Why Don't We Agree? Evidence from a Social Network of Investors*

J. Anthony Cookson† and Marina Niessner‡

June 20, 2016

Abstract

We study the sources of investor disagreement using sentiment expressed by investors on a social media investing platform, combined with information on the users' investment approaches (e.g., technical, fundamental). We examine how much of overall disagreement is driven by different information sets versus different investment models, by studying disagreement within and across investment approaches. We find that differences of opinion across investment approaches account for 54% of the overall disagreement at the firm-day level. Moreover, changes in our measures of disagreement robustly forecast abnormal trading volume, suggesting that our measures proxy well for disagreement in the wider market. Our findings suggest that improvements to informational efficiency of financial markets won't completely erode high trading volume and stock market volatility.

*We are grateful to James Choi, Diego Garcia, Rawley Heimer, Tyler Muir, Justin Murfin, Paul Tetlock (discussant), and Heather Tookes for helpful comments. We thank Jason Klusowski and Toomas Laarits for outstanding research assistance. This draft has also benefited from the comments of conference and seminar participants at the 2015 European Summer Symposium for Financial Markets (evening session), the 2015 IDC Summer Conference (early ideas), Universidad de Chile, University of Colorado Consumer Financial Decision Making Group, Yale School of Management, and 2016 NBER Spring Behavioral Finance Meeting.

†University of Colorado at Boulder - Leeds School of Business, tony.cookson@colorado.edu

‡Yale School of Management, marina.niessner@yale.edu

Introduction

Disagreement among investors is central to trading in financial markets. Indeed, it is difficult to motivate why investors would trade at all without some source of disagreement (Milgrom and Stokey, 1982; Karpoff, 1986). Motivated partly by this observation, a growing literature evaluates the effects of investor disagreement on financial market outcomes (e.g., Varian, 1985; Nagel, 2005; Banerjee and Kremer, 2010; Carlin et al., 2014; Bailey et al., 2016). Research has linked disagreement to trading volume and stock returns, and has studied its dynamic effects (Ajinkya et al., 1991; Diether et al., 2002; Banerjee and Kremer, 2010).

Despite the breadth of work on the consequences of investor disagreement, much less is known empirically about the *sources* of disagreement. That is, why do investors disagree in the first place? Leading theories have identified differences in information sets and differences in models that investors use to interpret information as the two main sources of disagreement (Hong and Stein, 2007). In this paper, we study a setting in which models (i.e., investment approaches) are explicitly reported by investors, to examine the role of different investment philosophies in overall investor disagreement. We find there is significant scope for disagreement across investment approaches.

Separating the roles of information asymmetry and heterogeneous models in investor disagreement is empirically challenging, given the typical data limitations. First, disagreement refers to differences in investors' opinions, which are difficult to observe. Even if a researcher had individual-level trading data (which itself is hard to come by), it is difficult to impute investors' opinions from their trades, as investors can trade for reasons unrelated to their opinions - like liquidity. Second, as Rothschild and Sethi (2014) and Baron et al. (2012) point out, in order to separate whether the differences in investor opinions are due to differences in information sets or differences in investors' models, the researchers would ideally observe investors trading strategies – not just the executed trades. We overcome these challenges by studying disagreement among investors on a social media investing platform (called StockTwits), in which users regularly express their opinions (bullish or bearish) about the same stocks, and where user profile information *explicitly* conveys the user's broad investment approach.

Our data set enables us to empirically draw the distinction between information-driven versus model-driven sources of disagreement, because disagreement across investment approaches is more likely to arise due to differing investment models, whereas within-investment approach disagreement is more likely to be related to different information sets. When we evaluate disagreement across groups, we find that differences in opinions across the investment approaches in our data are responsible for approximately half of overall disagreement. In addition, we find that both cross-group disagreement and within-group disagreement are robust and significant predictors of abnormal trading volume. Thus, both information-driven and model-driven differences in opinions are important sources of disagreement.

It is important to note that our investment approach measures are coarsely defined. For example, whereas we only know whether a user is a technical investor, there are several different subgroups of technical investors (e.g., "subjective" and "objective" technical approaches), which we do not observe. The same holds for other investment approach categories in our dataset. As a result, some of the disagreement that we attribute to within-groups, and therefore to differences in information sets, is actually caused by model disagreement. Thus we likely document a lower bound of the contribution of cross-group disagreement to overall disagreement in the market.

One concern with our data is that trading by investors that communicate their opinions on StockTwits probably does not account for the majority of trading volume in the market. Therefore, we examine whether our disagreement measures (cross-group and within-group) predict trading volume. We find that not only is there a strong contemporaneous relationship between our disagreement measures and trading volume in the market, but that out disagreement measures forecast *future* trading volume. For example, a one standard deviation increase in overnight cross-group disagreement is associated with approximately 4% greater abnormal trading volume during the day, controlling for abnormal trading volume during the previous day. This suggests that our disagreement measures proxy for disagreement in the wider market, even though the trades by the investors on StockTwits do not likely represent the majority of trades in the market.

Furthermore, we examine how our disagreement measures relate to disagreement proxies used in prior literature. We find a relatively weak correlation with previous proxies for disagreement, indicating that our disagreement measures capture a distinct notion of disagreement from prior work. One strength of our measure is that it directly captures dispersion of investor opinions, whereas leading alternative disagreement measures rely on indirect information, either observed trading patterns (i.e., volatility measures) or opinions of third parties (i.e., analyst forecast dispersion). Another advantage is that our measure can be reliably computed at the daily level, whereas alternative measures need to be measured at lower frequencies (typically, monthly or quarterly). Given that the puzzle in the literature is to explain the high trading volume at the daily level (e.g., Hong and Stein, 2007), this is an important distinction.

As our disagreement measure is based on messages posted by investors, and the number of messages in a given day is strongly correlated with investors' attention, it is important to ensure that our measure of disagreement among investors is distinct from investor attention. We analyze the joint effect of investor disagreement and investor attention on trading volume, where we proxy for investor attention by the number of daily messages on StockTwits and also by the number of daily searches for the companies' tickers on Google (e.g., Da et al., 2011; Niessner, 2016). Consistent with prior theory (Hong and Stein, 2007), we find that both investor disagreement and investor attention are strongly associated with greater trading volume, but trading volume is highest when both measures are high. This is a sensible implication, suggesting that trading volume is driven not only by the overall number of investors who are paying attention, but also by the dispersion of their opinions.

Finally, we examine time periods where most investors in the market observe the same information, yet prior literature has documented very large increases in trading volume: the time period around earnings announcements. We verify the results from prior literature that trading volume goes up by more than 20% on earnings announcement days and stays high for several weeks afterwards, and find that variation in our disagreement measure can explain up to one third of this spike in trading volume. This finding provides additional support for our earlier results that model-driven disagreement is an important source of overall disagreement in the market, and also provides fresh empirical evidence for emerging theories that argue for why disagreement rises precisely when information arrives to the market (Kondor, 2012; Banerjee et al., 2015).

Our results, measure of disagreement, and approach should be of broad interest to scholars studying individual investing behavior and market microstructure, as well as policy makers more generally. First, although there has been significant inquiry into the consequences of disagreement for financial market outcomes, we are the first to empirically study the sources of disagreement. In so doing, we provide empirical evidence of both channels posited theoretically in Hong and Stein (2007). This is an important step forward because showing that a substantial component of disagreement arises from differing investment approaches implies that enriching the information environment will not fully alleviate disagreement in financial markets, and in fact, as recent theoretical contributions have highlighted, disagreement may rise (Kondor, 2012; Banerjee et al., 2015).

We also contribute to the disagreement literature by providing a useful measure of disagreement among individual investors. Although the consequences of disagreement are well studied, the extant measures of disagreement have notable weaknesses. For example, some of these measures measure dispersion of opinion indirectly (e.g., volatility of accounting performance, historical trading volume, firm age, return volatility), and the most prominent measure of analyst forecast dispersion measures the stated opinions of analysts, which has been questioned as a reliable measure of market-wide disagreement (Ataise and Bamber, 1994; Bamber et al., 2011). We fill this gap by combining our setting – which yields daily measures of sentiment at the individual firm $\times$ investment approach level – with a theoretically grounded measure of disagreement from Antweiler and Frank (2004). Taken together, our disagreement measure can be computed at a higher frequency than most other measures of disagreement (analyst dispersion is usually computed monthly or quarterly), and because it is a direct sentiment measure, it is less likely to proxy for other market forces that are unrelated to disagreement, such as liquidity needs of investors.

Our results on abnormal trading volume and disagreement also relate to the literature on the abnormal trading of individual investors (Barber and Odean, 2000). In particular, this literature has identified numerous behavioral rationales for over-trading, including entertainment (Dorn and Sengmueller, 2009), sensation seeking (Barber and Odean, 2008; Grinblatt and Keloharju, 2009), gambling (Kumar, 2009; Cookson, 2016), and learning by doing (Linnainmaa, 2011). We contribute to this stream of research by showing clean evidence that model disagreement is an additional reason for abnormal trading volume. It is notable that model disagreement is not well aligned with entertainment motives, nor learning by doing motives for trading, and thus, is a theoretically distinct rationale for additional trading.

Our research complements recent work on the micro-level determinants and consequences of investor disagreement (e.g., Carlin et al., 2014; Giannini et al., 2015) by delving deeper into the underlying sources of disagreement. In this respect, a closely related paper to ours is Chang et al. (2014), who also study sources of disagreement among individual investors. Chang et al. (2014) find that linguistic diversity is a source of divergence of opinion because agreement is more difficult when there are communication barriers. By contrast, we show that differing investment approaches are an important source of disagreement, irrespective of language, which is homogeneous on StockTwits.

In the end, understanding the causes of investor disagreement has important policy implications. Regulators put substantial effort into trying to minimize information asymmetry among investors (e.g., see the analysis in Rogers et al., 2015). Abstracting from any notion of fairness, it is important to understand whether and by how much these policies could actually decrease disagreement among investors, and therefore trading volume and volatility in the stock market. For these reasons, it is natural that Hong and Stein (2007) pose the key question, "what are underlying mechanisms, either at the level of market structure or individual cognition, that give rise to disagreement among traders and hence to trading volume?" Our results suggest that different investment philosophies are partly responsible for the high trading volume because two investors reading the same piece of information likely draw different conclusions about the report's implications for a proper trading position.[1] Therefore, new information might not decrease volatility, but in fact, volatility may increase.

1 Data

1.1 StockTwits Data

Our data set comes from a company called StockTwits. StockTwits was founded in 2008 as a social networking platform for investors to share their opinions about stocks. The website has Twitter-like format, where participants post messages of up to 140 characters, and use "cashtags" with the stock ticker symbol (example $AAPL), to index ideas to a particular company. Although the website does not automatically integrate with other social media websites, users can share content to their personal Twitter, LinkedIn, and Facebook accounts. According to a website analytics tool, Alexa, StockTwits was ranked as the 2,004th most popular website in the US as of May, 2015. The users are predominantly male and the number of users with a graduate school degree is over-represented relative to other websites on the internet that Alexa tracks.

[1] A recent article by the Economist mentioned "This week a report showing a slump in China's imports and exports in November was read differently by bulls and bears" (The Economist, "In a hole", December 12, 2015).

Our original dataset spans from 1 January, 2010 until 30 September, 2014. In total, there are 18,361,214 messages by 107,920 unique users mentioning 9,755 unique tickers. For each message, we observe a user identifier and the message content. We also observe indicators for sentiment (bullish, bearish, or unclassified), and "cashtags" with tickers that link the message to particular stocks.

For most users, we observe a self-reported investment approach (technical, fundamental, momentum, value, growth, or global macro), user's holding period (day trader, swing trader, position trader, or long term investor), and their experience level (novice, intermediate, or professional).

We restrict our sample to cover the time period between January 2013 and September 2014 because the number of messages posted to StockTwits has grown substantially over time and the best quality data come from more recent years. As can be seen in Table 1, this restriction leaves us with 75% of messages. To focus on sentiment that can be directly linked to a particular stock, we restrict attention to messages that only mention one ticker. We also focus on the sample of messages by users who have indicated their investment approach, holding period, and experience. To link our data with earnings announcements information, we focus on firms that are headquartered in the United States, and thus, have regular filings with the SEC. It would be ideal to observe investors' opinions about individual firms every day for constructing a daily measure of disagreement. Thus, we concentrate on firms for which there is a high amount of StockTwits coverage. The top 100 firms mentioned comprise 60% of the overall number of messages in our sample. This leaves us with 1,460,349 messages by 11,874 unique users.[2] We present the names of the 100 firms and the frequency of messages about these firms in the Appendix (Table A.1). Not surprisingly, many of the most discussed firms are in technology and pharmaceutical industries.

Note that some users joined StockTwits after January 1, 2013. We control for the growing nature of our sample by including time fixed effects in our analysis. Out of 11,874 users, 4,566 joined before January 1, 2013.[3] Figure 2 portrays the number of messages over time in our data,

[2]To ensure that the cutoff of 100 firms does not arbitrarily influence our results, we reproduce our main findings in the Appendix (Table A.5) using the sample of the 150 most talked-about firms, obtaining similar results. We also replicate our main results using the top 50 firms and the top 51-100 firms, to ensure that our results are not driven by just the top stocks. Indeed, our measure of disagreement exhibits somewhat better properties for firms 51-100 than for the top 50 firms.

[3]In the Appendix, we conduct a robustness test to ensure that the potentially changing composition of the investors is not affecting our results by repeating the analysis using just users who joined Stocktwits before January 1, 2013. As can be seen in the Appendix (Table A.5), the results are very similar.

indicating no dramatic changes over time aside from the steady growth in the number of users and messages posted.

Table 2 presents summary statistics of the sample coverage. The median number of messages per firm per day is 10, with as many as 5,000 messages on some days for some firms. Since the typical firm has multiple messages per trading day in the data, we are able to calculate measures of disagreement at the day-firm level.

1.2 StockTwits Users

To register with StockTwits, a user reports via an online form his or her investment approach, investment horizon (or holding period), and experience level. In Table 2 Panel B, we present the breakdown of users by investment approach, holding period, and experience. On StockTwits, the most common approach is technical, representing 38% of users and also posting about 38% of messages. Momentum and growth investors represent the next two most common investment philosophies (20% and 18% of investors, respectively), followed by fundamental and value investors. Although some groups of investors post more than their fair share (momentum investors) and some less (value investors) the posting frequencies across investor groups is similar. To the best of our knowledge this is the first paper to directly measure investors' approaches, and therefore, we cannot compare whether this breakdown is representative of other samples in the market.

Next, we examine the holding horizons of investors. A plurality of investors (44%) are swing traders, who tend to have an investment horizon from a couple of days to a couple of weeks. The next biggest group is position traders, whose investment horizon is usually several months. The day traders and long term investors each makeup about 15% of the investors. About 20% are professional and about 30% are novices. Consistent with likely trading behaviors, professionals post disproportionately more messages than novices or intermediates.

It is important to evaluate the representativeness of the opinions on the platform if we hope to speak to broad-market disagreement using the unique features of StockTwits. Validating the empirical measures from StockTwits is especially important for self-reported measures, such as investment approach. We hand-checked a number of user profiles using identifying information from the profile (such as a real name), and found that the self-reported experience appears to be a reliable indicator of the user's experience. Figure 1 presents three examples of user profiles, one for each experience level in the data, to give a sense for this comparison. The *novice* investor is a student, who is mostly trading for fun, the *intermediate* investor reports real life trading experience, but seems to be less active. Finally, the *professional* investor has over 30 years of trading experience and has worked in the IBM PIT.

Although we do not have strong reasons to suspect that investors are untruthful about their investment approaches, one might worry that investors select these approaches without understanding what they mean. To address this potential limitation, we study the text of the messages users post to StockTwits to examine whether users actually adhere to the investment approach they select when they register (i.e., fundamental, technical, momentum, global macro, growth, or value). Specifically, we extract the 400 most commonly used words for each approach. We drop any words from these lists that are used commonly by users across multiple strategies because the distinctive, commonly-used words should most aptly capture the essence of an approach's language. At this stage, we also remove typos and references to stock tickers. We present these trimmed lists of words for each approach in the Appendix Table A.2.[4] Interestingly, the distinctive words used by investors align closely with the investment approaches that they claim to follow. For example some of the most common words for fundamental investors are "eps" (earnings per share) and "announcement", whereas technical investors refer to "charts", "area", and "head." This evidence is consistent with investors following the investment approaches that they self-report on their profile.

We also examine when investors post the messages. We are interested to know whether investors post the messages as they update their beliefs when news occur, or in the evenings after work, when they have more free time. In Figures 3 and 4 we plot the distribution of messages by the day of the week and by the hour of the day. It is evident that investors predominantly post messages when the markets are open (Monday-Friday and between 9am and 4pm). This is consistent with investors updating their messages in real time as financial events unfold.

1.3 Why do users post messages?

For constructing a measure of disagreement, it is essential that the sentiment expressed on StockTwits reveals the true opinions of investors. Thus, we want to rule out that users are trying to

[4] Since we have very few users who self-report to follow the Global Macro approach, that strategy has a lot of noise, and so we truncate the number of words we present.

manipulate the stock market by posting fake opinions. For example, if a user thinks the stock price will go down and thus wants to sell the stock, she could post really positive messages, in an attempt to increase the price temporarily, which would allow her to sell at a higher price. This would invalidate our measure, as we would capture her opinion as bullish, even though she is bearish on the stock. This does not appear to be an important concern in our data for several reasons. First, there is anecdotal evidence that investors post on social networks to attract followers, and gain internet fame or a job.[5] In all those cases, it is in their best interest to provide their best forecast of the future stock performance, and thus their honest opinion about the stock. Second, these firms have large market caps, and therefore it is very unlikely that individual investors think they can move the stock price.

2 Sentiment

2.1 Sentiment measure

When using StockTwits, users can post a message (limited to 140 characters) and indicate their sentiment as bullish, bearish, or unclassified (the default option). The following figure presents an image of the interface.

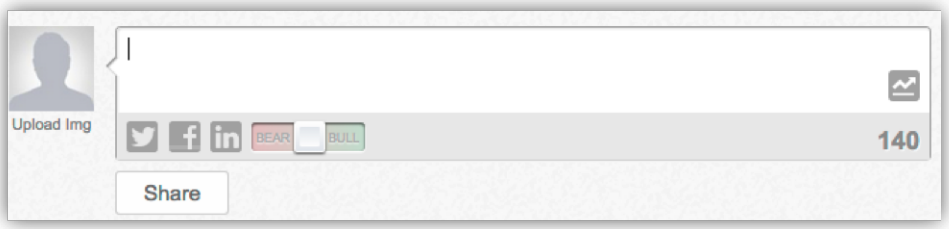


Table 2 Panel C, column 1, shows the distribution of sentiment across messages in the original sample. According to these summary statistics, 18.3 percent of classified messages are bearish and 81.7 percent are bullish. Even though the setting and time period are different, our classifications give similar relative frequencies to the distribution reported in Antweiler and Frank (2004) who hand classify individual trader messages on an Internet message board.

From reading the unclassified messages, it is clear that most of them are quite bullish or quite

[5]For an example of an article on the fame motive for posting to investment social networks see the Wall Street Journal article from April 21, 2015 (article here).

bearish, but the user did not select the option. To incorporate this information into the analysis, we follow prior literature and use natural language methods to classify the unclassified messages into bearish and bullish ones.[6] We use a maximum entropy-based method (described in the Appendix) to classify messages, that were unclassified in the original sample, as either bearish or bullish. Furthermore, we train our algorithm and use it to classify messages separately by investment approach to account for the possibility that investors with different approaches use different terminology to describe positive or negative sentiment. Table 2 Panel C, column 2, shows the distribution of sentiment in the final dataset. We have 458,218 bearish and 1,001,788 bullish messages.

2.2 Validating the Sentiment Measure

We validate the sentiment measure in four ways. First, we utilize a cross validation method while classifying the messages that were unclassified by the users. Second, we show that the unclassified messages do not seem to significantly differ from classified messages in the probability of bullish opinion, based on the language used. Third, we show that the measure correlates sensibly with proxies for investor short sale constraints. Finally, we examine whether sentiment is related to future stock performance.

2.2.1 Cross Validating the Sentiment Classification

Using most of the original classified data for training the model and a small subset to test the algorithm, we are able to comment on the accuracy of our classification method. On average, the overall accuracy rate is 83%. This high degree of accuracy enhances our confidence in using the classification scheme on unclassified messages.

2.2.2 Unclassified versus Classified Messages

One potential concern we have about the unclassified messages is that investors are more certain of their sentiment when they tag their message as bullish or bearish than when they leave sentiment

[6] Prior papers that use message data (e.g., Antweiler and Frank (2004), Giannini et al. (2015)) must construct a training dataset (usually ~1,000 messages) by classifying the messages by hand, calibrating a classification model (usually based on maximum-entropy methods) to this self-constructed training set of messages, and then using the calibrated model to classify the rest of the data. In our setting, we avoid the subjectivity of hand classification because 475,303 messages were pre-classified by the users as bullish or bearish. This training sample is both larger and more accurate because the users report their sentiment directly to StockTwits.

unclassified. We conduct a cross-validation exercise to examine whether this is likely to be the case. For users who classify at least one message in the data set, we randomly select 100,000 pre-classified messages to train the maximum entropy algorithm. Then we randomly draw another set of 100,000 messages, where users expressed their sentiment, and 100,000 messages that were left unclassified, and apply the maximum entropy algorithm to those messages. The algorithm assigns the probability that each message is bullish or bearish such that a higher probability of a message being bullish means that more words that are usually associated with bullish sentiment were used in the message. We examine whether the unclassified messages differ in their likelihood of being bullish or bearish from the user-classified ones. The two distributions are almost identical with the mean probability being 0.958 for unclassified and 0.959 for pre-classified messages, and the standard deviation being 0.104 and 0.105, respectively. This confirms that the unclassified messages are very similar in nature to the user-classified ones.[7]

2.2.3 Expressed Sentiment versus Trading

One potential concern with an expressed sentiment measure like ours is that expressed opinions might reflect a behavioral bias toward broadcasting positive information. We address this concern by relating the propensity to report negative news to the likelihood that an investor without an inventory of the stock cannot trade because of short selling constraints. Given that many investors face short selling constraints (Hong and Stein, 2003, Engelberg et al., 2014), a tilt toward bullish sentiment is natural. A bearish investor with a strict short sale constraint can only sell the stock until her inventory is zero. Investors with limited attention tend to neglect information on stocks for which they have zero inventory (Davies, 2015). Zero inventory stocks are likely to be the stocks for which investors are bearish, and because these stocks get less investor attention, bearish messages would be reported less frequently.

Using percent of institutional ownership of a firm as a proxy for shorting constraints (Nagel, 2005), we find that the the fraction of bullish messages for companies in the top quartile of institutional holdings (lax shorting constraints) is 0.54, compared with 0.28 for companies in the bottom quartile (tight shorting constraints). This evidence suggests that the bullish-bearish imbalance in

[7] In the Appendix, we replicate our main findings in Table A.5 using only messages that were classified by the investors themselves (user-classified messages), and we obtain qualitatively similar results.

our sentiment measure is most likely due to the short selling constraints.

2.2.4 Sentiment and Stock Returns

To further evaluate our sentiment measure, we examine whether investor sentiment forecasts stock returns by analyzing the abnormal cumulative returns of portfolios formed using the daily frequency of bullish and bearish messages. Specifically, we evaluate the performance of two portfolios based on the sentiment of StockTwits users: a bullish portfolio and a bearish portfolio. We use the bullish or bearish message frequencies as portfolio weights for each portfolio.[8] We construct cumulative returns over the following 60 days for each of the two portfolios and subtract out the value-weighted market index. We rebalance the portfolios daily.

Figure 6 (a) presents a graph of the cumulative abnormal returns for the bullish and bearish portfolios for the overall sample. The cumulative returns for each portfolio are initially flat, and then increase over the coming months. Firms for which investors are bullish exhibit similar performance as firms for which investors are bearish. This finding is consistent with prior findings that investors, especially retail investors, cannot predict returns, on average. Despite this average finding, it is possible that some subsets of investors are better at predicting returns. We present separate bearish and bullish portfolios by experience level (i.e., novice, intermediate, and professionals) to shed light on this question in Figure 6 (b) - (d). The portfolios that follow novice recommendations exhibit very poor performance. A bullish recommendation from a novice tends to forecast *lower* returns than a bearish recommendation. This is in line with prior research that individual investors lose money in the market, even before accounting for transaction costs (Barber and Odean, 2000). Interestingly, a portfolio that follows the recommendations of StockTwits professionals yields positive abnormal returns, suggesting that experienced investors have some ability to forecast returns (either by taking priced risks, or by identifying mispricing). In fact, professionals appear to do quite well, outperforming the market by almost 2% over a 60 trading-day period.

[8] To be concrete, consider an example where there are two potential firms (A and B) and 20 bullish messages were posted in total. In this scenario, if firm A had 15 bullish messages and firm B had 5 bullish messages then firm A will get a weight of 0.75 and firm B a weight of 0.25 in the "bullish portfolio."

2.3 Average Sentiment Measure

We follow Antweiler and Frank (2004) in combining individual opinions into one measure of sentiment. We code each bearish message as -1, and each bullish message as 1, and take the arithmetic average of these classifications at the $firm \times day \times group$ level:

$$AvgSentiment_{itg} = \frac{N_{itg}^{bullish} - N_{itg}^{bearish}}{N_{itg}^{bullish} + N_{itg}^{bearish}}. \tag{1}$$

The $AvgSentiment_{itg}$ measure ranges from -1 (all bearish) to $+1$ (all bullish). A group can either be all investors or investors with a given investment philosophy, holding period, or experience level. As a significant fraction of messages are posted outside of market hours (see Figure 4), we calculate the average sentiment measure for day t from messages posted between the market close of day $t-1$ to the market close of day t. Figure 5 presents the timing of our measurement.

From reading the messages, many investors post new messages as their sentiment changes. They might not necessarily go from being bullish to being bearish, but they might feel more or less strongly about a given stock. Therefore, the sentiment expressed by the users is a useful measure of how the average investor changes or updates sentiment, rather than the overall level of sentiment. From the standpoint of understanding daily trading volume, this is appropriate because changes in sentiment are what lead to trading rather than the overall level of sentiment. If no messages were posted for a given firm/day/group, we set the average sentiment measure equal to 0, as we assume that no change in sentiment occurred from the last time the users posted

Table 2, Panel D displays the summary statistics of average sentiment change for all users, and then broken down by investment philosophy. As seen with the distribution of bullish and bearish messages, investors tend to express more bullish sentiment, on average. Therefore, it is not surprising that the average sentiment for all users is 0.372. Interestingly, investors who self-report to follow a growth investment philosophy are the most likely to post bullish messages, whereas fundamental investors are the most likely to post bearish messages. We present the summary statistics of the sentiment measure broken down by experience level and holding period in the Appendix (Table A.3).

For our main measure, we equal-weight the sentiment of each message. As a robustness to our main measure of sentiment, we also calculate a "follower-weighted" average sentiment measure by weighting the sentiment of each message by the number of followers of the user who posted the message. As we show in the Appendix, our broad findings are not sensitive to the choice of weights in the calculation of the average sentiment (see Table A.5).

3 Cross-Group Disagreement

3.1 Investment Philosophies and Variation in Sentiment

The fact that we observe sentiment for the same firm separately for distinct groups of investors allows us to construct a direct test for whether adherence to a particular group of investors (i.e., a particular approach, holding period philosophy, or experience level) is a source of disagreement in the market. Taking it one step further, we are able to evaluate how important these investment philosophy affiliations are to overall disagreement. When investors disagree, there will be variation in the sentiment expressed for the same firm on the same day. One test for whether differing investment philosophies lead to disagreement is to evaluate whether these affiliations explain variation in expressed sentiment. Moreover, the amount of variation that is explained by accounting for investment philosophies can help quantify the extent to which these philosophies matter to overall changes in disagreement.

To test for this, we estimate the following regression specification of sentiment by approach, date, and firm:

$$AvgSentiment_{itg} = FirmFEs + TimeFEs + ApproachFEs + \varepsilon_{itg} \quad (2)$$

where $AvgSentiment_{itg}$ is the average sentiment on date t for firm i by investors of a given investment approach g. We include firm, time (month, year, and day-of-week), and approach fixed effects to explicitly compare the explanatory power of different investment models to the amount of variation in sentiment captured by differences across firms and across time.

Table 3, Panel A, presents the analysis of variance (ANOVA) decomposition from specifications that estimate equation (2). Firm and time dummies alone explain 10.1 percent of the variation in sentiment changes. Adding the approach fixed effects explains an additional percentage point of variation in sentiment. To put the importance of approach styles in context, differing approaches explain approximately 9.9 percent of the changes in disagreement (variation in sentiment) that is explained using firm and time fixed effects.

In the Appendix (Table A.4) we also estimate equation (2) for different experience levels and investment horizons. The F-statistics indicate that all types of categories are statistically significant, and thus, are sources of model disagreement. Yet, investment horizon and experience levels explain much less variation in sentiment – approximately 0.7 percent of firm fixed effects for experience and approximately 2.3 percent of firm fixed effects for horizon. Therefore, for the rest of the paper, we concentrate on disagreement across different investment approaches in our main analysis.

If differing investment philosophies are important for disagreement, they should also be important for trends in how sentiment is expressed over time. For example, if fundamental investors and technical investors respond to the same information differently, accounting for approach fixed effects should explain different trends in expressed sentiment. To test for this, we estimate the following regression specification of first-differenced sentiment by approach, date, and firm:

$$\Delta AvgSentiment_{itg} = FirmFEs + TimeFEs + ApproachFEs + \varepsilon_{itg} \tag{3}$$

where $\Delta AvgSentiment_{itg}$ is first-differenced average expressed sentiment on date t for firm i by investors of approach g.

In Panel B of Table 3, we present the ANOVA decomposition of sentiment trends from estimating equation (3). Similar to the regressions of average sentiment, differing investment approaches explains 0.8% of the variation in first-differenced average sentiment. In contrast to the regressions of average sentiment, time and firm fixed effects explains little of the variation in first-differenced average sentiment, only explaining 0.5%. That is, differing investment philosophies is as important for explaining sentiment trends as changes in sentiment, despite these trends being more difficult to explain using firm and time fixed effects.

3.2 Quantifying Disagreement Across Groups

The results from prior section indicate that there is significant disagreement across investors with different investment philosophies. In this section, we construct explicit measures of cross-group disagreement, with the goal of evaluating the importance of cross-group disagreement for trading volume. Specifically, we construct a measure of overall cross-group disagreement by computing the standard deviation of the average sentiment measures ($AvgSentiment_{itg}$) within each day-firm observation across investment approaches (i.e., fundamental, technical, momentum, global macro, growth, and value investors).

$$\begin{aligned} CrossDisagreeement_{it} = \quad & SD\{AvgSentiment_{itF}, AvgSentiment_{itT}, AvgSentiment_{itM}, \\ & AvgSentiment_{itGM}, AvgSentiment_{itV}, AvgSentiment_{itG}\} \end{aligned} \tag{4}$$

Our disagreement measure captures changes to the level of disagreement because, as we discuss above, the *AvgSentiment* measure reflects changes in sentiment. Hereafter, we refer to our measures as "disagreement," though it is appropriate to think of the measure as capturing changes in investors' level of disagreement.

3.3 Trading Volume and Cross-Group Disagreement

To examine how the cross-group disagreement measure is related to the abnormal trading volume, we estimate the following regression specification:

$$\begin{aligned} AbLogVol_{it} &= \alpha + \beta_1 CrossDisagreement_{it} + AbLogVol_{it-1} \\ &+ TimeFEs + FirmFEs + \varepsilon_{it} \end{aligned} \tag{5}$$

where $AbLogVol_{it}$ is the abnormal log trading volume on day t for firm i, $CrossDisagreement_{it}$ is the cross-group disagreement measure described in the previous section. We also control for trading volume on day $t-1$ to account for persistence in abnormal trading volume. We include year, month, day-of-the-week, and firm fixed effects, and cluster the standard errors at the day and firm levels.

We standardize the disagreement measure by subtracting out the mean and dividing by the standard deviation.

The results from estimating equation (5) are presented in Table 4. In column (1) we regress abnormal trading volume on day t on changes in disagreement on the same day. The coefficient estimate is statistically significant and suggests that a one standard deviation larger change in disagreement is associated with 2.2% increase in abnormal trading volume. When we add firm fixed effects in column (2), the effect almost doubles: now a one standard deviation larger change in disagreement is associated with a 4.1% increase in abnormal trading volume. In column (3) we examine whether changes in cross-group disagreement on day $t-1$ forecast abnormal trading volume on day t for stock of firm i. We find that the coefficient is statistically insignificant, which is consistent with the idea that changes in disagreement are reflected in trading volume the same day.

In column (4), we alleviate the concern that disagreement among investors is merely a reaction to changes in trading volume by regressing abnormal trading volume on day t on disagreement before the market opens on day t. The timing of pre-market messages is illustrated in Figure 5. In this case, the disagreement measure clearly leads the trading volume measure in time. To account for autocorrelation among trading volume, we also control for abnormal trading volume on day $t-1$. As can be seen in column (4), one standard deviation higher change in disagreement overnight (before the market opens on day t) is associated with a 4.2% increase in abnormal trading volume after the market opens on day t. This suggests that our disagreement measure is not driven by changes in trading volume.

4 Cross-Group versus Within-Group Disagreement

Hong and Stein (2007) propose that disagreement can arise from two non-mutually exclusive sources: differences in information sets among investors or differences in models that investors use to interpret that information. The cross-group disagreement results likely reflect model disagreement rather disagreement from having differing information sets. To evaluate the relative importance of model disagreement to overall level of disagreement, we now turn to comparing cross-group with within-group disagreement.

4.1 Measuring Disagreement within Groups

Following Antweiler and Frank (2004), we construct a disagreement measure based on the standard deviation of expressed sentiment. Using the fact that sentiment opinions are expressed as a binary variable (-1/1), the variance of the sentiment measure during a time period t can be calculated as $1 - AvgSentiment^2$. Although Antweiler and Frank (2004) studied disagreement using opinions expressed across the whole set of investors, we adapt this insight to study disagreement within investor groups. Specifically, the within-group measure for a given $firm \times day \times group$ is computed as:

$$Disagreement_{itg} = \sqrt{1 - AvgSentiment_{itg}^2} \tag{6}$$

where a "group" can represent all investors, or only investors with a given investment approach, experience, or investment horizon. This disagreement measure ranges from 0 to 1, with 0 being no change in disagreement and 1 being maximum change in disagreement. To illustrate the properties of the disagreement measure consider the following example. Assume that there are 10 messages by fundamental investors about Apple on a given day. In Figure 7, we show how the disagreement measure changes as the number of bearish messages goes from 0 (all bullish messages) to 10 (all bearish messages). There is no change in disagreement if everyone's opinion changes in the same direction – all messages are either bearish or bullish. The change in disagreement is maximized at 1, when investors' opinions change in opposite directions - when there are 5 bullish and 5 bearish messages. Since the measure is a square root function, the disagreement measure changes the most when there are few bullish or few bearish messages (the measure has the largest slope).

Because StockTwits users are most likely to post when their sentiment about a firm changes, we set our disagreement measure for the given group and day to be 0 if no messages were posted that day.[9] Intuitively, if no information came out that fundamental investors viewed as informative, we assume that their opinions about the firm (and thus their disagreement) have not changed.

[9] This choice deviates from how Antweiler and Frank (2004) handle stock-days where no messages come out. If there are no messages posted during a given time period, Antweiler and Frank (2004) set disagreement for that time period to be 1, and justify it by saying that no information came out during that time period, and thus there is latent disagreement. As the opinions on StockTwits likely reflect changes of opinion, it is more appropriate in our context to set the case of no messages to be a change in disagreement of 0.

4.2 Summary of Disagreement across and within Groups

In Table 5 we summarize our disagreement measure across and within groups. The first three rows summarize disagreement for all investors, disagreement across investment philosophies, and the average disagreement within investment philosophies. The average for our main disagreement measure for all investors is 0.47, and the median is 0.637. The average cross-group disagreement is 0.39, and the average within-group disagreement is 0.25.

The patterns of within-group disagreement for different investment approaches also provide interesting insight. Technical investors disagree the most with one another, whereas value, fundamental, and growth investors disagree much less with investors of the same investment philosophy. This finding resonates with the fact that there are many ways to be a technical investor, but much more standardization in what value, and growth investing means.[10] We also summarize within-group disagreement by investor experience and by investment horizon in the Appendix (Table A.3).

4.3 Within-Group versus Cross-Group Disagreement and Trading Volume

To quantify the importance of cross-group disagreement versus within-group disagreement in explaining trading volume, we estimate the following regression specification:

$$\begin{aligned} AbLogVol_{it} &= \alpha + \beta_1 CrossDisagreement_{it} + \beta_2 WithinDisagreement_{it} \\ &+ AbLogVol_{it-1} + TimeFEs + FirmFEs + \varepsilon_{it} \end{aligned} \tag{7}$$

Where $AbLogVol_{it}$ is the abnormal log trading volume on day t for firm i, $CrossDisagreement_{it}$ is the cross-group disagreement measure for stock i, on day t. In column (2) $WithinDisagremeent_{it}$ is the average of within-group disagreement measures, and in column (3) $WithinDisagremeent_{it}$ stands for within-group disagreement measures of individual investment approaches. All disagreement measures are standardized by subtracting out the mean and dividing by the standard deviation. We also control for trading volume on day $t-1$ to account for persistence in abnormal trading volume. As in our other specifications, we include year, month, day-of-the-week, and firm fixed effects, and

[10] For example, many technical investors use the subjective method of finding patterns in charts, and therefore often come to opposite conclusions.

cluster the standard errors at the day and firm levels.

The results from estimating equation (7) are presented in Table 6. In column (1) we confirm our prior result that a one-standard deviation increase in cross-group disagreement is associated with a 4.1% increase in trading volume. In column (2) we examine whether this relationship between cross-group disagreement and abnormal trading volume changes after controlling for within-group disagreement. To do so, we calculate the within-group disagreement for each investment philosophy and take the average of those for each firm i day t. The effect of cross-group disagreement doesn't change when we control for within-group disagreement, which suggests that these are two distinct sources of disagreement. In column (3), we control for disagreement within individual investment-philosophy groups, and find that the effect of cross-group disagreement is smaller than disagreement among technical and fundamental investors, and slightly larger than the disagreement within the other investment philosophies. Because the different investment philosophies that we observe are coarsely defined,[11] and some within-group disagreement likely still arises from differing investment models, these relative magnitudes are likely a lower bound on the importance of model disagreement to abnormal trading volume.

In conjunction with our findings in the variation of sentiment specifications in Table 3, these findings suggest that the models investors use to interpret information matter for the extent of disagreement among investors in the marketplace. Thus we expect that improvements to the informational efficiency of markets, will not completely erode stock market volatility.

5 Disagreement Among All Investors

The findings in the prior section indicate that both cross-group and within-group disagreements are significantly related to trading volume. In this section, we focus on the measure of disagreement across all investors, and draw some comparisons between our measure and notable alternatives used in the prior literature. We further examine how our measure of disagreement relates to investor attention in explaining trading volume, and show that our measure of disagreement among all investors provides useful insight into trading volume patterns around earnings announcements.

[11] For example, some technical investors use "subjective" and some use "objective" methods to find patterns in price and trading volume data.

5.1 Comparing Disagreement Measures

In this section, we examine how well our disagreement measure correlates with alternative measures of disagreement (e.g., analyst dispersion as in Diether et al. (2002), return volatility, and divergence of sentiment on StockTwits from sentiment expressed in the media, as in Giannini et al. (2015)), and with abnormal trading volume. Broadly, we find a weak correlation between our measure and existing measures of disagreement, but a strong correlation between our measure and abnormal trading volume, which enhances confidence that the measure is a useful proxy for disagreement.

First, we evaluate the correlation between our disagreement measure and analyst dispersion. We examine the disagreement among all investors, disagreement among groups with different investment philosophies, and the average disagreement within those groups. Following prior literature, we calculate a monthly measure of analyst dispersion using the standard deviation of analyst earnings forecasts made in a given month. To compare our measure to this monthly measure of analyst dispersion, we compute the average of our measure over the month, then calculate its correlation with analyst dispersion. As can be seen in Table 7, column (1), the two measures do not correlate with one another significantly. On some level, this low correlation between our measure and analyst dispersion is to be expected. Our measure captures high-frequency and recent disagreement about the prospects of a stock based on the sentiment of actual traders, whereas analyst dispersion is much lower frequency and is issued by analysts rather than traders. In column (2) we examine the correlation of our disagreement measures with return volatility. Interestingly, the cross-group disagreement is negatively correlated with both analyst dispersion and return volatility, whereas the within-group measure has a positive correlation. This suggests that analyst dispersion and return volatility are better measurements of information-driven disagreement.

As another comparison, we reconstruct a measure of divergence of opinion on StockTwits from sentiment expressed in the media, as in Giannini et al. (2015). We find that cross-group disagreement is more correlated with their measure than within-group disagreement, consistent with the fact that StockTwits users and media might have different models for processing financial information.[12]

[12] Giannini et al. (2015) measure the divergence between investor sentiment on StockTwits and the sentiment of breaking news articles and firm press releases. Their measure is akin to a cross-group disagreement measure where one group is all StockTwits users, and the other group is whomever posts in the media. Unlike our analysis, Giannini et al. (2015) do not evaluate how different groups of StockTwits investors disagree with one another. To quantitatively evaluate how their style of measuring disagreement contrasts with ours, we reproduce an alternative measure that – like Giannini et al. (2015) – contrasts investor sentiment on StockTwits with media sentiment as reported in the Ravenpack database. Ap-

We further evaluate the performance of disagreement measures by examining their correlation with abnormal trading volume. Aside from being a useful way to evaluate a measure of disagreement, explaining trading volume is interesting unto itself. What exactly drives trading volume and why it varies so much over time is still subject to much debate in the finance literature (e.g., see pages 111-112 of Hong and Stein, 2007). Relevant to our tests, one theory is that trading volume reflects differences in investors' opinions about the prospects of a stock. Despite the compelling logic, there is not much empirical support for a correlation between existing measures of disagreement and abnormal trading volume.

Consistent with this body of evidence, when we correlate analyst dispersion at the monthly level with abnormal trading volume, there is a weak and insignificant correlation (0.0388). In contrast, our measure of disagreement correlates much more strongly with abnormal trading volume. Specifically, in Table 7, column (4), we present the correlations between daily abnormal log trading volume and our daily measures of investor disagreement. We find that the correlation of market-wide disagreement and the abnormal log trading volume is 0.117. This represents a substantial improvement in the ability to explain abnormal trading volume. As we have seen before, the abnormal trading volume is more correlated with the average within-group disagreement than the cross-group disagreement measure.

In Appendix Section 7.4, we also provide evidence that disagreement within investment philosophies is not purely a reaction to trading volume in the stock market, by showing that changes in market-level disagreement predict future changes in trading volume. We further provide evidence that high levels of overall disagreement negatively forecast future stock returns, consistent with prior theory (Hong and Stein, 1999).

5.2 Measuring Disagreement versus Measuring Attention

In this section, we evaluate whether our measure of investor disagreement is distinct from investor attention, both by controlling for proxies for attention and examining the interaction between disagreement and attention. This exercise is also useful for explaining what causes investors to trade with one another. Intuitively, we should expect greater trading volume when both disagreement and

pendix 7.3 presents precise details on how we construct this alternative measure of disagreement, but our goal is to stay as close as possible to the Giannini et al. (2015) measure in an out-of-sample replication of their proxy for disagreement.

attention are high (related to this, see our results on earnings announcements in the next section). We approximate the amount of attention by using the total number of StockTwits messages posted about a stock on a particular day. In addition, we also use the Google Search Volume Index (SVI) measure proposed as a measure of attention by Da et al. (2011), which measures the frequency of stock ticker searches on Google for firm i on day t.[13] Using these proxies for attention, we estimate the following specification:

$$\begin{aligned} AbLogVol_{it} = {} & \alpha + \beta_1 Disagreement_{it} + \beta_2 InvestorAttention_{it} \\ & + \beta_3 Disagreement_{it} \times InvestorAttention_{it} \\ & + \gamma AbLogVol_{it-1} + TimeFEs + FirmFEs + \varepsilon_{it} \end{aligned} \tag{8}$$

where $Disagreement_{it}$ is the disagreement among all investors about stock i on day t, and $InvestorAttention_{it}$ is either the StockTwits message volume or the Google SVI for the stock on that particular day.

Table 8 presents the results from estimating equation (8). We conduct our analysis on firms for which we observe Google Search Volume Index. In columns (1) and (3), we see that our measure of disagreement predicts trading volume, even after holding constant proxies for investor attention. Columns (2) and (4) show that the relationship between disagreement and abnormal trading volume is strongest when proxies for attention are high. Furthermore, column (5) presents a horse race between the StockTwits measure of attention (message volume) and the Google SVI measure. The two measures appear to measure a similar kind of attention, as the magnitudes on both interactions decline slightly when the other is included. The results in this section provide evidence that while investor attention matters for trading volume (as has been shown in prior literature), disagreement among investors also plays an important role in predicting trading volume, and is a separate concept from investor attention.

5.3 Disagreement and Trading around Earnings Announcements

We now turn to examining the relationship between disagreement and the well-known spike in trading volume around earnings announcements. On its face, the fact that trading volume increases after earnings announcements is puzzling because firms release important financial information to

[13] For the exact construction of Google SVI at the daily level see Niessner (2016).

the market during this time, which should resolve uncertainty. Nonetheless, it is a robust feature of the market that volume goes up after earnings announcements, and remains high for several weeks (Drake et al., 2012; Kaniel et al., 2012). Recent theoretical work on this phenomenon points to a role for disagreement to resolve the puzzle (Banerjee et al., 2015). However, without a reliably high-frequency measure of disagreement like our measure, it is difficult to provide evidence for this conjecture. Thus, our setting positions us to provide one of the first empirical tests of the role of disagreement in explaining volume changes around earnings announcements.

To examine whether greater disagreement can explain the spike in volume around earnings announcements, we use our measure of disagreement for all investors to examine how volume changes around earnings announcements in the following regression:

$$\begin{aligned} AbLogVol_{it} = \alpha + \beta_1 1WeekBeforeEA_{it} + \beta_2 EA_{it} + \beta_3 1WeekAfterEA_{it} \\ + \beta_4 2WeekAfterEA_{it} + \beta_5 3WeekAfterEA_{it} + \beta_6 Dis_{it} + SUE_{iq} + \Gamma'\mathbf{X} + TimeFEs + FirmFEs + \varepsilon_{it} \end{aligned} \tag{9}$$

where $AbLogVol_{it}$ is the abnormal log trading volume on day t for firm i, $1WeekBeforeEA$ is a dummy variable equal to 1 if day t for firm i falls in the week before an earnings announcement for that firm, EA_{it} is a dummy variable equal one if firm i announces earnings on day t, $1WeekAfterEA_{it}$, $2WeekAfterEA_{it}$, $3WeekAfterEA_{it}$ are dummy variables for whether day t for firm i falls in week 1, week 2, or week 3 after an earnings announcement, respectively. SUE_{iq} is the earnings surprise for firm i in quarter q defined as the difference in reported earnings minus the median analyst forecast. Finally, in some specifications, we control for the amount of disagreement Dis_{it} at the firm-day level, and include interactions between disagreement and the timing dummy variables (captured in $\Gamma'\mathbf{X}$).

The results from estimating equation (9) are presented in Table 9. Column (1) replicates the existing finding in the literature that volume spikes on the earnings announcement date, and remains high for three weeks after the earnings announcement. The coefficients on $WeekBeforeEA_{it}$, EA_{it}, $1WeekAfterEA_{it}$, $2WeekAfterEA_{it}$, and $3WeekAfterEA_{it}$ are relative to the time outside of these weeks. Based on the coefficient estimate on $WeekBeforeEA_{it}$, the trading volume before an EA is approximately the same as it is during the time outside of the earnings announcement period. On the day of the announcement, trading volume increases by 67%, and stays high (39% higher) for one week and then slowly decreases over time. Note that abnormal trading volume is still 5% higher than normal three weeks after the earnings announcement.

Columns (2) and (3) of Table 9 present a test of the role of disagreement as measured by our measure of disagreement. To the extent that our measure of disagreement captures this spike in volume, we should expect the coefficient on EA_{ti} to diminish as we control for our disagreement measure. Indeed, we find that controlling for disagreement can explain approximately one eighth of the spike in abnormal volume around the earnings announcement (0.572 versus 0.670 on the earnings announcement date). Controlling for interactive effects of disagreement allows for the effect of disagreement to be different by date relative to the earnings announcement. In this specification, we observe that our within-group measure of disagreement can explain up to 26 percent of the spike in abnormal volume on the earning announcement day. These findings are important, especially because there are very few predictors that can explain changes in abnormal volume.

In columns (4) through (7), we estimate the model on subsamples, split by whether the earnings surprise was positive (columns (4) and (5)) or negative (columns (6) and (7)). In either case, controlling for our measure of disagreement explains a significant fraction of the volume spike on the earnings announcement day, but the explanatory power is better for negative earnings surprises than positive earnings surprises (40% vs. 20%).

While we do not expect that trading by investors who post on StockTwits makes up a significant proportion of the overall trading in the stock market, the results in this section suggest that our measure of disagreement is a good proxy for overall changes in disagreement in the market.

6 Conclusion

In this paper, we utilize the unique features of a data set of messages posted by individual investors on a social investing network to construct measures of disagreement within and across investor groups with different investment approaches. We exploit the fact that users frequently self-classify their sentiment about a given firm as bullish or bearish, and that we also observe their self-reported investment philosophy to study causes of disagreement among investors. Although there has been much theoretical work suggesting that disagreement among investors can be due to differences in information sets or differences in models, there has been very scant empirical research evaluating sources of disagreement, mainly due to data limitations.

Using coarsely-defined investment approach classifications, we find that approximately half of overall disagreement among investors is driven by differences across models, and that both cross-group and within-group disagreement contribute to the high trading volume in the market. Our measure of disagreement strongly predicts future intra-day trading volume, suggesting that our disagreement measures proxy well for changes in market-wide disagreement. Finally, we address an empirical puzzle in the literature that trading volume spikes right after earnings announcements and stays high for several weeks, even though information uncertainty is resolved during these time periods. We show that changes in disagreement can explain up to one third of the increase in trading volume after earnings announcements, providing corroborative evidence that disagreement across different models drives a significant amount of daily trading volume,

The findings in this paper that model-driven disagreement is a large source of overall disagreement in the market have implications for regulators. Given that regulators expand vast resources in an attempt to make markets more informationally efficient, it is important to keep in mind, that even if markets become perfectly informationally efficient, this will not erode high trading volume and volatility, on account of significant model-driven disagreement.

References

Ajinkya, B. B., R. K. Atiase, and M. J. Gift (1991). Volume of Trading and the Dispersion in Financial Analysts' Earnings Forecasts. *The Accounting Review*.

Antweiler, W. and M. 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.

Ataise, R. K. and L. S. Bamber (1994). Trading volume reactions to annual accounting earnings aannouncement: the incremental role of predisclosure information asymmetry. *Journal of Accounting and Economics*.

Bailey, M., R. Cao, T. Kuchler, and J. Stroebel (2016). Social networks and housing markets.

Bamber, L. S., O. E. Barron, and D. E. Stevens (2011). Trading Volume around Earnings Announcements and Other Financial Reports: Theory, Research Design, Empirical Evidence, and Directions for Future Research. *Contemporary Accounting Research*.

Banerjee, S., J. Davis, and N. Gondhi (2015). When transparency improves, must prices reflect fundamentals better? *Working Paper*.

Banerjee, S. and I. Kremer (2010). Disagreement and Learning: Dynamic Patterns of Trade. *Journal of Finance*.

Barber, B. M. and T. Odean (2000). Trading is Hazardous to Your Wealth: The Common Stock Investment Performance of Individual Investors. *The Journal of Finance*.

Barber, B. M. and T. Odean (2008). All that Glitters: The Effect of Attention and News on the Buying Behavior of Individual and Institutional Investors. *Review of Financial Studies 21*(2), 785–818.

Baron, M., J. Brogaard, and A. Kirilenko (2012). The trading profits of high frequency traders. *Unpublished Manuscript*.

Carlin, B. I., F. A. Longstaff, and K. Matoba (2014). Disagreement and asset prices. *Journal of Financial Economics 114*(2), 226–238.

Chang, Y.-C., H. G. Hong, L. Tiedens, N. Wang, and B. Zhao (2014). Does diversity lead to diverse opinions? evidence from languages and stock markets. *Rock Center for Corporate Governance at Stanford University Working Paper* (168), 13–16.

Cookson, J. A. (2016). When Saving Is Gambling. *SSRN Working Paper 2517126*.

Da, Z., J. Engelberg, and P. Gao (2011). In Search of Attention. *The Journal of Finance*.

Davies, S. W. (2015). Retail Traders and the Competitive Allocation of Attention. *Working Paper*.

Diether, K. B., C. J. Malloy, and A. Scherbina (2002). Differences of opinion and the cross section of stock returns. *The Journal of Finance 57*(5), 2113–2141.

Dorn, D. and P. Sengmueller (2009). Trading as Entertainment? *Management Science*.

Drake, M. S., D. T. Roulstone, and J. R. Thornock (2012, September). Investor Information Demand: Evidence from Google Searches Around Earnings Announcements. *Journal of Accounting Research 50*(4), 1001–1040.

Engelberg, J., A. V. Reed, and M. Ringgenberg (2014). Short selling risk. *Working Paper, University of California, San Diego, University of North Carolina, and Washington University in St. Louis.*.

Giannini, R. C., P. J. Irvine, and T. Shu (2015). The convergence and divergence of investors? opinions around earnings news: Evidence from a social network. *Available at SSRN 2017748*.

Grinblatt, M. and M. Keloharju (2009). Sensation Seeking, Overconfidence and Trading Activity. *The Journal of Finance*.

Hong, H. and J. C. Stein (1999). A Unified Theory of Underraction, Momentum Trading, and Overaction in Asset Markets. *The Journal of Finance*.

Hong, H. and J. C. Stein (2003). Differences of opinion, short-sales constraints, and market crashes. *Review of financial studies 16*(2), 487–525.

Hong, H. and J. C. Stein (2007). Disagreement and the stock market. *The Journal of Economic Perspectives*, 109–128.

Kaniel, R., S. Liu, G. Saar, and S. Titman (2012). Individual Investor Trading and Return Patterns around Earnings Announcements. *Journal of Finance 67*(2), 639–680.

Karpoff, J. M. (1986). A Theory of Trading Volume. *Journal of Finance*.

Kondor, P. (2012). The more we know about the fundamental, the less we agree on the price. *The Review of Economic Studies 79*(3), 1175–1207.

Kumar, A. (2009). Who Gambles in the Stock Market? *The Journal of Finance*.

Linnainmaa, J. (2011). Why do (some) households trade so much? *Review of Financial Studies*.

Milgrom, P. and N. Stokey (1982). Information, Trade, and Common Knowledge. *Journal of Economic Theory*.

Nagel, S. (2005). Short sales, institutional investors and the cross-section of stock returns. *Journal of Financial Economics 78*(2), 277–309.

Niessner, M. (2016). Strategic Disclosure Timing and Insider Trading. *SSRN Working Paper 2439040*.

Nigam, K., J. Lafferty, and A. McCallum (1999). Using maximum entropy for text classification. In *IJCAI-99 workshop on machine learning for information filtering*, Volume 1, pp. 61–67.

Rogers, J. L., D. J. Skinner, and S. L. C. Zechman (2015). Run EDGAR Run: SEC Dissemination in a High-Frequency World. *Chicago Booth Research Paper No. 14-36*.

Rothschild, D. M. and R. Sethi (2014). Trading strategies and market microstructure: Evidence from a prediction market. *Available at SSRN 2322420*.

Varian, H. R. (1985). Divergence of Opinion in Complete Markets: A Note. *Journal of Finance*.

7 Appendix

7.1 Alternative Disagreement Measure

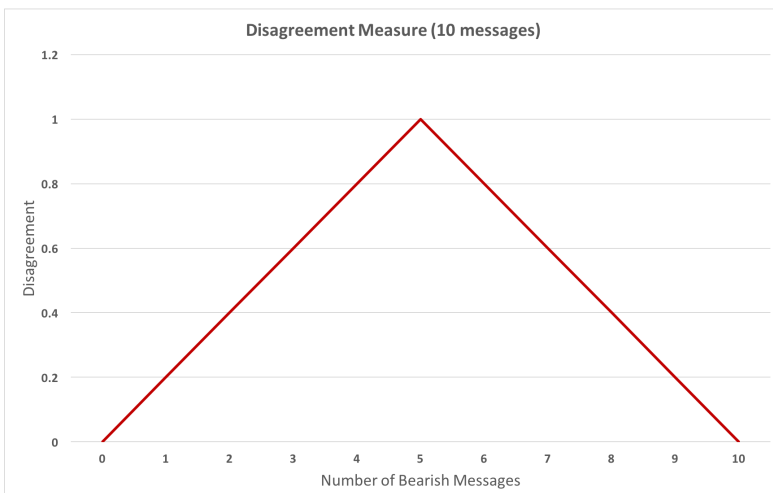
As mentioned in Section 4, the Antweiler-Frank disagreement measure is calculated as

$$D = \sqrt{1 - AvgSentiment^2}$$

Since it's a square-root function, it has the largest slope if there are very few bullish or very few bearish messages. We follow that method in our main analysis. However, as a robustness test, we also use a function that is liner in the average sentiment measure.

$$D^* = 1 - |AvgSentiment|$$

This disagreement measure for an example with 10 messages is depicted in the figure below.



Using this measure the slope of the disagreement function remains the same as the fraction of bearish messages increases in the market. In Table A.5 we rerun our analysis using this measure of disagreement and get qualitatively similar results as our main disagreement measure.

7.2 Maximum Entropy Method

There are a plethora of text and document learning algorithms that have been shown (empirically and theoretically) to yield desirable misclassification rates. Some of the more popular methods are maximum entropy, naive Bayes, k-nearest neighbor, and support vector machines. Here, we give a brief outline of the maximum entropy approach.

Excluding neutral opinions, “sentiment" is a binary variable and therefore a standard logistic regression model can be used to estimate the proportion of bullish investors. Classification can be done by thresholding these probabilities. This technique, also known as a maximum entropy classifier, uses labeled training data to fix a collection of constraints for the model that define the class-specific averages. We will use training data to fix constraints on the conditional distributions of the learned distribution (the condition probability of bullish or bearish classification given a particular message). The goal is to find the distribution $p^{\star}$, satisfying these constraints, that maximizes the entropy quantity

$$H(p) = \sum_{x \in \mathscr{X}} p(x) \log\left(\frac{1}{p(x)}\right),$$

where p is a probability mass function that belongs to a collection of mass functions $\mathscr{C}$ satisfying the constraint. That is,

$$p^{\star} = \mathrm{argmax}_{p \in \mathscr{C}} H(p).$$

Let $\mathscr{M}$ denote our dataset. Let $m \in \mathscr{M}$ denote a message and define $f_w(m, c(m))$ to be equal to the proportion of times the word w appears in the message m when it is classified as $c(m)$. Here, $c(m)$ can be either “bearish" or “bullish". We explicitly write $c(m)$ to emphasize the dependence of the class on the message m. We stipulate that the conditional distribution of the class given the message $p(c|m)$ satisfy

$$\frac{1}{|\mathscr{M}|} \sum_{m \in \mathscr{M}} f_w(m, c(m)) = \frac{1}{|\mathscr{M}|} \sum_{m \in \mathscr{M}} \sum_{c} p(c|m) f_w(m, c),$$

for all words w we consider informative. In the above notation, $\mathscr{C}$ is the collection of all probabilities $p(c|m)$ satisfying the above constraints. Then we choose

$$p^{\star}(c|m) = \mathrm{argmax}_{p(c|m) \in \mathscr{C}} H(p(c|m)).$$

Using the concavity of the logarithm, it can be shown that

$$p^{\star}(c|m) = \frac{\exp\{\sum_w \lambda_w f_w(m,c)\}}{\sum_c \exp\{\sum_w \lambda_w f_w(m,c)\}},$$

where the λ_w are estimated from the data. We classify a message m as bearish or bullish according to a 0.5 threshold for $p^{\star}(c|m)$. For more details on this method, we refer the reader to Nigam et al. (1999). We performed the maximum entropy algorithm separately within the six types of investment approach: growth, technical, value, momentum, fundamental, and global macro.

7.3 Producing a Disagreement Measure in the Spirit of Giannini et al. (2015)

In Giannini et al. (2015), the authors download all breaking news and company press releases that mention the company name or the company ticker from PR News Wire, Dow Jones News Wire, and Reuters News Wire from Factiva news database. They then use the maximum entropy approach to estimate the sentiment of every news article. We adopt a conceptually-similar approach that is more easily replicable by turning to Ravenpack (a news database that collects and classifies news articles and company press releases) as that is much more readily available. The advantage of using Ravenpack is that Ravenpack produces a standardized classification methodology for sentiment of articles about firms, which avoids the need to replicate the time-intensive maximum entropy approach in constructing a measure analogous to Giannini et al. (2015). Further, the advantages extend to other researchers and practitioners, who can adopt a similar methodology to construct a Giannini et al. (2015)-like measure of disagreement.

Using Ravenpack, we collect company press releases from PR News Wire and Dow Jones News Wire. Ravenpack uses proprietary methods to assign a sentiment score to every article, which we use to classify articles into "bearish" and "bullish" categories. We then follow Giannini et al. (2015) in constructing the IMPACT and the NEWS measures, where the former measures the StockTwits sentiment and the later captures the news media sentiment. We calculate these measures at the firm-day level.

To calculate the IMACT measure at the daily level, we first assign each StockTwits message a -1 or 1, based on whether the message was bearish or bullish, and then weigh each message by 1 plus the number of followers the author of the message has. In other words, for an individual message

$IMPACT = (1 + Followers) \times Sentiment$. We then add the IMPACT score for every message to the firm-day level.

We repeat the above procedure with press releases, by assigning -1 or 1 to each article, based on it's sentiment, and then add up those sentiment scores for each firm to the daily level. To calculate the final disagreement measure, at the firm-day level, we follow Giannini et al. (2015) and define disagreement (DIVOP) to be 0 if both IMPACT and NEWS are either positive or both are negative (there is agreement), and 1 otherwise (there is disagreement).

Note that our reproduction of the Giannini et al. (2015) measure is not an exact replication of their original measure, as we use the Ravenpack data instead of manually downloading the Factiva articles. However, the replicated measure has the same concept - difference in sentiment between the media and the StockTwits messages, and we believe that this is a reasonable approach to take for someone who wants to replicate the original measure.

7.4 Disagreement, Volume, and Returns

Because disagreement is measured for the same time period as abnormal trading volume, a potential concern remains that any changes in the disagreement of StockTwits investors are purely a reaction to the trading volume in the stock market. We alleviate this concern by examining whether disagreement predicts future changes in trading volume.

To examine whether our disagreement measure forecasts future trading volume, we estimate the following regression specification:

$$\begin{aligned} AbLogVol_{it} &= \alpha + \beta Disagreement_{it} + \gamma AbLogVol_{it-1} \\ &+ TimeFEs + FirmFEs + \varepsilon_{it} \end{aligned} \tag{10}$$

where $Disagreement_{it}$ is our disagreement measure for firm i in time period t. For ease of interpretation, we standardize the measure by subtracting the mean and dividing by the standard deviation, over the entire sample period. *AbLogVol* is the difference between log volume in timer period t and the average log volume from $t-140$ to $t-20$ trading days (6-month period, skipping a month). Since trading volume tends to be autocorrelated, we also control for abnormal trading volume on day $t-1$. We include year, month, day-of-the-week, and firm fixed effects. The standard errors are clustered at the date and firm levels.

The results are presented in Table A.6. In column (1) we examine whether market-wide disagreement on day $t-1$ forecast abnormal trading volume on day t for stock of firm i. We find that the coefficient is statistically insignificant, which is consistent with the idea that changes in disagreement are reflected in trading volume the same day. In column (2) we regress abnormal trading volume on day t on changes in disagreement on the same day. The coefficient estimate is statistically significant and suggests that a one standard deviation larger change in disagreement is associated with 1.1% increase in trading volume.

To alleviate the concern that disagreement among investors merely reflects changes in trading activity, in column (3), we regress abnormal trading volume on day t on disagreement among messages that were posted before the market opened on day t, as illustrated in Figure 5. In this case, the disagreement measure clearly leads the trading volume measure in time, but the connection between the two is more immediate than in column (1). To account for autocorrelation among trading volume, we also control for abnormal trading volume on day $t-1$. As can be seen in column (3), one standard deviation higher change in disagreement overnight (before the market opens on day t) is associated with a 5.4% increase in abnormal trading volume after the market opens on day t. This suggests that our disagreement measure is not fully driven by changes in trading volume. In fact, approximately half of the effect of contemporaneous disagreement (0.054 versus 0.110) can be attributed to messages that were posted before the trading volume is observed.

Finally, we also examine the relationship between investor disagreement and subsequent stock returns. In theory, greater disagreement could forecast either higher or lower future returns. Theories based on disagreement among optimists and pessimists suggest that greater disagreement should forecast negative returns (Hong and Stein, 1999), whereas other theories where disagreement is a priced risk factor suggest a positive return premium when there is more disagreement (Carlin et al., 2014). To evaluate this tension empirically, we estimate the following regression specification for abnormal stock returns on day $t+1$ and cumulative abnormal returns over days $t+1$ to $t+5$:

$$\begin{aligned} Abret_{i,t+1} &= \alpha + \beta DisMeasure_{it} + \nu AvgSentiment_{it} + \phi Abret_{it} \\ &\quad + \gamma AbLogVol_{it} + \delta LogME_{it} + \varepsilon_{it} \end{aligned} \tag{11}$$

where $Abret_{i,t+1}$ is the abnormal return (minus the value-weighted market index) for firm i on day $t+1$, $DisMeasure_{it}$ is our disagreement measure on day t. Some specifications also control for $AvgSentiment_{it}$ to alleviate the concern that the result arises from a mechanical correlation of our disagreement measure with sentiment. Moreover, we examine the stock market response starting the following day, which alleviates the concern of disagreement reacting to returns.

Table A.7 presents the results from estimating equation (11), with and without controlling for average sentiment. In column (1) we see that a standard deviation increase in disagreement is associated with a 6 basis points decrease in next business day's returns. The estimates in column (2), which also control for average sentiment, show that this relationship is not mechanically due to the relationship between our disagreement measure and sentiment, but arises from disagreement, conditional on sentiment.

In columns (3) and (4), we predict cumulative abnormal returns for days $t+1$ to $t+5$ ($CAR[1,5]$) using our measure of disagreement. According to these specifications, a one standard deviation increase in disagreement is associated with a 12 basis point decrease in returns over the following week. Moreover, this effect is not due to a mechanical relationship to investor sentiment as the effect of disagreement is a very similar magnitude when controlling for sentiment.

8 Tables and Figures

8.1 Figures

Figure 1: Examples of StockTwits User Profiles

Note: This figure presents screenshots of representative user profiles from StockTwits, illustrating the difference between novice, intermediate and professional StockTwits users.

(a) Novice Trader Profile

(b) Intermediate Trader Profile

(c) Professional Trader Profile

Figure 2: Monthly Time Series of Messages Posted to StockTwits

Note: This figure portrays the aggregate number of messages posted to StockTwits for each month in our 21-month sample (from January 2013 to September 2014).

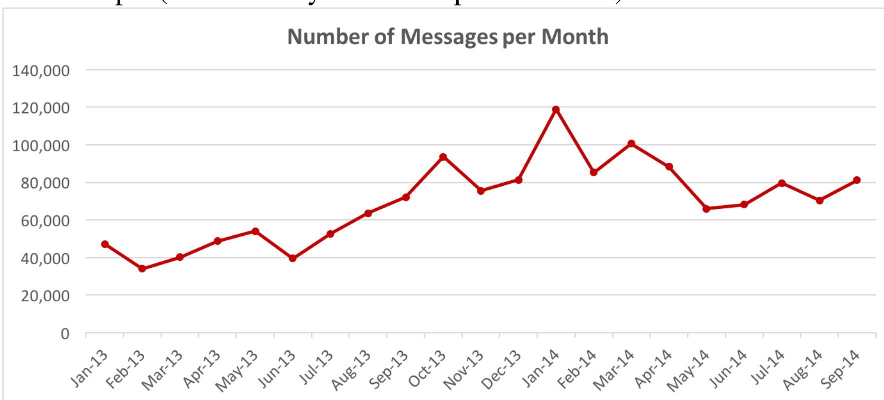


Figure 3: Day-of-Week Frequency Distribution of Messages Posted

Note: This figure presents a frequency distribution of the days of the week that messages are posted to StockTwits.

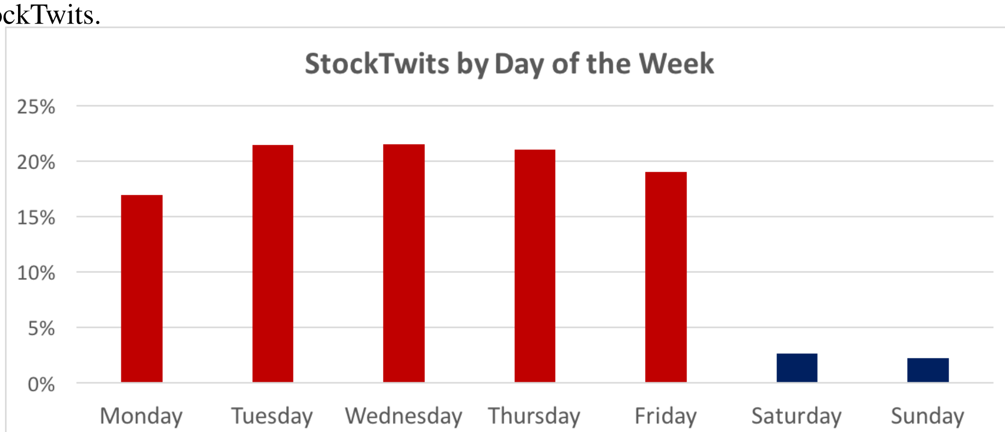


Figure 4: Hour-of-Day Frequency Distribution of Messages Posted

Note: This figure presents a frequency distribution across the hour of the day (Eastern Standard Time) at which messages are posted to StockTwits. Trading hours are plotted in red, whereas non-trading hours are plotted as blue bars.

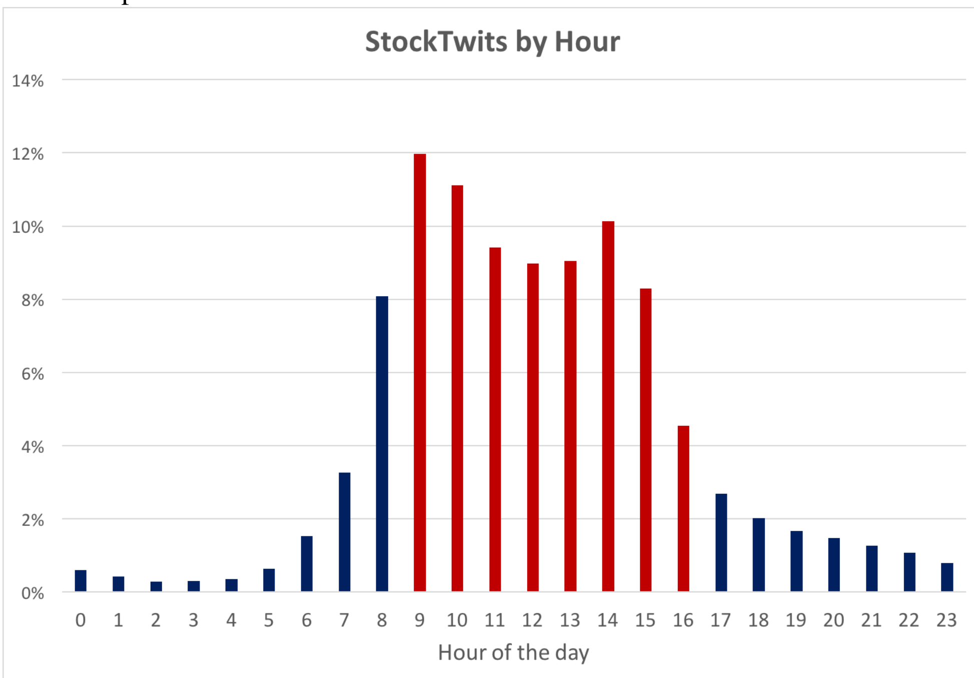


Figure 5: Timeline for Calculating Disagreement

Note: This figure presents how we calculate changes in disagreement. Since trading stops at 4pm on day $t-1$, we assign any messages that are posted on day $t-1$ after 4pm to trading day t. The same way we assign any messages posted after 4pm on day t to day $t+1$. To calculate "overnight" changes in disagreement, before the market opens (BMO) on day t, we include messages that are posted after 4pm on the previous day until 9am on day t.

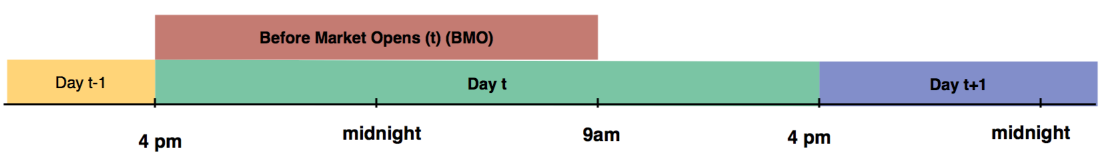
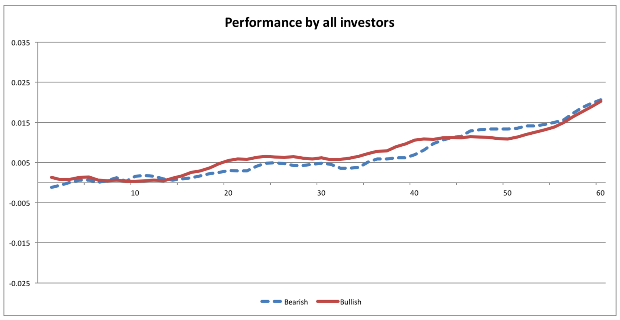


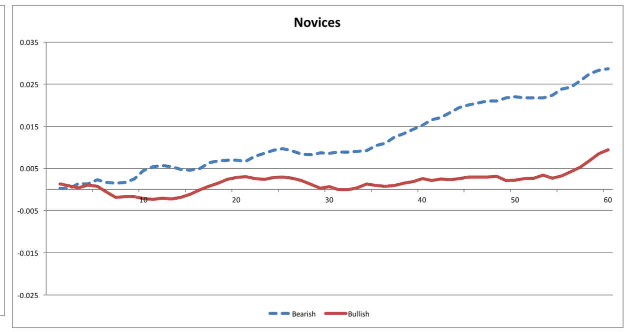
Figure 6: Performance of StockTwits Sentiment Strategies

Note: This figure presents the cumulative abnormal returns of strategies that buy when sentiment is bullish and sell when sentiment is bearish for several sentiment classifications: (a) the sentiment of all StockTwits users ("All Investors"), (b) the sentiment of Novices, (c) the sentiment of Intermediates, and (d) the sentiment of Professionals.

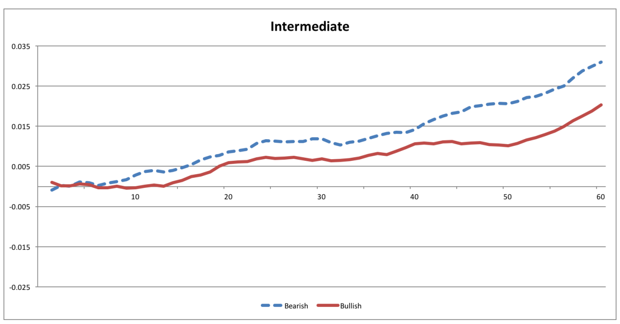
(a) All Investors



(b) Novices



(c) Intermediate



(d) Professionals

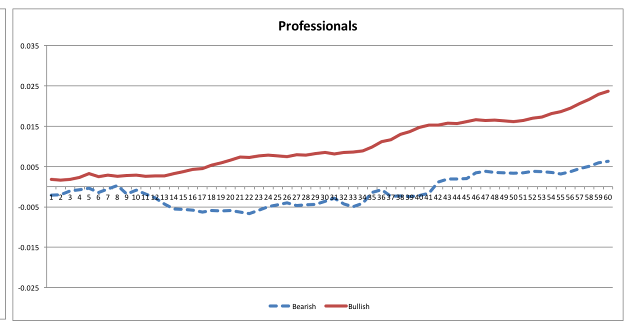
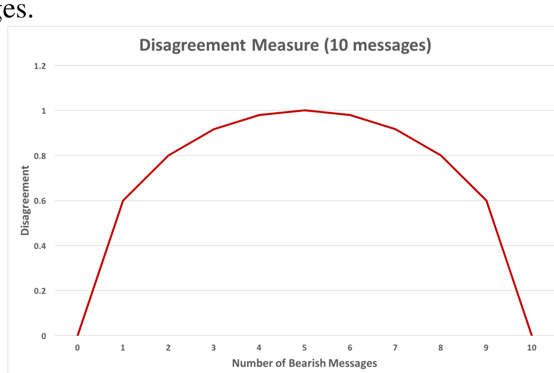


Figure 7: An Example of the Disagreement Measure

Note: This figure portrays how our main disagreement measure depends on the average sentiment of the underlying messages.



9 Tables

Table 1: Sampling Restrictions and the Size of the Analysis Sample

Note: In this table, we present the number of messages, number of unique StockTwits users, and number of company tickers covered as we clean the full sample to our final analysis sample.

Messages	Users	Tickers	Action
18,361,214	107,920	9,755	Original Sample
13,763,653	73,964	9,137	Years 2013 and 2014
7,315,198	56,551	8,558	Keep messages with 1 ticker per message
4,550,746	27,369	8,055	User must have non-missing approach and holding period and experience
3,928,842	25,109	6,326	Merge on CRSP
2,870,856	22,669	3,708	Stocks with at least one earnings announement
1,460,349	11,874	100	Keep top 100 firms

Table 2: Summary Statistics

Note: In this panel we report summary statistics from the StockTwits data. In particular, Panel A presents summary information on the coverage by stock and user, as well as user-level information. Panel B presents frequency distributions of users and messages posted by investment philosophy, holding period and experience, which are observed user profile characteristics. Panel C shows the distribution of bearish, bullish, and unclassified messages in the original sample in column 1, and the distribution of messages after we apply the maximum entropy (MaxEnt) approach to the unclassified messages, in column 2. Panel D presents the sentiment (average bullishness) by investment philosophy.

Panel A: Characteristics of Messages and Users

	Mean	Stdev	Min	p25	p50	p75	Max
Number of messages per stock	14,487	32,577	616	1,576	5,296	14,864	275,969
Number of meesages per user	121	391	1	5	19	82	11,759
Number of messages per stock per day	44	135	1	3	10	31	4,728
Sentiment stock/day	0.441	0.516	-1	0.170	0.5	1	1
Number of followers user has	187	1,972	0	1	5	18	84,657
Number of people user follows	43	193.7	0	4	15	45	9,990
Total Days Active	462	412	1	137	349	685	1,909

Panel B: Frequencies of User Profile Characteristics

Approach	Num. Users	Percent Users	Num. Messages	Percent Messages
Fundamental	1,475	12.42%	206,075	14.11%
Technical	4,510	37.98%	540,003	36.98%
Momentum	2,388	20.11%	381,290	26.11%
Global Macro	271	2.28%	13,008	0.89%
Growth	2,145	18.06%	221,174	15.15%
Value	1,085	9.14%	98,799	6.77%
Total	11,874	100%	1,460,349	100%

Holding Period	Num. Users	Percent Users	Num. Messages	Percent Messages
Day Trader	1,840	15.50%	266,075	18.22%
Swing Trader	5,257	44.27%	673,558	46.12%
Position Trader	2,644	22.27%	291,237	19.94%
Long Term Investor	2,133	17.96%	229,479	15.71%
Total	11,874	100%	1,460,349	100%

Experience	Num. Users	Percent Users	Num. Messages	Percent Messages
Novice	3,406	28.68%	239,170	16.38%
Intermediate	6,147	51.77%	806,534	55.23%
Professional	2,321	19.55%	414,645	28.39%
Total	11,874	100%	1,460,349	100%

Panel C: Sentiment Classification

	Number of Messages	
Sentiment	Original Sample	MaxEnt Classification
Bearish	87,193	458,218
Bullish	388,110	1,001,788
Unclassified	984,703	

Panel D: Sentiment Summary Statistics

	Average Sentiment	
	Mean	Stdev
All Investors	0.372	0.928
Fundamental	0.277	0.960
Technical	0.345	0.444
Momentum	0.387	0.921
Global Macro	0.417	0.908
Growth	0.505	0.862
Value	0.351	0.936

Table 3: Quantifying Disagreement Across Investment Models

Note: This table examines whether individuals with different investment *approaches* have different changes in their assessment of stocks. To do this, we run the following regression in Panel A:

$$AvgSentiment_{itg} = FirmFEs + TimeFEs + ApproachFEs + \varepsilon_{itg}$$

where $AvgSentiment_{itg}$ is the change in average sentiment for investment philosophy g, firm i, on date t. In this regression investment approach fixed effects capture whether differences in investors' investment philosophies explain changes in average sentiment. In Panel B we examine whether individuals with different investment philosophies have different accelerations in their disagreement. We run the following regression:

$$\Delta AvgSentiment_{itg} = FirmFEs + TimeFEs + ApproachFEs + \varepsilon_{itg}$$

where $\Delta AvgSentiment_{itg}$ is the difference between the average sentiment measure on day t and day $t-1$. The regressions include time (year, month and day-of-the-week) and firm fixed effects as noted in the columns. Standard errors are clustered by firm. *, **, and *** indicate statistical significance at the ten, five, and one percent level respectively. Standard errors are in parentheses.

Panel A: Analysis of Variance for Sentiment

		$Sentiment_{itg}$	
Sentiment Categories	(1)	(2)	(3)
Firm FEs	X	X	X
Year, month, day of week FEs		X	X
Investment philosophy FEs			X
R-squared	0.099	0.101	0.111
F-stat across categories			6.64
Observations	107,090	107,090	107,090

Panel B: Analysis of Variance for Sentiment Trends

		Δ $Sentiment_{itg}$	
Sentiment Categories	(1)	(2)	(3)
Firm FEs	X	X	X
Year, month, day of week FEs		X	X
Investment philosophy FEs			X
R-squared	0.005	0.005	0.013
F-stat across categories			6.47
Observations	106,988	106,988	106,988

Table 4: Cross-group Disagreement and Trading Volume

Note: This table examines whether our measure of changes in investor disagreement forecasts trading volume. We run the following regression:

$$AbLogVol_{it} = \alpha + \beta CrossDisagreement_{it} + \gamma AbLogVol_{it-1} + TimeFEs + FirmFEs + \varepsilon_{it}$$

Where in columns (1) and (2) $CrossDisagreement_{it}$ is our cross-group disagreement measure across different investment philosophies for firm i on day t, in column (3) it's the cross-group disagreement measure on day $t-1$, and in column (4) it is the cross-group disagreement measure constructed from messages that were posted before the market opens (BMO) (between 4pm on day $t-1$ and and 9am on day t). We standardize the disagreement measures by subtracting the mean and dividing by the standard deviation, over the entire sample period. $AbLogVol_{it}$ is the difference between log volume in time period t and the average log volume from $t-140$ to $t-20$ trading days (6-month period, skipping a month) for firm i. Since trading volume tends to be autocorrelated, we also control for abnormal trading volume on day $t-1$. The regressions include year, moth, day-of-the-week, and firm fixed effects. Standard errors are clustered by firm and date. *, **, and *** indicate statistical significance at the ten, five, and one percent level respectively. Standard errors are in parenthesis.

	Abnormal Log Volume (t)			
Disagreement measure	(1)	(2)	(3)	(4)
Cross-group Disagreement (t)	0.022***	0.041***		
	(0.008)	(0.009)		
Cross-group Disagreement (t-1)			0.001	
			(0.004)	
Cross-group Disagreement (BMO, t)				0.042***
				(0.006)
Abnormal Log Volume (t-1)	0.762***	0.737***	0.742***	0.735***
	(0.012)	(0.013)	(0.013)	(0.014)
Observations	42,415	42,415	42,322	42,415
R-squared	0.595	0.601	0.600	0.601
Year, month, day-of-the-week FEs	X	X	X	X
Firm FEs		X	X	X

Table 5: Disagreement Measure

Note: This table presents summary information on the StockTwits measure of disagreement. The first three rows show summary statistics for disagreement for all investors, disagreement across groups with different investment philosophies, and the average disagreement within groups with different investment philosophies. The table further shows the distribution of within-group disagreement by the individual investment philosophies.

	Mean	Stdev	Min	p25	p50	p75	Max
All Investors	0.469	0.447	0	0	0.637	0.932	1
Cross-group Disagreement	0.392	0.266	0	0.149	0.447	0.547	1
Average within-group Disagreement	0.249	0.285	0	0	0.181	0.400	0.997
Fundamental	0.214	0.385	0	0	0	0.631	1
Technical	0.360	0.439	0	0	0	0.866	1
Momentum	0.287	0.420	0	0	0	0.800	1
Global Macro	0.066	0.241	0	0	0	0.000	1
Growth	0.210	0.376	0	0	0	0.000	1
Value	0.173	0.361	0	0	0	0.000	1

Table 6: Within-Group Disagreement, Cross-Group Disagreement, and Trading Volume

Note: In this table we examine whether changes in cross-group disagreement, on top of changes in within-group disagreement, help explain changes in trading volume. We run the following regression

$$\begin{aligned} AbLogVol_{it} &= \alpha + \beta_1 CrossDisagreement_{it} + \beta_2 WithinDisagreement_{it} \\ &+ AbLogVol_{it-1} + TimeFEs + FirmFEs + \varepsilon_{it} \end{aligned}$$

Where $AbLogVol_{it}$ is the abnormal log trading volume on day t for firm i, $CrossDisagreement_{it}$ is the cross-group disagreement measure across different investment philosophies for stock i, on day t. In column (2) $WithinDisagremeent_{it}$ is the average of within-group disagreement measures for different investment philosophies, and in column (3) $WithinDisagremeent_{it}$ is the within-group disagreement measures for individual investment philosophies. We standardize all disagreement measures by subtracting the mean and dividing by the standard deviation, over the entire sample period. The regressions include year, moth, day-of-the-week, and firm fixed effects. Standard errors are clustered by company and date. *, **, and *** indicate statistical significance at the ten, five, and one percent level respectively. Standard errors are in parentheses.

	Abnormal Log Volume		
Disagreement Measure	(1)	(2)	(3)
Cross-Group Disagreement	0.041***	0.047***	0.046***
	(0.009)	(0.008)	(0.008)
Average within-group Disagreement		0.164***	
		(0.010)	
Fundamental Disagreement			0.034***
			(0.005)
Technical Disagreeement			0.064***
			(0.006)
Momentum Disagreeement			0.060***
			(0.005)
Growth Disagreeement			0.033***
			(0.004)
Value Disagreeement			0.036***
			(0.004)
Abnormal Log Volume (t-1)	0.737***	0.704***	0.703***
	(0.013)	(0.016)	(0.016)
Observations	42,415	42,415	42,415
R-squared	0.601	0.617	0.617
Year, month, day-of-the-week FEs	X	X	X
Firm FEs	X	X	X

Table 7: Other Disagreement Measures

Note: This table presents correlations between our main disagreement measure for all investors and other commonly used measures of disagreement (analyst dispersion, return volatility, and Gianini et. al. measure), as well as with abnormal log trading volume.

Disagreement among	Analyst Dispersion	Return Volatility	Giannini et al. measure	Abnormal Log Volume
All Investors	0.042	0.036	0.234	0.117
Cross-group Disagreement	-0.052	-0.173	0.390	0.058
Average within-group Disagreement	0.087	0.207	0.146	0.180

Table 8: Disagreement and Investor Attention

Note: This table examines whether our measure of disagreement complements investor attention in explaining abnormal trading volume. We run the following regression:

$$
\begin{aligned}
AbLogVol_{it} &= \alpha + \beta_1 Disagreement_{it} + \beta_2 InvestorAttention_{it} + \beta_3 Disagreement_{it} \times InvestorAttention_{it} \\
&+ AbLogVol_{it-1} + TimeFEs + FirmFEs + \varepsilon_{it}
\end{aligned}
$$

Where $Disagreement_{it}$ is the disagreement measure among all investors for a given firm i on day t. In columns (1) and (2) $InvestorAttention_{it}$ is the total number of messages posted on StockTwits about firm i on day t. In columns (3) and (4) $InvestorAttention_{it}$ is the abnormal Google Search Volume Index for ticker of firm i on day t. $AbLogVol_{it}$ is the difference between log volume in time period t and the average log volume from $t-140$ to $t-20$ trading days (6-month period, skipping a month) for firm i. Since trading volume tends to be autocorrelated, we also control for abnormal trading volume on day $t-1$. We standardize the disagreement measure and the total number of messages by subtracting the mean and dividing by the standard deviation, over the entire sample period. The regressions include year, moth, day-of-the-week, and firm fixed effects. Standard errors are clustered by company and date. *, **, and *** indicate statistical significance at the ten, five, and one percent level respectively. Standard errors are in parenthesis.

	Abnormal Log Volume (t)				
	(1)	(2)	(3)	(4)	(5)
Disagreement	0.035***	0.051***	0.088***	0.083***	0.060**
	(0.010)	(0.013)	(0.009)	(0.010)	(0.022)
Number of Messages	0.117***	0.176***			0.137***
	(0.031)	(0.041)			(0.035)
Disagreement x Number of Messages		0.107***			0.083***
		(0.030)			(0.025)
AbLog(Google SVI)			0.291***	0.236***	0.188***
			(0.028)	(0.025)	(0.025)
Disagreement x AbLog(Google SVI)				0.124***	0.075***
				(0.019)	(0.019)
AbLogVol(t-1)	0.691***	0.688***	0.675***	0.670***	0.658***
	(0.027)	(0.028)	(0.031)	(0.032)	(0.033)
Observations	27,437	27,437	27,437	27,437	27,437
R-squared	0.605	0.609	0.606	0.609	0.621
Year, month, day of week FEs	X	X	X	X	X
Firm FEs	X	X	X	X	X

Table 9: Disagreement and Trading Volume around Earnings Announcements

Note: In this table, we examine disagreement among investors and trading volume around earnings announcements. We run the following regression:

$$\begin{aligned} AbLogVol_{it} &= \alpha + \beta_1 1WeekBeforeEA_{it} + \beta_2 EA_{it} + \beta_3 1WeekAfterEA_{it} \\ &+ \beta_4 2WeekAfterEA_{it} + \beta_5 3WeekAfterEA_{it} + \gamma Disagreement_{it} \\ &+ \delta_1 Disagreement_{it} \times 1WeekBeforeEA_{it} + \delta_2 Disagreement_{it} \times EA_{it} \\ &+ \delta_3 Disagreement_{it} \times 1WeekAfterEA_{it} + \delta_4 Disagreement_{it} \times 2WeeksAfterEA_{it} \\ &+ \delta_5 Disagreement_{it} \times 3WeeksAfterEA_{it} + SUE_{iq} + TimeFEs + FirmFEs + \varepsilon_{it} \end{aligned}$$

Where $AbLogVol_{it}$ is the abnormal log trading volume on day t for firm i, $1WeekBeforeEA$ is a dummy variable equal to 1 if day t for firm i falls in the week before an earnings announcement for that firm, EA_{it} is a dummy variable equal one if firm i announces earnings on day t, $1WeekAfterEA_{it}$, $2WeekAfterEA_{it}$, $3WeekAfterEA_{it}$ are dummy variables for whether day t for firm i falls in week 1, week 2, or week 3 after an earnings announcement, respectively. $Disagreement_{it}$ is our measure of investor disagreement about stock i on day t. is the earnings surprise in quarter q for firm i, defined as the earnings minus the median analyst forecast. Columns (1)-(3) include all observations that are around earnings announcements with a non-missing earnings surprise, while columns (4) and (5) have observations with a positive earnings surprise and columns (6) and (7) have observations with a negative earnings surprise. We standardize the disagreement measure by subtracting the mean and dividing by the standard deviation, over the entire sample period. The regressions include year, moth, day-of-the-week, and firm fixed effects. Standard errors are clustered by company. *, **, and *** indicate statistical significance at the ten, five, and one percent level respectively. Standard errors are in parenthesis.

	Abnormal Log Volume						
	Full Sample			Positive Earnings Surprise		Negative Earnings Surprise	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
1 Week before EA	0.038	0.034	0.033	0.060**	0.065***	0.008	-0.024
	(0.024)	(0.022)	(0.024)	(0.023)	(0.022)	(0.037)	(0.035)
EA	0.670***	0.572***	0.499***	0.747***	0.595***	0.554***	0.335***
	(0.050)	(0.047)	(0.049)	(0.054)	(0.052)	(0.064)	(0.070)
1 Week after EA	0.399***	0.354***	0.314***	0.442***	0.355***	0.338***	0.241***
	(0.032)	(0.031)	(0.028)	(0.037)	(0.032)	(0.041)	(0.038)
2 Weeks after EA	0.117***	0.101***	0.073***	0.148***	0.114***	0.072*	0.013
	(0.022)	(0.022)	(0.022)	(0.025)	(0.024)	(0.036)	(0.035)
3 Weeks after EA	0.046**	0.038**	0.008	0.084***	0.056**	-0.008	-0.060*
	(0.020)	(0.019)	(0.019)	(0.024)	(0.023)	(0.037)	(0.035)
Disagreement		0.214***	0.199***		0.176***		0.262***
		(0.018)	(0.019)		(0.014)		(0.030)
Disagreement × 1 Week before EA			-0.024		-0.029*		-0.022
			(0.016)		(0.017)		(0.028)
Disagreement × EA			0.129***		0.112**		0.162**
			(0.043)		(0.048)		(0.061)
Disagreement × 1 Week after EA			0.104***		0.112***		0.090**
			(0.025)		(0.026)		(0.042)
Disagreement × 2 Week after EA			0.020		-0.002		0.039
			(0.019)		(0.022)		(0.035)
Disagreement × 3 Weeks after EA			0.007		-0.005		0.018
			(0.017)		(0.023)		(0.033)
SUE	-0.001	-0.020	-0.033	0.090	0.025	0.169	0.161
	(0.056)	(0.055)	(0.055)	(0.129)	(0.109)	(0.144)	(0.148)
Observations	33,111	33,111	33,111	20,129	20,129	12,908	12,908
R-squared	0.162	0.206	0.200	0.206	0.239	0.212	0.262
Year, month, dow FEs	X	X	X	X	X	X	X
Firm FEs	X	X	X	X	X	X	X

Appendix to:

Why Don't We Agree? Evidence from a Social Network of Investors

Table A.1: 100 Most Discussed Firms

Note: In this table we present tickers, names, and number of messages of the top 100 firms ranked by the number of messages posted to StockTwits that reference the firm's ticker.

Ticker	Name	Messages	Frequency	Ticker	Name	Messages	Frequency
AAPL	Apple Inc.	331,212	18.8%	ICPT	Intercept Pharmaceuticals Inc	6,045	0.34%
FB	Facebook Inc	140,258	7.96%	QCOR	Questcor Pharmaceuticals Inc	5,989	0.34%
TSLA	Tesla Motors Inc	109,200	6.2%	FCEL	FuelCell Energy Inc	5,897	0.33%
PLUG	Plug Power Inc	95,565	5.43%	CHTP	Chelsea Therapeutics International	5,876	0.33%
VRNG	Vringo, Inc	62,890	3.57%	TTWO	Take-Two Interactive Software	5,760	0.33%
TWTR	Twitter Inc	48,953	2.78%	GS	Goldman Sachs Group Inc	5,644	0.32%
NFLX	Netflix, Inc	38,572	2.19%	CMG	Chipotle Mexican Grill, Inc	5,608	0.32%
ARIA	Ariad Pharmaceuticals, Inc.	35,603	2.02%	GEVO	Gevo, Inc.	5,604	0.32%
KNDI	Kandi Technologies Group Inc	35,530	2.02%	Z	Zillow Group, Inc.	5,561	0.32%
INO	Inovio Pharmaceuticals Inc	33,746	1.92%	CLF	Cliffs Natural Resources Inc	5,418	0.31%
MNKD	MannKind Corporation	30,742	1.75%	FIO	Fusion-IO, Inc.	5,405	0.31%
JCP	JC Penney Company Inc	29,260	1.66%	HK	Halcon Resources Corp	5,354	0.3%
ZNGA	Zynga Inc	26,394	1.5%	RAD	Rite Aid Corporation	5,220	0.3%
GOOG	Alphabet Inc	26,291	1.49%	SWHC	Smith and Wesson Holding Corp	5,152	0.29%
AMD	Advanced Micro Devices	25,327	1.44%	CPRX	Catalyst Pharmaceuticals Inc	5,146	0.29%
GLUU	Glu Mobile Inc	23,692	1.35%	ACHN	Achillion Pharmaceuticals, Inc	5,098	0.29%
SCTY	SolarCity Corp	23,357	1.33%	KERX	Keryx Biopharmaceuticals	5,077	0.29%
AMZN	Amazon.com, Inc.	22,234	1.26%	RMTI	Rockwell Medical Inc	5,073	0.29%
BAC	Bank of America Corp	21,107	1.2%	APP	American Apparel Inc.	5,022	0.29%
UNXL	UniPixel Inc	20,672	1.17%	CYTR	CytRx Corporation	4,991	0.28%
PCLN	Priceline Group Inc	20,158	1.14%	IBM	International Business Machines Corp.	4,852	0.28%
YHOO	Yahoo! Inc.	19,804	1.12%	OPK	Opko Health Inc.	4,749	0.27%
DDD	3D Systems Corporation	19,448	1.1%	ACAD	ACADIA Pharmaceuticals Inc.	4,688	0.27%
RNN	Rexahn Pharmaceuticals, Inc	18,741	1.06%	MSTX	Mast Therapeutics Inc	4,665	0.26%
GALE	Galena Biopharma Inc	17,253	0.98%	VHC	VirnetX Holding Corporation	4,458	0.25%
GTAT	GT Advanced Technologies Inc	16,395	0.93%	NIHD	NII Holdings Inc.	4,436	0.25%
LNKD	LinkedIn Corp	15,085	0.86%	CRM	salesforce.com, inc.	4,402	0.25%
ARNA	Arena Pharmaceuticals, Inc	14,772	0.84%	IDRA	Idera Pharmaceuticals Inc	4,389	0.25%
GOGO	Gogo Inc	12,532	0.71%	CLSN	Celsion Corporation	4,383	0.25%
GPRO	GoPro Inc	12,312	0.7%	DGLY	Digital Ally, Inc.	4,372	0.25%
FSLR	First Solar, Inc.	12,184	0.69%	BBY	Best Buy Co Inc	4,352	0.25%
GILD	Gilead Sciences, Inc.	11,969	0.68%	SBUX	Starbucks Corporation	4,229	0.24%
GMCR	Keurig Green Mountain Inc	11,578	0.66%	SPWR	SunPower Corporation	4,214	0.24%
YELP	Yelp Inc	10,807	0.61%	USU	Centrus Energy Corp	4,214	0.24%
P	Pandora Media Inc	10,361	0.59%	MNGA	MagneGas Corporation	4,176	0.24%
FEYE	FireEye Inc	10,205	0.58%	NAVB	Navidea Biopharmaceuticals	4,151	0.24%
ONVO	Organovo Holdings Inc	10,004	0.57%	AA	Alcoa Inc	4,096	0.23%
MU	Micron Technology, Inc	9,818	0.56%	DRL	Diadem Resources Limited	3,979	0.23%
F	Ford Motor Company	9,342	0.53%	S	Sprint Corp	3,963	0.22%
LULU	Lululemon Athletica inc	9,249	0.53%	ISRG	Intuitive Surgical, Inc.	3,945	0.22%
WLT	Walter Energy Inc	9,222	0.52%	NEON	Neonode, Inc	3,884	0.22%
GRPN	Groupon Inc	8,681	0.49%	ZGNX	Zogenix, Inc.	3,843	0.22%
ISR	IsoRay, Inc.	8,394	0.48%	BA	Boeing Co	3,797	0.22%
MCP	McPherson's Ltd	8,109	0.46%	SHLD	Sears Holdings Corp	3,788	0.22%
RXII	RXi Pharmaceuticals Corp	8,084	0.46%	V	Visa Inc	3,697	0.21%
MSFT	Microsoft Corporation	7,500	0.43%	CAT	Caterpillar Inc.	3,669	0.21%
INVN	InvenSense Inc	7,253	0.41%	ZLCS	Zalicus Inc.	3,660	0.21%
SRPT	Sarepta Therapeutics Inc	6,315	0.36%	CPST	Capstone Turbine Corporation	3,631	0.21%
EBAY	Ebay Inc.	6,266	0.36%	SGYP	Synergy Pharmaceuticals Inc	3,631	0.21%
CYTK	Cytokinetics, Inc.	6,140	0.35%	PCYC	Pharmacyclics, Inc.	3,598	0.2%

Table A.2: Common and Distinctive Words Used by Approach

Note: This table examines whether StockTwits users actually seem to adhere to the investment approach they select when they register (i.e., fundamental, technical, momentum, global macro, growth, or value). We concentrate on 400 mostly commonly used words for each approach. We then focus on words that are not duplicated across strategies, and remove any typos, numbers or references to stock tickers. The leftover lists of words for each approach are presented below. Since we have so few users who self-report to follow the Global Macro approach, that strategy has a lot of noise, and so we truncate the number of words we present.

Approach	**Most Common Unique Words**
Fundamental	mobile, boys, billion, phase, mid, fine, thought, pump, upgrade, fundamentals, eps, means, announcement
Technical	flag, fill, area, intraday, setup, channel, swing, candle, posted, consolidation, updated, charts, setting, head, closed, uptrend, fall, downside, loss, triangle, key, update, base
Momentum	covering, bid, stops, minutes, place, enjoy, rise, float, heading, lose, train
Value	sorry, chance, pipeline, undervalued, forward, fun, launch, bring, upgrades, wonder, approval
Growth	seen, either, car, else, street, world, quot, fuel, due, old,
Global Macro	tech, demand, per, called, corner, find, option, annual, cfo, doubt, event, net, opening, gold, ...

Table A.3: Sentiment and Disagreement Measures by Experience and Holding Period

Note: This table presents summary information on the StockTwits measure of sentiment and disagreement for different experience levels and for different holding periods, as reported in the StockTwits user profile.

Panel A: Sentiment Summary Statistics

	Average Sentiment	
	Mean	Stdev
Novice	0.390	0.920
Intermediate	0.396	0.917
Professional	0.314	0.949
Day Trader	0.294	0.955
Swing Trader	0.376	0.926
Position Trader	0.419	0.907
Long Term Investor	0.389	0.921

Panel B: Disagreement Summary Stastics

	Within-group Disagreement						
	Mean	Stdev	Min	p25	p50	p75	Max
Novice	0.254	0.407	0	0	0	0.745	1
Intermediate	0.398	0.445	0	0	0	0.904	1
Professional	0.323	0.437	0	0	0	0.866	1
Day Trader	0.265	0.415	0	0	0	0.800	1
Swing Trader	0.386	0.445	0	0	0	0.904	1
Position Trader	0.281	0.419	0	0	0	0.800	1
Long Term Investor	0.220	0.389	0	0	0	0.943	1

Table A.4: Quantifying Disagreement Across Investment Models

Note: This table examines whether individuals with different experience levels or different holding periods have different changes in disagree over their assessment of stocks. To do this, we run the following regression in Panel A:

$$AvgSentiment_{itg} = FirmFEs + TimeFEs + GroupFEs + \varepsilon_{itg}$$

where $AvgSentiment_{itg}$ is the change in average sentiment for for group g (e.g., experience level or holding period), firm i, on date t. In this regression Group fixed effects capture whether differences in groups that investors belong to explain changes average sentiment. In Panel B we examine whether individuals with different experience levels or different holding periods have different accelerations in their disagreement. We run the following regression:

$$\Delta AvgSentiment_{itg} = FirmFEs + TimeFEs + GroupFEs + \varepsilon_{itg}$$

where $\Delta AvgSentiment_{itg}$ is the difference between the average sentiment measure on day t and day $t-1$. The regressions include time (year, month and day-of-the-week) and firm fixed effects as noted in the columns. Standard errors are clustered by firm. *, **, and *** indicate statistical significance at the ten, five, and one percent level respectively. Standard errors are in parentheses.

Panel A: Analysis of Variance for Sentiment

			$Sentiment_{itg}$			
Sentiment Categories	(1)	(2)	(3)	(4)	(5)	(6)
Firm FEs	X	X	X	X	X	X
Year, month, day of week FEs		X	X		X	X
Experience FEs			X			
Holding Period FEs						X
R-squared	0.153	0.155	0.156	0.005	0.132	0.133
F-stat across categories			3.85			4.47
Observations	75,278	75,278	75,278	90,941	90,941	90,941

Panel B: Analysis of Variance for Sentiment Trends

			Δ $Sentiment_{itg}$			
Sentiment Categories	(1)	(2)	(3)	(4)	(5)	(6)
Firm FEs	X	X	X	X	X	X
Year, month, day of week FEs		X	X		X	X
Experience FEs			X			
Holding Period FEs						X
R-squared	0.001	0.003	0.003	0.001	0.001	0.004
F-stat across categories			5.83			4.79
Observations	75,278	75,278	75,278	90,941	90,941	90,941

Table A.5: Robustness of Main Results to Different Sampling Restrictions and Measurement Choices

Panel A

Note: In this panel we present the average changes in sentiment, average changes in disagreement, and the correlation between our disagreement measure and the abnormal log volume for different robustness specifications. Column (1) presents results for our main specifications. In column (2) when we construct the sentiment and disagreement measures we weigh each message by the number of followers the author of the message has. In column (3) we only include opinions by investors who joined StockTwits before 1 January, 2013. In column (4) we only use messages that were classified by users themselves as bullish or bearish. In column (5) we use a linear disagreement measure described in the appendix. In column (6) we only consider top 50 most talked-about firms. In column (7) we only consider top 51-100 most talked-about firms. In column (8) we consider the top 150 most talked-about firms.

	(1) Main dataset	(2) Weighted Disagreement	(3) Joined before 1 Jan 2013	(4) User-classified messages	(5) Linear Disagreement	(6) Top 50 firms	(7) Top 51-100 firms	(8) Top 150 firms
Avg. Sentiment	0.442	0.425	0.404	0.650	0.442	0.370	0.543	0.482
Avg Disagreement	0.469	0.368	0.382	0.199	0.219	0.722	0.227	0.336
Corr(Dis, Ablogvol)	0.117	0.127	0.143	0.137	0.103	0.114	0.169	0.099

Panel B

Note: In this panel we examine how disagreement within different types of investors change around earnings announcements. We run the following regression:

$$\begin{aligned} AbLogVol_{it} &= \alpha + \beta_1 1WeekBeforeEA_{it} + \beta_2 EA_{it} + \beta_3 1WeekAfterEA_{it} \\ &+ \beta_4 2WeekAfterEA_{it} + \beta_5 3WeekAfterEA_{it} + \gamma Disagreement_{it} \\ &+ \delta_1 Disagreement_{it} \times 1WeekBeforeEA_{it} + \delta_2 Disagreement_{it} \times EA_{it} \\ &+ \delta_3 Disagreement_{it} \times 1WeekAfterEA_{it} + \delta_4 Disagreement_{it} \times 2WeeksAfterEA_{it} \\ &+ \delta_5 Disagreement_{it} \times 3WeeksAfterEA_{it} + SUE_{iq} + TimeFEs + FirmFEs + \varepsilon_{it} \end{aligned}$$

Where $AbLogVol_{it}$ is the abnormal log trading volume on day t for firm i, $1WeekBeforeEA$ is a dummy variable equal to 1 if day t for firm i falls in the week before an earnings announcement for that firm, EA_{it} is a dummy variable equal one if firm i announces earnings on day t, $1WeekAfterEA_{it}$, $2WeekAfterEA_{it}$, $3WeekAfterEA_{it}$ are dummy variables for whether day t for firm i falls in week 1, week 2, or week 3 after an earnings announcement, respectively. $Disagreement_{it}$ is our measure of investor disagreement about stock i in the market on day t. We standardize the disagreement measure by subtracting the mean and dividing by the standard deviation, over the entire sample period. SUE_{iq} is the earnings surprise in quarter q for firm i, defined as the earnings minus the median analyst forecast In this table we present average sentiment, average disagreement measure, and the daily correlation between our disagreement measure and the abnormal log volume for different robustness specifications. Columns (1) and (2) presents results for our main specifications. In column (3) when we construct the sentiment and disagreement measures we weigh each message by the number of followers the author of the message has. In column (4) we only include opinions by investors who joined StockTwits before 1 January, 2013. In column (5) we only use messages that were classified by users themselves as bullish or bearish. In column (6) we use a linear disagreement measure described in the appendix. In columns (7) and (8) we only consider top 50 most talked-about firms. In columns (9) and (10) we only consider top 51-100 most talked-about firms. In columns (11) and (12) we consider the top 150 most talked-about firms. The regressions include year, moth, day-of-the-week, and firm fixed effects. Standard errors are clustered by company and date. *, **, and *** indicate statistical significance at the ten, five, and one percent level respectively. Standard errors are in parentheses.

Abnormal Log Volume

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
hline	Main dataset		Weighted Disagreement	Joined before 1 Jan 2013	User-classified Messages	Linear Disagreeement	Top 50 firms		Top 51-100 firms		Top 150 firms	
1 Week before EA	0.038	0.033	0.034	0.033	0.041*	0.039*	0.042	0.052*	0.038	0.017	0.041*	0.044*
	(0.024)	(0.024)	(0.024)	(0.023)	(0.024)	(0.023)	(0.030)	(0.030)	(0.032)	(0.032)	(0.024)	(0.024)
EA	0.670***	0.499***	0.524***	0.498***	0.492***	0.582***	0.716***	0.622***	0.623***	0.405***	0.677***	0.475***
	(0.050)	(0.049)	(0.051)	(0.049)	(0.049)	(0.048)	(0.062)	(0.060)	(0.074)	(0.073)	(0.049)	(0.056)
1 Week after EA	0.399***	0.314***	0.319***	0.315***	0.305***	0.356***	0.450***	0.395***	0.346***	0.243***	0.401***	0.288***
	(0.032)	(0.028)	(0.027)	(0.028)	(0.027)	(0.028)	(0.046)	(0.040)	(0.040)	(0.035)	(0.032)	(0.027)
2 Weeks after EA	0.117***	0.073***	0.079***	0.083***	0.080***	0.112***	0.113***	0.085***	0.120***	0.063**	0.118***	0.067***
	(0.022)	(0.022)	(0.022)	(0.022)	(0.022)	(0.022)	(0.028)	(0.027)	(0.031)	(0.031)	(0.022)	(0.023)
3 Weeks after EA	0.046**	0.008	0.014	0.020	0.015	0.046**	0.037	0.004	0.051*	0.012	0.048**	0.008
	(0.020)	(0.019)	(0.020)	(0.019)	(0.020)	(0.019)	(0.026)	(0.025)	(0.029)	(0.028)	(0.020)	(0.020)
Disagreement		0.199***	0.174***	0.215***	0.143***	0.118***		0.149***		0.189***		0.193***
		(0.019)	(0.017)	(0.020)	(0.015)	(0.013)		(0.019)		(0.026)		(0.018)
Disagreement × 1 Week before EA		-0.024	-0.014	-0.024	-0.035**	-0.008		-0.035*		-0.049*		-0.026*
		(0.016)	(0.016)	(0.015)	(0.014)	(0.014)		(0.018)		(0.029)		(0.015)
Disagreement × EA		0.129***	0.085**	0.121***	0.097***	0.058*		0.084		0.077*		0.118***
		(0.043)	(0.039)	(0.039)	(0.032)	(0.033)		(0.066)		(0.044)		(0.041)
Disagreement × 1 Week after EA		0.104***	0.105***	0.114***	0.096***	0.087***		0.075**		0.064*		0.097***
		(0.025)	(0.024)	(0.022)	(0.020)	(0.021)		(0.029)		(0.034)		(0.024)
Disagreement × 2 Week after EA		0.020	-0.002	0.013	0.012	-0.018		-0.020		0.031		0.020
		(0.019)	(0.018)	(0.016)	(0.017)	(0.016)		(0.016)		(0.031)		(0.018)
Disagreement × 3 Weeks after EA		0.007	0.009	0.012	-0.019	0.010		0.005		0.004		0.005
		(0.017)	(0.017)	(0.014)	(0.015)	(0.015)		(0.019)		(0.034)		(0.017)
SUE	-0.001	-0.033	-0.032	-0.034	-0.029	-0.012	0.048	0.019	-0.040	-0.081	-0.132*	-0.151**
	(0.056)	(0.055)	(0.055)	(0.055)	(0.054)	(0.055)	(0.089)	(0.090)	(0.068)	(0.064)	(0.074)	(0.068)
Observations	33,111	33,111	33,111	33,111	33,111	33,111	16,811	16,811	16,300	16,300	33,489	33,489
R-squared	0.162	0.200	0.195	0.212	0.192	0.185	0.194	0.222	0.141	0.186	0.164	0.202
Year, month, dow FEs	X	X	X	X	X	X	X	X	X	X	X	X
Firm FEs	X	X	X	X	X	X	X	X	X	X	X	X

Table A.6: Disagreement and Forecasting Trading Volume

Note: This table examines whether our measure of changes in investor disagreement forecast trading volume. We run the following regression:

$$AbLogVol_{i,t} = \alpha + \beta DisMeasure_{i,t} + \gamma AbLogVol_{i,t-1} + TimeFEs + FirmFEs + \varepsilon_{i,t}$$

Where in column (1) $DisMeasure_{it}$ is our disagreement measure for firm i on day $t-1$, in column (2) it's our disagreement measure on day t, and in column (3) it is the disagreement among investors who expressed their opinions before the market opens (BMO) (between 4pm on day $t-1$ and and 9am on day t). We standardize the disagreement measures by subtracting the mean and dividing by the standard deviation, over the entire sample period. $AbLogVol_{it}$ is the difference between log volume in timer period t and the average log volume from $t-140$ to $t-20$ trading days (6-month period, skipping a month) for firm i. Since trading volume tends to be autocorrelated, we also control for abnormal trading volume on day $t-1$. The regressions include year, moth, day-of-the-week, and firm fixed effects. Standard errors are clustered by company and date. *, **, and *** indicate statistical significance at the ten, five, and one percent level respectively. Standard errors are in parenthesis.

	Abnormal Log Volume (t)		
Disagreement measure	(1)	(2)	(3)
Disagreement (t-1)	-0.006		
	(0.004)		
Disagreement (t)		0.110***	
		(0.008)	
Disagreement (BMO, t)			0.054***
			(0.005)
Abnormal Log Volume (t-1)	0.745***	0.726***	0.733***
	(0.013)	(0.014)	(0.014)
Observations	42,322	42,415	42,415
R-squared	0.600	0.608	0.602
Year, month, day of week FEs	X	X	X
Firm FEs	X	X	X

Table A.7: Disagreement and Forecasting Abnormal Stock Returns

Note: In this table we examine whether changes in investor disagreement predict stock returns. We run the following regression:

$$Abret_{it+1} = \alpha + \beta DisMeasure_{it} + \nu AvgSentiment_{it} + \phi Abret_{it} + \gamma AbLogVol_{it} + \delta LogME_{it} + TimeFEs + \varepsilon_{i,t}$$

Where is the disagreement measure on day t for firm i. In column (1) $Abret_{it+1}$ is the abnormal return (minus the value-weighted market index) on day $t+1$ for firm i. In column (2) we put cumulative abnormal returns for days $t+1$ to $t+5$ ($CAR[1,5]$) on the left-hand side. We standardize the disagreement measure by subtracting the mean and dividing by the standard deviation, over the entire sample period. *AvgSentiment* is the average sentiment measure for firm i on day t. *AbLogVol* is the difference between log volume in time period t and the average log volume from $t-140$ to $t-20$ trading days (6-month period, skipping a month). $Log(ME)$ is the log of market capitalization of the firm. The regressions include year, moth, and day-of-the-week fixed effects. Standard errors are clustered by date. *, **, and *** indicate statistical significance at the ten, five, and one percent level respectively. Standard errors are in parenthesis.

	(1)	(2)	(3)	(4)
	AbRet_{t+1}	AbRet_{t+1}	CAR[1,5]	CAR[1,5]
Disagreeement All Investors (t)	-0.0005**	-0.0006**	-0.0011*	-0.0012*
	(0.000)	(0.000)	(0.001)	(0.001)
Avg. Sentiment (t)		-0.0002		-0.0011
		(0.000)		(0.001)
AbRet (t)	0.0472**	0.0473**	0.0317	0.0326
	(0.020)	(0.020)	(0.030)	(0.030)
Abnormal Log Volume (t)	0.0011*	0.0011*	0.0032**	0.0032**
	(0.001)	(0.001)	(0.001)	(0.001)
Observations	42,432	42,432	42,432	42,432
R-squared	0.005	0.005	0.010	0.010
Year, month, day of week FEs	X	X	X	X