0.004	-0.020	0.022	0.094*	-0.029	-0.044
	(0.043)	(0.056)	(0.056)	(0.056)	(0.056)	(0.056)
Ret_t		-0.001	0.001	0.003	-0.036	0.018
		(0.035)	(0.034)	(0.034)	(0.034)	(0.035)
Observations	932	932	932	932	932	932
R-squared	0.045	0.013	0.021	0.018	0.015	0.010

Standard errors are in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Table 5: Sharpe ratio of the trading strategy that uses the TSI as a trading signal.

"TSI Trading" strategy takes daily long or short positions on the SPDR S&P500 index ETF (SPY) at market open depending on the sign and magnitude of the Twitter sentiment before the market open. Positive (Negative) Twitter sentiment provides a signal for long (short) position and the size of the position is proportional to the magnitude of the TSI. The daily positions are closed at the market close on every trading day. "Passive Trading" is a trading strategy that takes a long position on the ETF at market open and closes the position at market close on the same day.

	Daily	Annualized
TSI Trading	0.087	1.374
	(0.031)	(0.495)
Passive Trading	0.059	0.929
	(0.033)	(0.528)

Standard errors are in parentheses

Table 6: Regression results of VIX futures contracts' daily return on the Pre-market Twitter Sentiment Index (PreTSI), and control variables.

VXX is the daily return of an ETN that provides exposure to a daily rolling long position in 30 days VIX futures contracts. Control variables are changes in the economic policy uncertainty index, changes in the ADS business conditions index from the Federal Reserve Bank of Philadelphia, one day lagged Volatility index and lagged return of VXX.

	VXX_t
$PreTSI_t^{fw}$	-0.594**
	(0.235)
ΔADS_t	-10.941
	(11.174)
ΔEPU_t	0.289
	(0.217)
VIX_{t-1}	0.022
	(0.028)
Observations	932
Adjusted R-squared	0.018

Standard errors are in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Table 7: Correlation between the TSI and alternative measures of sentiment.

"TSI $>$ 500 followers" is the daily Twitter Sentiment Index constructed from the opinionated tweets posted by accounts with more than 500 followers.
"TSI $<$ 500 followers" is the sentiment of those with less than 500 followers.
"Δ Asset Managers Net Long" and "Δ Leverage Money Net Long" are weekly changes in net long positions of asset managers and hedge funds on the S&P500 futures and options contracts respectively.
"AAII bull-bear spread" is the difference between the percentage of bullish and bearish investors in the weekly AAII surveys.
logVIX is the log CBOE volatility index.
ADVDEC is the number of advancing stocks divided by the number of declining stocks in NYSE.
Put/Call Ratio is the number of put options on the S&P500 index divided by the number of call options.
HILO is the number of stocks at their 52 weeks high price divided by the number of stocks at their 52 weeks low price.
ΔADS is changes in the business conditions index from the Federal Reserve Bank of Philadelphia.

	TSI $>$ 500 followers	TSI $<$ 500 followers
CFTC Commitment of Traders		
Δ Asset Managers Net Long	0.279***	0.240***
Δ Leverage Money Net Long	-0.148**	-0.106
Survey-based Sentiment Index		
AAII bull-bear spread	0.154**	-0.021
Financial Market-based Measures		
logVIX	-0.075**	-0.052
ADVDEC	0.273***	0.212***
Put/Call Ratio	-0.112**	-0.075**
HILO	0.014	0.082**
Business Conditions Index		
ΔADS	0.071**	0.008

*** p<0.01, ** p<0.05, * p<0.1

Table 8: GARCH(1,1) forecast for volatility of the SPY daily returns (denoted by h_t).

The table shows the daily volatility as a function of daily disagreement measured by Dispersion of Sentiment Index (DSI). DSI_t is the standard deviation of bullish and bearish tweets throughout a day. $logDOT$ is the log Daily number of Opinionated Tweets.

	h_t
DSI_t	7.601***
	(2.600)
$logDOT_t$	3.497***
	(0.302)
ΔADS_t	-18.79**
	(7.803)
ΔEPU_t	0.448**
	(0.202)
Observations	920

Table 9: Regression of abnormal log Trading Volume on Disagreement.

The dependent variable is abnormal log trading volume of SPY ETF. In column (1), the control variables are one day lagged abnormal volume, week day dummy, month dummy and year dummy. FOMC decision day dummy and unemployment data release day dummy are included as additional controls in column (2) and option expiry day dummy is included in column (3).

	(1)	(2)	(3)
DSI_t	0.490***	0.498***	0.511***
	(0.125)	(0.125)	(0.124)
Observations	919	919	919
R-squared	0.463	0.467	0.474