The Journal of FINANCE
The Journal of THE AMERICAN FINANCE ASSOCIATION

THE JOURNAL OF FINANCE • VOL. LXXV, NO. 1 • FEBRUARY 2020

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

J. ANTHONY COOKSON and MARINA NIESSNER*

ABSTRACT

We study sources of investor disagreement using sentiment of investors from 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 differential interpretation of information by studying disagreement within and across investment approaches. Overall disagreement is evenly split between both sources of disagreement, but within-group disagreement is more tightly related to trading volume than cross-group disagreement. Although both sources of disagreement are important, our findings suggest that information differences are more important for trading than differences across market approaches.

DISAGREEMENT AMONG INVESTORS HAS LONG been thought to be central to trading in financial markets. Indeed, it is difficult to explain why investors would trade at all without some source of disagreement (Milgrom and Stokey (1982), Karpoff (1986)). Motivated in part by this observation, a growing literature examines the effects of investor disagreement on financial market outcomes

*J. Anthony Cookson is affiliated with University of Colorado at Boulder - Leeds School of Business. Marina Niessner is with AQR Capital Management. We are grateful to Nick Barberis, Joey Engelberg (discussant), James Choi, Diego Garcia, Harrison Hong (discussant), Rawley Heimer, Toby Moskowitz, Tyler Muir, Justin Murfin, Matt Spiegel, Johannes Stroebel (discussant), Paul Tetlock (discussant), Heather Tookes, Martin Weber (discussant), Paul Irvine (discussant), and Scott Yonker (discussant) 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, 2016 National Bureau of Economic Research (NBER) Spring Behavioral Finance Meeting, 2016 NBER Summer Institute Asset Pricing, 2016 IDC Summer Conference, 6th Helsinki Finance Summit on Investor Behavior, 2016 SITE New Models in Financial Markets Session, 2016 CMU Summer Symposium, Michigan State University, University of Washington, MIT (Accounting), Boston College, American Finance Association 2017, Jackson Hole Finance Conference 2017, European Winter Finance Conference 2017, AQR Capital Management, Finance Down under Finance Conference, Front Range Finance Seminar 2017, Harvard Business School (NOM), Harvard Business School (Finance), The European Summer Symposium in Financial Markets 2017, and the Red Rock Finance Conference 2017. AQR Capital Management is a global investment management firm, which may or may not apply similar investment techniques or methods of analysis as described herein. The views expressed here are those of the authors and not necessarily those of AQR. We have read *The Journal of Finance*'s disclosure policy and have no conflicts of interest to disclose.

DOI: 10.1111/jofi.12852

(e.g., Varian (1985), Harris and Raviv (1993), Kandel and Pearson (1995), Nagel (2005), Banerjee and Kremer (2010), Carlin, Longstaff, and Matoba (2014)). Prior research has linked disagreement to trading volume and stock returns, and has studied its dynamic effects (Ajinkya, Atiase, and Gift (1991), Diether, Malloy, and Scherbina (2002), Banerjee and Kremer (2010)).

However, despite abundant evidence 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 identify two main sources of disagreement—differences in information sets and differences in the models that investors use to interpret information (Hong and Stein (2007)). To examine these questions empirically, we study disagreement among investors on the social media investing platform StockTwits, where users regularly express their opinions (e.g., bullish or bearish) about stocks and where user profile information *explicitly* conveys the user's broad investment approach (e.g., fundamental, technical). Using this setting, we provide novel insights into the relative importance of different information sets versus different investment models.[1]

Separating the roles of different information sets and different models in determining investor disagreement is empirically challenging, given the typical data limitations. First, disagreement corresponds 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—such as liquidity. Second, as Rothschild and Sethi (2016) and Baron et al. (2019) point out, to determine whether differences in investor opinions are due to differences in information sets or differences in investors' models, researchers would ideally observe investors' trading strategies—not just their executed trades.

Our data set enables us to empirically distinguish between information-driven and model-driven sources of disagreement because, as we will show, disagreement across investment approaches is more likely to arise due to differing investment models, whereas disagreement within investment approach disagreement is more likely due to different information sets. We find that differences of opinion across the broad investment approaches in our data are responsible for approximately half of overall disagreement. At the same time, within-group differences of opinion are much more strongly related to trading

[1] Specifically, Hong and Stein (1999) posit that gradual information diffusion is an important source of disagreement that can drive trading decisions. More recently, Chang et al. (2014) provide evidence that different information sources lead to a divergence of opinion and greater trading volume. On the other hand, differential interpretation of information is central to the models of Harris and Raviv (1993) and Kandel and Pearson (1995). Kandel and Pearson (1995) provide evidence of differential interpretation by financial analysts and argue that this differential interpretation leads to greater trading volume after public announcements of information (earnings announcements). A central aim of our paper is to use our decomposition of overall disagreement to speak to the relative weight of these two theories of trading.

volume than are differences of opinion across groups, suggesting that model disagreement is less likely to induce trading than different information sets.

Given that these investment philosophies are self-reported, we carefully check that adherence to an investment philosophy on StockTwits reflects adherence to an investment model in reality. We first analyze the textual content of tweets by users of different philosophies. We find that users of different philosophies use language that is consistent with the underlying philosophy (e.g., fundamental traders discuss earnings, technical traders discuss charts, and momentum traders discuss trends). Next, speaking to the external validity of the language used, we find that the language used on the StockTwits platform closely resembles public writings of prominent investors with particular investment philosophies. Furthermore, using hand-classified lists of information words (i.e., referring to news sources or timing) and model words (i.e., referring to substantive analyses), we find that information words tend to be used across investment philosophies, while model words tend to focus on one or two investment philosophies. Beyond language usage, we show that investor sentiment reactions to earnings news concentrate among fundamental investors, while sentiment reactions to "technical view" events identified by the news analytics database RavenPack concentrate among technical investors.[2] Taken together, these findings support the view that the differences across investment philosophies are significant, substantive, and a function of differential beliefs about investing.

Turning to our main findings, we observe that both within-group and cross-group disagreement significantly predict abnormal trading volume, but that within-group disagreement exhibits a much stronger relation to trading volume. Specifically, we find that a one-standard-deviation increase in within-group disagreement is associated with *four times* the increase in abnormal trading volume as a one-standard-deviation increase in cross-group disagreement. This finding is robust to alternative specifications for the differences between within-group and cross-group disagreement. Moreover, we continue to find a similarly large effect on within-group disagreement when we restrict attention to opinions from before the market opens. We therefore conclude that both types of disagreement are important determinants of trading, but that within-group (informational) differences matter more than differences in investment philosophies. This result suggests that disagreement due to slow information diffusion is important for trading volume.

We provide two additional pieces of evidence on the slow information diffusion hypothesis using self-reported experience classifications to split our sample of investors into sophisticated and unsophisticated investors. First, we find that within-strategy disagreement across sophisticated and unsophisticated investors predicts trading volume. This result suggests that information diffuses from sophisticated investors to unsophisticated slowly over time, consistent

[2] This test is analogous to the work of Jia, Wang, and Xiong (2015) who show that local and foreign investors react differently to recommendations of local and foreign analysts in the context of the Chinese stock market.

with slow information diffusion. Second, building on this result, we show that sophisticated investor sentiment leads unsophisticated investor sentiment in time, not the other way around, consistent with information diffusing from sophisticated investors to less sophisticated investors over time.

These findings are robust to a wide array of measurement choices and controls. Notably, we show that our disagreement measure is distinct from other factors that influence trading volume, such as attention or news articles about the firm. We analyze the joint effect of investor disagreement and investor attention on trading volume, where we proxy for investor attention using the number of daily messages on StockTwits and the number of daily searches for companies' tickers on Google (e.g., Da, Engelberg, and Gao (2011), Niessner (2016)). We find that both investor disagreement and investor attention are strongly associated with greater trading volume, and that the relation between disagreement and trading volume is robust to granular controls for the number of messages posted. We also control for the presence of media articles in our analysis, and find that it does not change the effect of disagreement on volume.

We next examine how our disagreement measures relate to other disagreement proxies used in prior literature. We find a relatively weak correlation with previous proxies for disagreement, which indicates that our disagreement measures capture different aspects of disagreement relative to 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 (e.g., Diether, Malloy, and Scherbina (2002), Giannini, Irvine, and Shu (2018)).[3] Given that the aim the literature is to explain high trading volume at the daily level (e.g., see Hong and Stein (2007)), this is an important distinction.

In addition to providing broad evidence on cross-group and within-group disagreement, we use our measure to study a classical setting of excess trading volume following earnings announcements (e.g., Kandel and Pearson (1995)). This application highlights the advantage of having a daily measure of disagreement, which contrasts sharply with analyst forecasts that are updated less frequently. We show that daily changes in disagreement can explain up to a third of the increase in trading volume after earnings announcements, which supports the view that disagreement across different models drives a significant amount of daily trading volume. Digging deeper into reactions to earnings news, we find evidence consistent with differential interpretations of earnings announcement news. In particular, we find that all investor types post more messages after earnings announcements, but the sentiment of their reactions is in line with investors interpreting new information differently in a manner

[3] Furthermore, in a recent paper Cen, Wei, and Yang (2016) show that the earnings analyst forecast dispersion measure captures not only disagreement but also other return-predictive information contained in the normalization scalars of the measure.

that is consistent with their different investment philosophies. These findings provide additional support for our result that model-driven disagreement is an important source of overall disagreement in the market, as well as fresh empirical evidence for emerging theories on why disagreement rises precisely when information arrives in the market (Kondor (2012), Banerjee, Davis, and Gondhi (2018)).

Our results, disagreement measure, and approach should be of broad interest to scholars studying individual investing behavior and market microstructure. First, although there has been significant research on the consequences of disagreement for financial market outcomes, our paper is one of the first to empirically study the sources of disagreement. The broader literature recognizes that significant differences in early-life experiences and in genetic predisposition to risk can affect how individuals approach financial markets (Malmendier and Nagel (2011), Cronqvist, Siegel, and Yu (2015), Cronqvist et al. (2016), Brown, Cookson, and Heimer (2019)). Our work builds on these underlying differences among investors to better understand how disagreement influences aggregate trading outcomes. In this respect, our work is most closely related to two contemporaneous papers on sources of disagreement that focus on differential exposure to information (Chang et al. (2014), Bailey et al. (2017)). Bailey et al. (2017) examine how differential exposure to friends' real estate experiences influences optimism about real estate investing, and show that differences in these social network experiences lead to disagreement. In a similar vein, Chang et al. (2014) exploit differences in exposure to ideas and find that linguistic diversity is a source of divergence of opinion because agreement is more difficult in the face of communication barriers. Viewed broadly, these studies show empirically that features of the information environment lead to differential information, and, in turn, to investor disagreement. In relation to these findings, we are the first to provide direct evidence of model disagreement among investors. Thus, in contrast to other sources of disagreement, disagreement in our setting would exist even if investors had the same information.

Second, our work provides a useful perspective on the large theoretical literature on disagreement. Heterogeneous investor beliefs play a central role in explaining speculative bubbles (e.g., Chen, Hong, and Stein (2002), Scheinkman and Xiong (2003)). For example, it is difficult to otherwise explain the high levels of trading volume in financial markets (Varian (1989), Harris and Raviv (1993), Kandel and Pearson (1995)). The models in these studies suggest that investors with different information-processing models interpret public information differently. Furthermore, a different set of models focuses on the consequences of gradual information diffusion and investor inattention and their effects on trading volume and prices (e.g., Hong and Stein (1999), Hirshleifer and Teoh (2003), Peng and Xiong (2006)). An excellent survey of this literature can be found in Xiong (2013). Our characterization of model disagreement relates most closely to Kandel and Pearson (1995), who provide indirect evidence of differential interpretation of information by stock market analysts in response to earnings announcements. Relative to this work, our approach is novel in that it delivers an explicit measure of differences of opinion across investment models that can be related to trading volume at a daily level. As we show in our decomposition of trading volume effects, this measure allows us to compare the importance of within-group disagreement and cross-group disagreement. Our finding that within-group differences of opinion matter more for trading than differences across groups are novel evidence on the relative importance of each of these mechanisms. Because gradual information diffusion theories and differential interpretation theories each have empirical support (Kandel and Pearson (1995), Hong and Stein (1999)), this is an important result.[4]

We also contribute to the empirical disagreement literature by providing a useful measure of disagreement among individual investors. Although the consequences of disagreement are well studied, existing measures of disagreement have notable weaknesses. For example, some of these measures proxy for dispersion of opinion indirectly (e.g., volatility of accounting performance, historical trading volume, firm age, and return volatility), and the most prominent measure of analyst forecast dispersion is based on analysts' stated opinions and may not be a reliable measure of market-wide disagreement (Ataise and Bamber (1994), Bamber, Barron, and Stevens (2011). We address these issues 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). In short, our disagreement measure can be computed at a higher frequency than most other measures of disagreement and, because it is a direct measure of sentiment, it is less likely to proxy for other market forces that are unrelated to disagreement, such as investors' liquidity needs.

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 overtrading, including entertainment (Dorn and Sengmueller (2009)), sensation seeking (Barber and Odean (2008), Grinblatt and Keloharju (2009)), gambling (Kumar (2009), Cookson (2018)), and learning by doing (Linnainmaa (2011)). We contribute to this stream of research by providing clean evidence that model disagreement is an additional determinant of 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 trading.

[4] The literature recognizes that information differences need to interact with some other forms of heterogeneity, like heterogeneous beliefs, to generate trading (see the review by Xiong (2013)). Thus, both informational differences and model differences are necessary to generate trades, but the extent to which how much each source of disagreement matters for trading is an open question.

I. Data

A. 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 a Twitter-like format: participants can post messages of up to 140 characters and can use "cashtags" with the stock ticker symbol (example $AAPL) to link a user's message to a particular company. According to the website analytics tool Alexa, StockTwits was ranked the 2,004th most popular website in the United States as of May 2015. The users are predominantly male, and the number of users with a graduate degree is overrepresented relative to other websites.

StockTwits provided data on the universe of messages posted between January 1, 2010 and September 30, 2014. In total, we have 18,308,948 messages by 107,808 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.

We restrict our sample to messages posted between January 2013 and September 2014 because the best coverage and highest quality data come from more recent years. The resulting sample retains 75% of the messages in the StockTwits data. We also restrict attention to messages that mention only one ticker to focus on sentiment that can be directly linked to a particular stock. Because it will be useful for our decomposition of disagreement into different types, we retain StockTwits messages by users who report their investment approach, holding period, and experience in their profile information. We focus on firms that are headquartered in the United States, and thus, have regular filings with the Securities and Exchange Commission to facilitate linking the data with earnings announcement information. Finally, because daily observations of investors' opinions about individual firms are ideal for constructing a daily measure of disagreement, 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. After these sampling restrictions, our final sample contains 1,442,051 messages by 12,029 unique users.[5] We present the names of the 100 firms and the frequency of messages

[5] In the Internet Appendix (Table IA.IX), which is available on *The Journal of Finance* website, we conduct a number of robustness exercises to ensure that our findings are not sensitive to our sampling choices. First, we reproduce our main findings analyzing the top 150 firms, the top 50 firms, and the top 51-100 firms, obtaining similar results. These results indicate that our findings are not driven by the top stocks, nor are they sensitive to the 100-firm cutoff. Second, we evaluate how StockTwits user growth affects our findings. Our main tests control for the growing nature of our sample by including date fixed effects in our analysis (out of 11,876 users, 4,566 joined before January 1, 2013). To ensure that the potentially changing composition of investors is not affecting our results, we repeat the analysis using just those users who joined StockTwits before January 1, 2013. We again obtain similar results. Third, we see no sharp changes to StockTwits message volume over time (i.e., Internet Appendix Figure IA.1), which provides a reliable basis for measuring disagreement.

Table I
Characteristics of StockTwits Data

In this table, we report summary statistics for the StockTwits data. Panel A presents summary information on 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 A: Characteristics of Messages and Users

	Mean	*SD*	Min	p25	p50	p75	Max
Number of messages per stock	14,420	32,493	616	1,589	5,296	14,686	275,969
Number of messages per user	119	391	1	5	19	80	11,759
Number of messages per stock per day	43	134	1	3	10	31	4,690
Sentiment stock/day	0.439	0.518	−1	0.167	0.5	1	1
Number of followers user has	212	2,126	0	2	6	21	96,433
Number of people user follows	45	197	0	5	15	45	9,990
Total Days Active	457	411	1	131	343	679	1,908

Panel B: Frequencies of User Profile Characteristics

Investment Philosophy	Number of Users	Percent Users	Number of Messages	Percent Messages
Fundamental	1,505	12.51%	203,383	14.10%
Technical	4,610	38.32%	538,425	37.02%
Momentum	2,395	19.91%	368,939	26.12%
Global Macro	269	2.24%	12,974	0.90%
Growth	2,158	17.94%	217,504	15.08%
Value	1,092	9.08%	100,826	6.99 %
Total	12,029	100%	1,442,051	100%
Holding Period	**Number of Users**	**Percent Users**	**Number of Messages**	**Percent Messages**
Day Trader	1,872	15.56%	267,896	18.58%
Swing Trader	5,313	44.17%	660,898	45.83%
Position Trader	2,668	22.29%	288,238	19.99%
Long Term Investor	2,176	18.09%	225,019	15.60%
Total	12,029	100%	1,442,051	100%
Experience	**Number of Users**	**Percent Users**	**Number of Messages**	**Percent Messages**
Novice	3,392	28.20%	228,041	15.81%
Intermediate	6,272	52.14%	803,198	55.70%
Professional	2,365	19.66%	410,812	28.49%
Total	12,029	100%	1,442,051	100%

about these firms in the Internet Appendix (Table IA.II). Not surprisingly, many of the firms that are discussed most are in the technology and pharmaceutical industries.

Table I, Panel A, presents summary statistics on the sample coverage. The median number of messages per firm-date observation is 10, with as many as 4,690 messages for some firms on some days. Since the typical firm has multiple messages per trading day in the data, we are able to calculate measures of disagreement at the firm-date level.

B. Investor Philosophies

StockTwits users can fill out user profiles with information about themselves as investors—investment approach, investment horizon (or holding period), and experience level. In Panel B of Table I, we present the breakdown of users by investment approach, holding period, and experience. Our analysis makes use of these user characteristics, particularly the investment approach information, which describe quite different investing philosophies. On StockTwits, the most common approach is technical, representing 38% of users and about 37% 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. Given that global macro investors make up only 2.24% of overall investors and 0.90% of messages, we exclude these investors from the analysis below. To the best of our knowledge, ours is the first paper to directly measure investors' approaches, and as a result, we cannot assess whether this breakdown is representative of other samples in the market. Because there are no natural comparisons to ours, in Section IV.B, we present several alternative weighting schemes to evaluate the external validity of our results.

To examine whether the StockTwits investment approaches reliably categorize users into truly different investment philosophies, we systematically analyze the content of the messages posted by users across approaches. The content of the messages posted by users of different approaches suggests that users adhere to the investment approach they select when they register (i.e., fundamental, technical, momentum, growth, or value). Specifically, we study word saliency to focus on words that are distinctive of each approach as in Goldsmith-Pinkham, Hirtle, and Lucca (2016). To highlight the differences across strategies, Panel A of Table II presents the 15 most salient words for each strategy, that is, the words most frequently used relative to language used by investors of other approaches. These salient words give a sense of the contextual differences between strategies. For example, some of the most salient words for fundamental investors are "eps" (earnings per share) and "cash," whereas technical investors refer to "charts," "area," and "head."

While these most salient words suggest that investors follow the investment approaches that they self-report on their profile, one may be concerned that these words happen to be salient by chance or do not represent the full spectrum of word usage. We mitigate this potential concern with evidence that the full distribution of word usage across strategies is also distinctive. We compute the word frequency distribution using all of the tweets by users of each investment philosophy, which provides a basis for this comparison. To capture the distinctiveness of the language used across different investment philosophies, we borrow a measure from information theory, namely, the Kullback-Liebler divergence (KL divergence), which captures the degree to which the

Table II

Textual Validation of User Approaches

This table presents three textual validation exercises for the self-defined user approaches. Panel A presents the 15 most salient words by investment approach, which is a technique to parse the useful content of a source text (as in Goldsmith-Pinkham, Hirtle, and Lucca (2016)). Panel B presents Kullback-Liebler divergence calculations for each strategy relative to fundamental. Standard errors are computed by drawing 100 bootstrap samples from the fundamental word distribution, recomputing the Kullback-Liebler divergence, and computing the sample standard deviation of Kullback-Liebler divergence across bootstrap samples. Panel C presents Kullback-Liebler divergence calculations relative to two focal investors who write extensively about investments outside of StockTwits—Gregory Harmon (technical) and Todd Sullivan (value). For ease of comparison, the Kullback-Liebler divergence calculations are grouped by trends-oriented approaches (technical and momentum) versus more fundamentally oriented approaches (growth, value, and fundamental).

Panel A: Most Salient Words Used by Approach

Approach	Most Common Unique Words
Fundamental	eps, sales, growth, sentiment, read, revenue, earnings, million, quarter, consensus, billion, share, cash, results, analysts
Technical	chart, support, nice, break, looking, looks, gap, move, day, stop, calls, daily, close, resistance, bounce
Momentum	play, calls, time, via, week, day, news, squeeze, hod (high of day), hit, shares, cover, highs, run, money
Value	view, attempts, bulls, rising, aboard, stair, intraday, correction overextended, breakdown, fresh, mayb, steak, moved, rollout
Growth	news, er (earnings report), hope, green, shares, plug, money, article, time, bears, waitings, ve, wait, board, share, future

Panel B: Kullback-Liebler Divergences of Word Distributions by Approach
Fundamental as the Baseline Approach

	Growth	Momentum	Technical	Value
Divergence from Fundamental	0.0854	0.1146	0.1919	0.2336
Standard Error (100 bootstrap replications)	0.00008	0.00009	0.00009	0.00008

Panel C: Divergence from Writing by Focal Investors

	Fund, Value, and Growth	Technical and Momentum	Difference	Std. Err.	t-Stat
Technical Focal Investor	1.186	1.113	0.073	0.028	2.624
Value Focal Investor	1.284	1.512	−0.228	0.039	−5.871

information contained in one distribution differs from the information in another. In our context, the more different are two word distributions, the greater is their KL divergence.

Using the word distribution of tweets from each approach, we compute the pairwise KL divergences for each pair of word frequency distributions. Panel B of Table II presents the KL divergence of each strategy relative to the fundamental investor distribution, along with standard errors, obtained using a bootstrapping procedure.[6] By these calculations, the word distributions of growth, value, momentum, and technical investors are highly statistically different from the word distribution used by fundamental investors. The tight standard error bounds also indicate that each strategy uses language that is quite different from that of the others.

In addition to showing that the language used across strategies is distinctive, we show that the language of the messages on StockTwits is consistent with investment views of identifiable focal investors who write extensively outside of StockTwits. To do so, we identify two focal investors (one technical investor and one value investor) who have significant public writings and investor reputations outside of the StockTwits platform. For the technical focal investor, we select Gregory W. Harmon (133K followers on StockTwits as of May 2017), who maintains a blog and subscription newsletter that espouses his technical trading views (e.g., the subtitle of his website is "Do you see what I see? What the charts are telling about the current state of the market"). The other focal investor we select is Todd Sullivan (98K followers), who is decidedly not a technical investor. Sullivan writes publicly about investing on the blog "Value Plays." For both focal investors, we download the universe of their public posts on their respective websites from May 2017, and we process the text in the same way we process the StockTwits text.

Using these reference texts, we test whether StockTwits users with more technically oriented approaches (i.e., momentum and technical) use language that is closer to the technical focal investor, Harmon, than StockTwits users with more fundamentally oriented approaches (i.e., value, growth, and fundamental), and vice versa for the more fundamentally oriented focal investor, Sullivan. To do so, we group StockTwits messages by whether they were posted by a user who was technically oriented versus fundamentally oriented. For each grouping, we compute the KL divergence between the StockTwits messages and the external reference texts.

As we show in Panel C of Table II, the KL divergence from the technical reference text is significantly smaller for the technically oriented StockTwits group than for the fundamentally oriented StockTwits group. The magnitude of the difference (0.073) is similar to cross-group differences on StockTwits, and this difference is significant at the 1% level (t-statistic $= 2.624$). Consistent with the self-identified strategies, we also find that the KL divergence from the fundamentally oriented reference text is significantly smaller for the fundamentally oriented StockTwits group than for the technically oriented StockTwits group. The magnitude of the difference (0.228) is similar to cross-group differences on

[6] Specifically, we randomly draw a bootstrap sample of words from the reference distribution (with replacement and with the same number of words as the original). Using each bootstrapped reference distribution, we recompute the KL divergence. Across 100 bootstrap samples, the standard deviation of the bootstrapped KL divergence calculations is the standard error reported in Table II.

StockTwits, and this difference is statistically significant at better than the 1% level (t-statistic $= -5.871$).[7]

As a final piece of textual evidence on StockTwits investment philosophies, we investigate whether the textual differences across strategies reflect different information sources or differential interpretation of market information. To address this question, we hand-classify the top 1,000 most frequently used words across strategies into three categories—information words, model words, and unclassified words. Information words are words that describe the timing, source, or direction of information (e.g., "positive," "today," "yesterday," "news," and "cnbc"), whereas model words are words that imply a particular approach or analysis of market information (e.g., "sma" [simple moving average], "pattern," "reversal," "upgrades," "squeeze," "ichan," and "director"). We report these word lists in Panel A of Table III; note that we selected the clearest examples to be included in these lists, leaving the remaining words unclassified.[8]

Using these lists of model and information words, we examine the degree to which each type of words is commonly used (i.e., in the top 250 words) by each of the five StockTwits investment philosophies. Panel B of Table III presents the frequency distribution of the number of investment philosophies that commonly use information words versus model words. Consistent with the differences across strategies reflecting different investment philosophies (i.e., different analyses of the same information environment), we find that model words tend to be commonly used by fewer StockTwits investment philosophies than information words. In contrast, information words tend to be used frequently, regardless of the strategy. Specifically, the mean number of strategies that commonly use model words is 3.351, whereas the mean number of strategies that use information words is 4.792. As the two-sample t-test in Panel C of Table III indicates, this difference in means is statistically significant at the 1% level.

[7] To obtain standard errors for these difference-in-divergence tests, we draw 500 random bootstrap samples of words with replacement from the reference distribution (each with replacement and with the same number of words as the original distribution). For each bootstrapped reference distribution, we recompute both KL divergences and their difference (i.e., the KL divergence for fundamentally oriented approaches minus KL divergence for technically oriented approaches). Across the 500 bootstrap samples, the standard deviation of the differences provides the standard error we use for the test of differences in KL divergences.

[8] Selecting words for the word list is an inherently subjective exercise. The type of classification error we are most likely to make is to classify a word as a "model word" when it truly belongs in the "information word" list. Because model words are about interpretation of information, they will necessarily refer to some types of information, even when they imply a particular approach toward processing information. Especially for fundamental strategies that discuss topics such as "revenue," "sales," "profit," and "acquisitions" as their mode of analysis, it is difficult to draw the line between information and interpretation of information. With this in mind, we select 16 words from our original list of model words that are most likely to be confused for information words. Using this list, we rerun our tests with those words included among the information word list instead of the model word list. We obtain quantitatively similar results (see Internet Appendix Table IA.VI), which helps alleviate concerns that our conclusions are sensitive to the subjective exercise of compiling these word lists.

Table III

Model and Information Words

In this table, we examine whether model and information words that we identify are either more related to differences across models or differences in information. In Panel A, we hand-classified the top 1,000 most frequently used words across strategies into three categories—information words, model words, and unclassified words. We display the information and model words. Information words are words that describe the timing, source or direction of information (e.g., "positive," "today," "yesterday," "news," and "cnbc"), whereas model words are words that describe a particular approach or analysis of market information (e.g., "sma" (simple moving average), "pattern," "reversal," "upgrades," "squeeze," "ichan," and "director"). In Panel B, we present the frequency distribution of the number of strategies that commonly use information words versus model words. A word is commonly used by a strategy if it is one of the 250 most commonly used words among StockTwits users who adhere to that strategy. In Panel C, we present a two-sample t-test for the difference in the mean number of strategies—out of technical, momentum, fundamental, value, and growth—which commonly use information words versus model words. *** indicates statistical significance at 1% level.

Panel A: Words That Are Related to Models or Information

Information Words			Model Words				
yesterday	month	beat	value	growth	dip	set	stop
weeks	money	cnbc	trend	gap	analyst	setup	profit
weekly	monday	volume	transcript	fill	analysts	stage	profits
week	lower	real	top	rsi	director	reversal	pop
watch	looks	positive	bottom	macd	candle	reports	low
watching	looking	position	test	bollinger	charts	pullback	earnings
waiting	look	options	swing	crossover	business	pattern	drop
wait	friday	news	support	director	dma	moving	er
tomorrow	day	move	statement	deal	corporation	line	worth
time	days	morning	squeeze	cover	consensus	icahn	wall
term	daily	months	sales	chart	conference	guidance	street
strong	coming	bears	revenue	ceo	double	min	upgrades
stocks	close	bearish	resistance	breakout	data	acquisition	sma
start	call	bad	report	breaking	expected	levels	trigger
soon	calls	article	pattern	break	expect	level	
share	bulls	added	ownership	bounce	fast	key	
ready	bullish	add	near	bottom	flag	hod	
quarter	bull	current	ma	volume	eps	highs	

Panel B: Distributions of Model and Information Words across Investment Philosophies

Distribution of Model Words across Investment Philosophies					
Number of approaches that commonly use the word	1	2	3	4	5
Count of words	19	9	7	10	32

Distribution of Information Words across Investment Philosophies					
Number of approaches that commonly use the word	1	2	3	4	5
Count of words	1	0	2	3	47

Panel C: Test for Differential Use of Information versus Model Words across StockTwits Strategies

	Mean # of Strategies That Commonly Use		
	Information Words	Model Words	Difference
Estimate	4.792	3.351	1.441***
t-Statistic			(6.763)

Beyond the user language, we also evaluate whether investor sentiment reactions to different types of events are consistent with their self-ascribed investment philosophies. We use the news analytics database RavenPack to identify two types of events that should differentially associate with the opinions of users with different investment philosophies: technical view events (labeled as such in RavenPack) and earnings announcements. Specifically, the RavenPack data provide both positive and negative events: "technical view bullish," "technical view bearish," "earnings up," and "earnings down." When we compute sentiment around earnings announcements, we pool the reactions to "earnings down" events with the reactions to "earnings up" events by multiplying the sentiment around "earnings down" events by -1. We similarly multiply "technical view bearish" by -1 when constructing the reaction to the signed technical events. These transformations ensure that the sentiment reaction to negative events does not cancel the sentiment reaction to positive events in the aggregate. Figure 1 plots the sentiment of investors of different types around these two types of news. Consistent with investors having a technical investing philosophy, technical sentiment is high throughout the entire nine-day window around technical view events, whereas nontechnical sentiment is similar to outside of the event window. On the other end of the spectrum, fundamental sentiment reacts more strongly to earnings announcements news than nonfundamental sentiment.

Taken together, the sentiment reactions to fundamental and technical events and the contextual analysis of the contents of StockTwits tweets provide compelling evidence that self-ascribed investment philosophies accurately capture substantively different investor types. Moreover, these investment philosophies use language that is consistent with differential interpretation of the same information, and exhibit sentiment reactions to events that are consistent with these StockTwits investors primarily having different investment models across investment philosophies.

C. Other Useful Features of the StockTwits Data

Two other features of the StockTwits data are useful in our tests. First, we observe the precise time stamp when a message is posted, which is informative as to whether investors post messages as they update their beliefs when news occurs, versus in the evenings after work, when they have more free time. In Figure 2, we plot the distribution of messages by the day of the week and by the hour of the day. As can be seen, investors tend to post messages when the markets are open (Monday-Friday, between 9 a.m. and 4 p.m.). This timing is consistent with investors updating their messages in real time as financial events unfold. In our empirical tests, we use the message time stamps to evaluate the degree to which sentiment and disagreement changes before trading volume does.

Second, we observe investors' self-reported experience level: novice, intermediate, and professional. About 20% of StockTwits users classify themselves as professionals, 52% as intermediate, 28% as novices. Consistent with likely

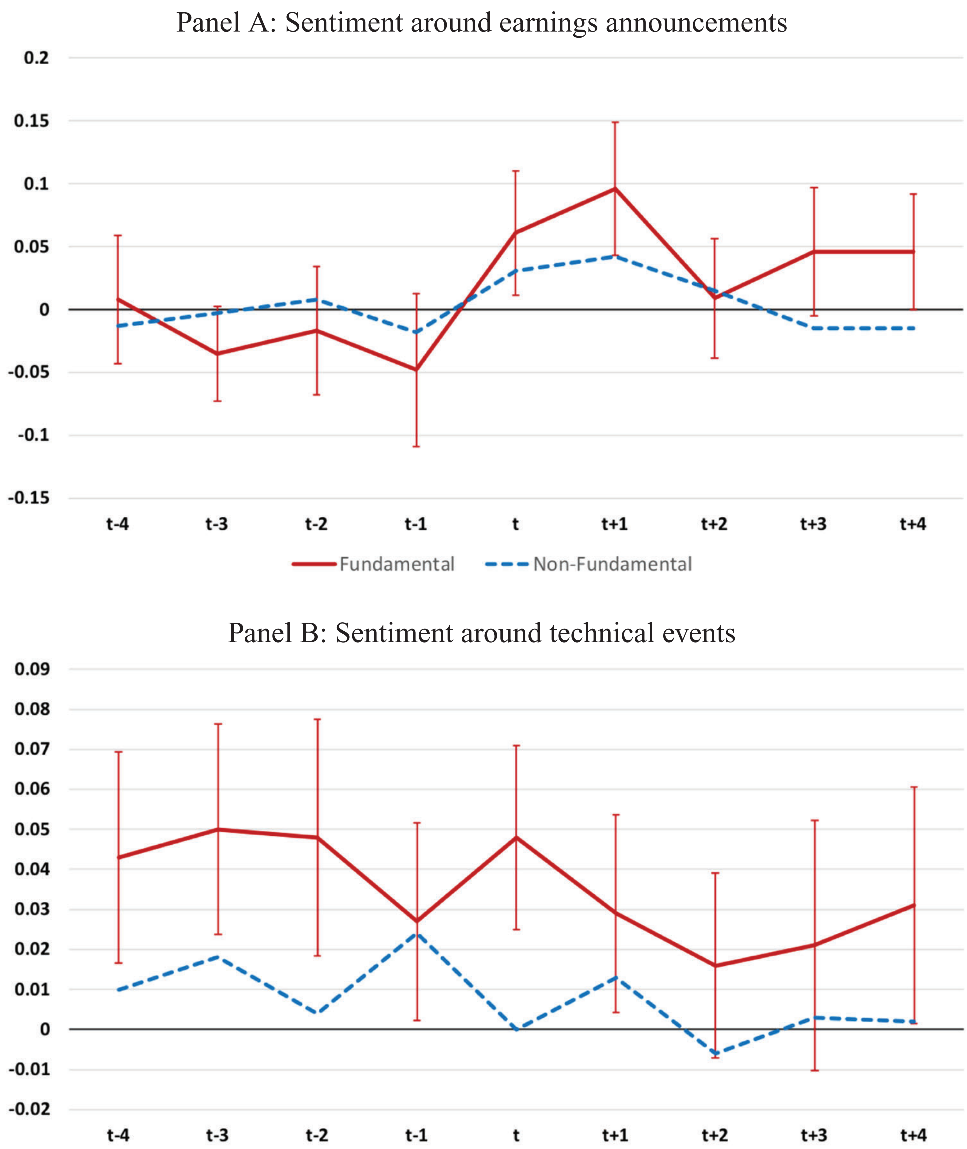


Figure 1. Sentiment around fundamental and technical events. This figure depicts the average sentiment around earnings announcements and technical events as defined by media analytics provider RavenPack. In Panel A, we plot the average sentiment by fundamental and nonfundamental investors around events that RavenPack designates as "earnings-up" and "earnings-down." "Earnings-up" events are defined as follows: "The Company announces an increase in financial earnings results for the period." "Earnings-down" events are defined similarly: "The Company announces a decrease in financial earnings results for the period." The bars are 1.645*standard errors of the sentiment measure. In Panel B, we plot the average sentiment by technical and nontechnical investors around events that RavenPack designates as "technical-bullish" and "technical-bearish." The "technical-bullish" events are defined as follows: "Technical analysis indicates the Entity's

trading behaviors, professionals post disproportionately more messages than novices or intermediates. In our empirical tests, we use the experience classifications to distinguish among investors with different levels of sophistication. As with the self-reported investor philosophies, it is important to check that the experience levels approximate actual investor experience. We provide both contextual validation, by reading profiles, and quantitative validation, by tracking the abnormal market performance of mimicking portfolios.

Contextually, Figure 3 presents three examples of user profiles, one for each experience level in the data, to give a sense of this comparison. Consistent with our reading of many user profiles, the self-reported experience level well aligns with likely real-world investing experience. For example, the *novice* investor is a student, who is trading mostly for fun, the *intermediate* investor reports real-life trading experience but does not trade as a primary source of income, and the *professional* investor has worked for Wells Fargo and Morgan Stanley as well as on floors of COMEX, CSCE, and NYFE.

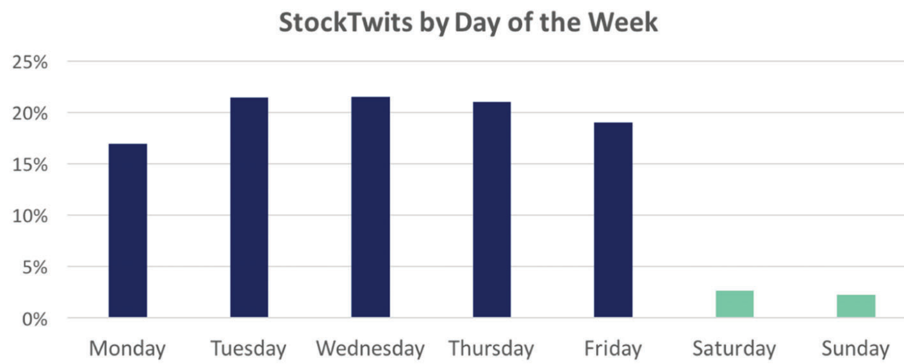
We further validate the experience levels by constructing mimicking portfolios that separately follow professional, intermediate, and novice opinions. Specifically, for each experience level, we evaluate the performance of two portfolios based on the sentiment of StockTwits users: a bullish portfolio and a bearish portfolio. Within each group, we use the bullish or bearish message frequencies as portfolio weights for each portfolio.[9] Figure 4 graphs the cumulative abnormal returns for bullish and bearish portfolios for the overall sample, novices, intermediates, and professionals. Consistent with self-reported experience being valuable, Panel D shows that the professional portfolio exhibits positive cumulative abnormal returns (bullish outperforms bearish over a 60-trading-day period, t-statistic $= 1.55$), whereas the novice portfolio exhibits negative abnormal returns (bearish outperforms bullish over a 60-trading-day period, t-statistic $= -2.11$).[10] Forming a long-short portfolio, the difference in 60-day performance between novices and professionals is highly statistically significant, with a t-statistic of 3.844, a finding that implies substantial

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

[10] The novice portfolio findings are in line with prior research, showing that individual investors lose money in the market, even before accounting for transaction costs (Barber and Odean (2000))

price will appreciate or gain value." "Technical-bearish" events are defined as "Technical analysis indicates the Entity's price will depreciate or lose value." We look only at events that did not have any other technical bullish and technical bearish events in the proceeding three days and the following three days. The bars are 1.645*standard errors of the sentiment measure. Since we have both positive and negative events, we multiply the sentiment around negative events by −1 to ensure that the different types of events do not cancel each other out. (Color figure can be viewed at wileyonlinelibrary.com)

Panel A: Day-of-week frequency distribution of messages posted to StockTwits



Panel B: Hour-of-day frequency distribution of messages posted to StockTwits

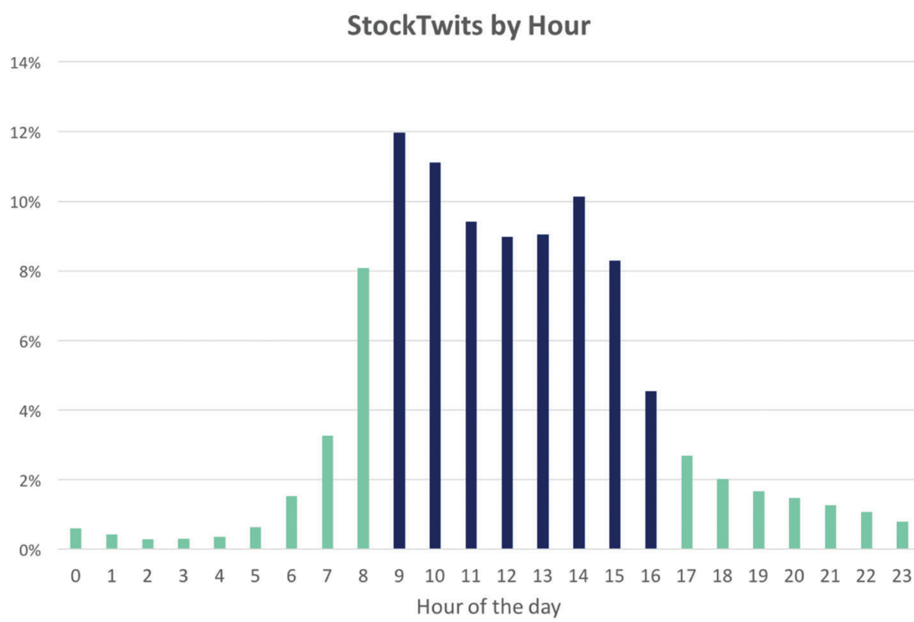
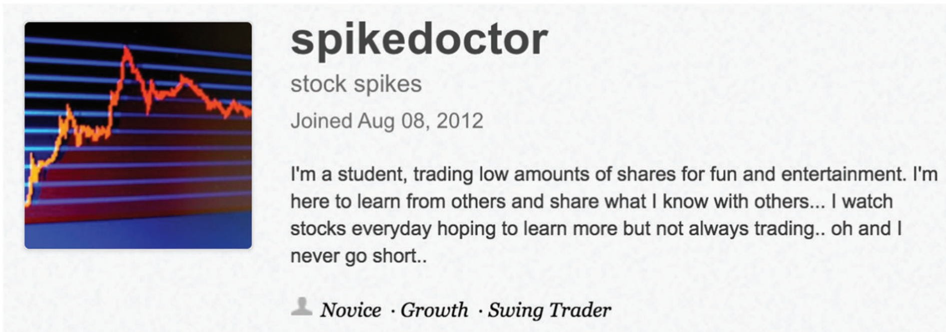


Figure 2. Timing of messages posted. This figure presents a frequency distribution of messages posted by day of the week (Panel A) and hour of the day (Eastern Standard Time; Panel B) that messages are posted to StockTwits. Trading hours are plotted as dark bars and nontrading hours are plotted as light bars. (Color figure can be viewed at wileyonlinelibrary.com)

differences between professionals and novices. The differences between StockTwits professionals and intermediates are similarly statistically significant, with a t-statistic of 3.10. Moreover, these differences are economically significant, with the professional portfolio outperforming the market by nearly 2% over a 60-trading-day period.

Panel A: Novice Trader Profile



Panel B: Intermediate Trader Profile

Panel C: Professional Trader Profile

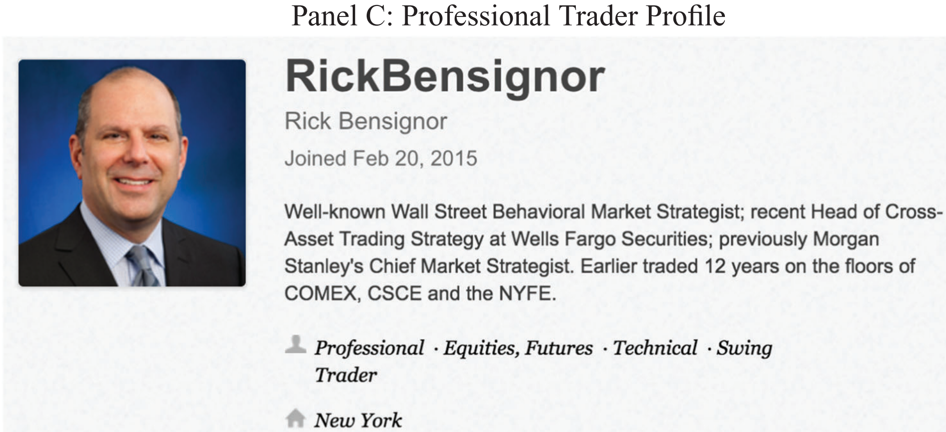


Figure 3. Examples of StockTwits user profiles. This figure presents screenshots of representative user profiles from StockTwits that illustrate the difference between novice, intermediate, and professional StockTwits users. (Color figure can be viewed at wileyonlinelibrary.com)

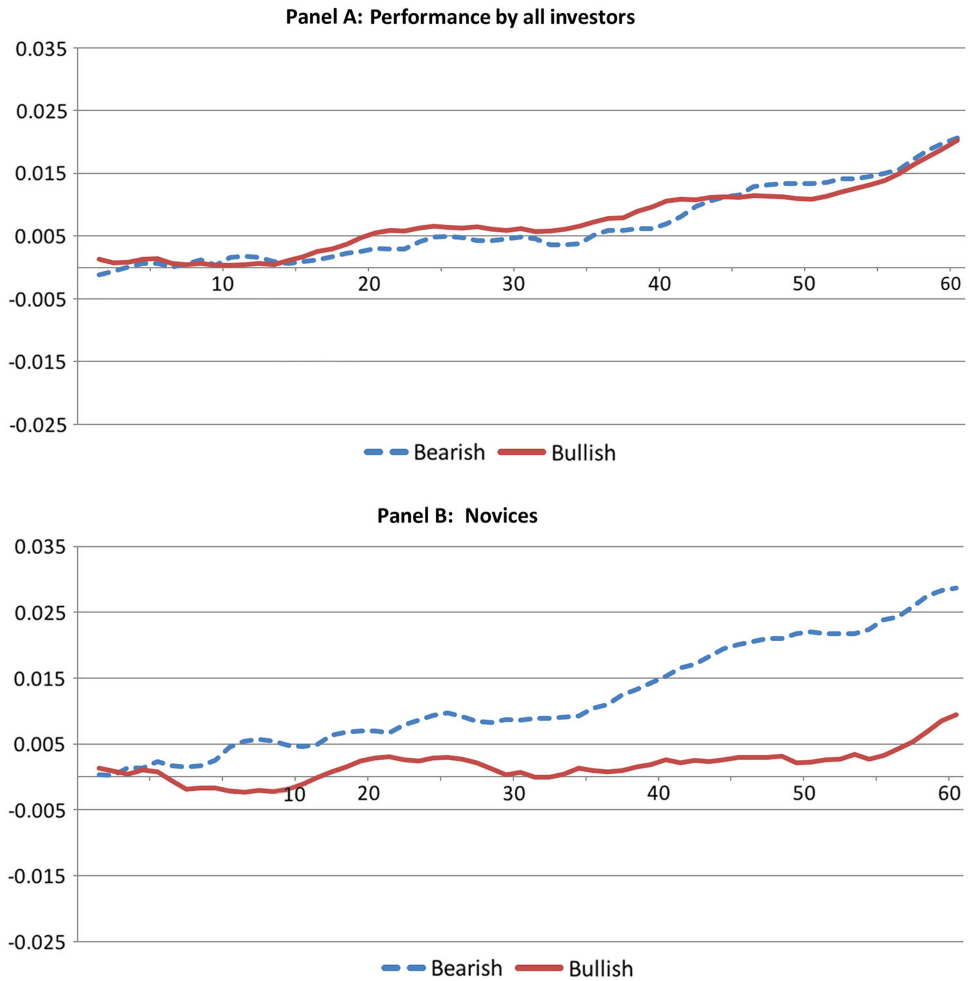


Figure 4. Performance of StockTwits sentiment strategies. This figure presents the cumulative abnormal returns of strategies that buy when sentiment is bullish and sell when sentiment is bearish for several sentiment classifications: the sentiment of all StockTwits users (Panel A), the sentiment of novice investors (Panel B), the sentiment of intermediate investors (Panel C), and the sentiment of professional investors (Panel D). (Color figure can be viewed at wileyonlinelibrary.com)

D. Why Do Users Post Messages?

In constructing a measure of disagreement, it is important that the sentiment expressed on StockTwits reveals investors' true opinions. We therefore want to rule out the possibility that users are trying to manipulate the stock market by posting fake opinions. For example, if a user thinks that the stock price will go down, and as a result, wants to sell the stock, she could post really bullish messages in an attempt to increase the price temporarily, which would

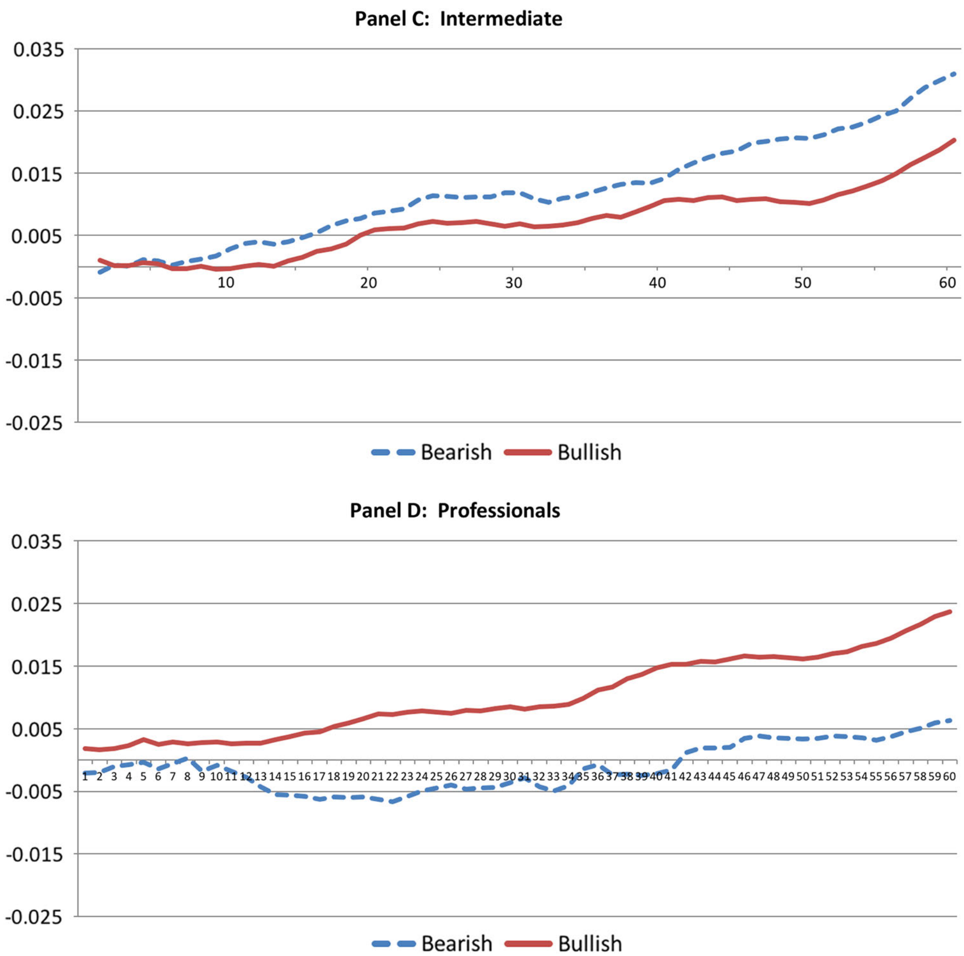


Figure 4. *Continued.*

allow her to sell at a higher price. This would invalidate our measure, as we would record her opinion as bullish even though she is bearish on the stock. However, this does not appear to be an important concern in our data, for several reasons. First, anecdotal evidence suggests that investors post on social networks to attract followers and gain Internet fame or a job,[11] which makes it in their best interest to provide their best forecast of a stock's future performance and their honest opinion about the stock. Second, per StockTwits policy, messages cannot be retroactively withdrawn by the user, which further enhances incentives to post true, reliable opinions. Third, since we concentrate

[11] For an example of an article on the fame motive for posting to investment social networks, see the *Wall Street Journal* article "Retail Traders Wield Social media for Investing Fame" from April 21, 2015 (article here).

on the 100 most-discussed firms, the firms we examine are highly liquid and have large market caps, and thus, it is unlikely that individual investors would think that they can move the stock price.[12]

II. Measuring Sentiment and Disagreement

A. Sentiment Classification

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.

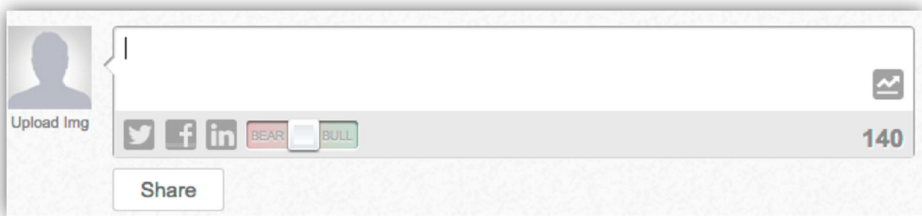


Table IV, Panel A, column (1), presents the distribution of sentiment across messages in the original sample. According to these summary statistics, 18.3% of classified messages are bearish and 81.7% are bullish. Even though the setting and time period are different, our classifications yield similar relative frequencies as 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 bearish, but the user did not select the option. To incorporate this information into the analysis, 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.[13] 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 IV, Panel A, column (2), presents the distribution of sentiment in the final data set. The fully classified data set has 452,258 bearish messages and 989,793 bullish messages.

[12] Consistent with this observation, a recent paper by Kogan, Moskowitz, and Niessner (2018) shows that the effects of "fake news" on financial markets are confined to small, illiquid stocks. Manipulation of large stocks through the posting of false opinions is typically infeasible.

[13] Prior papers that use message data (e.g., Antweiler and Frank (2004), Giannini, Irvine, and Shu (2018)) must construct a training data set (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 preclassified 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.

Table IV

Sentiment and Disagreement Summary Statistics

In this table, we present summary statistics for our sentiment and disagreement measures. Panel A reports 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 approach to the unclassified messages, in column (2). Panel B reports the sentiment (average bullishness) by investment philosophy. Panel C presents summary information on the StockTwits measure of disagreement. The first three rows provide summary statistics on disagreement for all investors, disagreement across groups with different investment philosophies, and the weighted-average disagreement within groups. The weights are proportional to the number of investors adhering to each approach. The table further shows the distribution of within-group disagreement by the individual investment philosophies. Panel D presents the autocorrelation of the disagreement and sentiment measures. Panel E presents correlations between our three main disagreement measures and other commonly used measures of disagreement (analyst dispersion, return volatility, the Giannini, Irvine, and Shu (2018) measure, and abnormal log trading volume).

Panel A: Sentiment Classification

	Number of Messages	
Sentiment	Original Sample	MaxEnt Classification
Bearish	86,615	452,258
Bullish	385,753	989,793
Unclassified	969,683	

Panel B: Sentiment Summary Statistics

	Average Sentiment	
	Mean	*SD*
All investors	0.342	0.492
Fundamental	0.146	0.494
Technical	0.264	0.535
Momentum	0.237	0.504
Growth	0.252	0.489
Value	0.118	0.457

Panel C: Disagreement within and across Approaches

	Mean	*SD*	Min	p25	p50	p75	Max
All investors	0.467	0.446	0	0	0.628	0.932	1
Cross-group disagreement	0.382	0.262	0	0.151	0.435	0.545	1.117
Within-group disagreement	0.245	0.299	0	0	0	0.480	0.994
Fundamental	0.172	0.354	0	0	0	0.531	1
Technical	0.341	0.434	0	0	0	0.866	1
Momentum	0.249	0.401	0	0	0	0.699	1
Growth	0.171	0.346	0	0	0	0.000	1
Value	0.124	0.313	0	0	0	0.000	1

(*Continued*)

Table IV—*Continued*

Panel D: Autocorrelations for Disagreement and Sentiment

Disagreement			Sentiment					
Overall	Within-Group	Cross-Group	Average	Fundamental	Technical	Momentum	Growth	Value
0.345	0.500	0.265	0.311	0.183	0.185	0.164	0.187	0.160

Panel E: Correlations with Other Disagreement Measures

Disagreement Measure	Analyst Dispersion	Return Volatility	Giannini, Irvine, and Shu Measure	Abnormal Log Volume
All investors	0.030	−0.018	0.206	0.116
Cross-group disagreement	−0.054	−0.151	0.391	0.052
Within-group disagreement	0.062	0.076	0.129	0.188

Table V

Examples of Bullish and Bearish Messages

In this table, we present examples of some of the more bullish and bearish messages, according to our classification algorithm.

Bullish Messages
"$ZNGA BUY BUY BUY BUY BUY BUY BUY BUY BUY BUY BUY BUY BUY BUY BUY BUY BUY BUY BUY BUY BUY BUY BUY BUY BUY BUY BUY BUY BUY BUY BUY BUY BUY"
"$RXII way oversold but no worries I'm confident in the company's future!"
Bearish Messages
"$ZNGA Investors need to see growth, there is none with this game anymore, sad but true, they need to keep investors happy, just business"
"$AAPL It happened to BBRY & it appears it is happening to aapl, the downhill slide can happen in a few qtrs 1st a loss of mrkt share then :("

Based on a reading of the classified messages, the sentiment classifier appears to accurately capture truly bullish and bearish messages. Table V provides several examples of classified messages. In addition to reading the classified messages, we systematically evaluate the accuracy of the sentiment classifier using a cross-validation exercise. For the cross-validation, we hold out 10% user-classified messages and train the algorithm on the remaining 90% of the messages. The cross-validation shows that, on average, the predictive accuracy of our classifier is 83%. This high degree of accuracy increases our confidence in using the classification scheme on unclassified messages.

Beyond the cross-validation evidence, a potential concern with 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 unclassified. To examine this possibility, we begin by randomly selecting 100,000 preclassified messages by users who classify at least one message in the data set, to train the maximum entropy algorithm. Then, using this training set, we deploy the maximum entropy algorithm to classify a randomly selected set of 200,000 messages, where half (100,000) of these were a second set of preclassified messages and half were unclassified by StockTwits users. For each message, the algorithm computes a probability that describes the level of confidence that the classification is either bullish or bearish based on the maximum entropy algorithm and the text of the message. Finally, we examine whether the unclassified messages and user-classified messages differ in the algorithm's confidence in assigning a bullish versus bearish label. The distributions of these confidence levels are almost identical with the mean being 0.958 for unclassified and 0.959 for preclassified messages, and the standard deviation being 0.104 and 0.105, respectively. These results confirm that the unclassified messages are very similar in nature to the user-classified messages.[14]

B. Average Sentiment Measure

We follow Antweiler and Frank (2004) in constructing a sentiment measure from bearish and bullish message data. We first code each bearish message as −1 and each bullish message as 1. We then take the arithmetic average of these classifications at the firm× day× 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. For our base measure, 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. Alternatively, we use the message time stamp to construct a before-market-opens (BMO) version of the sentiment measure, which is useful for empirical tests that exploit timing. Figure 5 presents a time line that illustrates our measurement.

Table IV, Panel B, presents summary statistics on average sentiment across for all users, as well as average sentiment broken down by investment philosophy. As investors tend to post bullish messages more frequently than bearish messages, it makes sense that the average sentiment for all users is 0.342 (closer to 1 than −1). During our sample period, technical investors are the most likely to post bullish messages, whereas value investors are the most likely to post bearish messages. We present summary statistics for the sentiment measure broken down by experience level and holding period in the Internet Appendix Table IA.IV.

[14] Furthermore, in the Internet Appendix Table IA.IX, we replicate our main findings using only messages that were classified by the investors themselves (user-classified messages). We obtain similar results.

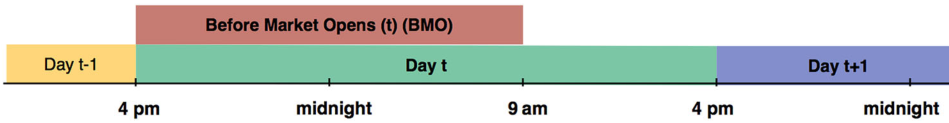


Figure 5. Time line for calculating disagreement. This figure presents a time line illustrating how we compute disagreement. We assign any messages that are posted on day $t-1$ after 4 p.m. to trading day t because trading stops at 4 p.m. on day $t-1$. Similarly, we assign any messages posted after 4 p.m. 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 4 p.m. on day $t-1$ and before 9 a.m. on day t. (Color figure can be viewed at wileyonlinelibrary.com)

For our main analysis, we compute the average sentiment measure by assigning each message an equal weight. As a robustness check, 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 Internet Appendix, our main findings are not sensitive to the choice of weights in the calculation of the average sentiment (see Table IA.IX).

Our reading of messages shows that investors tend to post new messages that reflect their opinion about the future prospects of a stock, which maps naturally to near-term trading sentiment. Specifically, a bullish message about the stock typically indicates that the investor intends to buy the stock, whereas a bearish message indicates that the investor intends to sell the stock. If no messages were posted for a given firm-day-group, we set the average sentiment measure equal to zero, as we assume that users who do not post are not interested in buying or selling in the near term.[15]

One potential concern with an expressed sentiment measure 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 who does not hold the stock cannot trade because of short-selling constraints. Given that many investors face short-selling constraints (Hong and Stein (2003), Engelberg, Reed, and Ringgenberg (2018)), a tilt toward bullish sentiment is natural. A bearish

[15] In the literature on trading and sentiment constructed from analyst forecasts (e.g., see Gleason and Lee (2003)), changes in sentiment (i.e., analyst forecast revisions) are what lead to trading, rather than the overall level of sentiment (i.e., analyst forecasts themselves). The literature focuses on sentiment changes to isolate the portion of sentiment that is new information (i.e., not yet impounded into prices or traded upon in the market). From this point of view, our sentiment measure is similar to a "sentiment changes" measure, as StockTwits users tend to post when their information is new. Our normalization of average sentiment to zero in the no-messages case is consistent with this interpretation. However, one may be concerned that most messages do not contain new information (i.e., they are reiterations of prior statements), but rather reflect stale past sentiment. Stale sentiment should be less directly related to trading volume than sentiment updates. Nevertheless, if sentiment reflects stale opinions, it would be appropriate to set average sentiment equal to the prior day's measure if no messages were posted for a given firm-day-group. As we show in the Internet Appendix Table IA.IX, our main findings continue to hold if we make this choice instead. As theory would predict, the empirical link between trading volume and disagreement is weaker under this assumption.

investor with a strict short-sale constraint can only sell the stock if she holds shares of the stock in her portfolio. 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 on which investors are bearish, and because these stocks get less investor attention, bearish messages should be reported less frequently.

Using the percent of institutional ownership of a firm as a proxy for shorting constraints (Nagel (2005)), we find that the fraction of bearish messages for companies in the top quartile of institutional holdings (lax shorting constraints) is 0.37, compared with 0.23 for companies in the bottom quartile (tight shorting constraints). This evidence suggests that the bullish-bearish imbalance in our sentiment measure is most likely due to the presence of short-selling constraints.

C. Measuring Disagreement

We construct the overall disagreement measure by computing the standard deviation of expressed sentiment across messages, as in Antweiler and Frank (2004). Because the underlying variable is binary (−1/1), the variance of the sentiment measure for period t equals $(1 - AvgSentiment^2)$. Although Antweiler and Frank (2004) used this formulation to study disagreement using opinions expressed across the whole set of investors, we adapt this insight to also measure disagreement within subgroups of investors. Specifically, the within-group measure for a given firm× day× group is computed as

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

where a group can represent either all investors or those investors with a given investment approach. This disagreement measure ranges from zero to one, with one being maximal disagreement. We apply the formula to firm-day-group observations that have nonzero messages. When there are no messages for a particular firm-day-group, the logic for this formula breaks down (i.e., it no longer represents the standard deviation of sentiment). To compute disagreement in this corner case, we maintain the assumption that nonposting means that traders do not wish to buy or sell in the near term. Accordingly, we normalize disagreement in the no-message case to zero, consistent with latent agreement. This choice of how to normalize the no-message case is consistent with the idea that minimal disagreement should correspond to minimal trading.[16] In Table IA.IX, we challenge our assumption and replace the disagreement on days with no messages to be the last value of disagreement that

[16] This choice deviates from how Antweiler and Frank (2004) handle stock-days in which no messages come out. If no messages are posted during a given time period, Antweiler and Frank (2004) set disagreement for that time period to one, and justify their choice by arguing that no information came out during that period and hence there is latent disagreement. As we believe that the opinions on StockTwits likely reflect new information, it is more appropriate in our context to set the case of no messages to be a change in disagreement of zero.

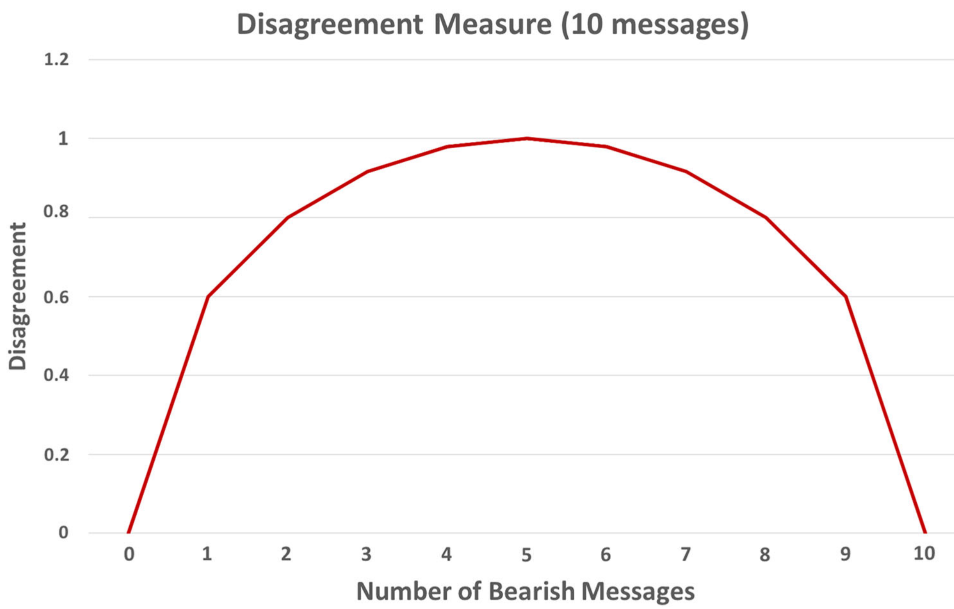


Figure 6. An example of the disagreement measure. This figure illustrates how our main disagreement measure depends on the average sentiment of the underlying messages. Specifically, for the case of 10 total messages, the figure presents how the value of the disagreement measure depends on the number of bearish messages. (Color figure can be viewed at wileyonlinelibrary.com)

we observe for the given stock. As the results in column (2) show, our main findings are conceptually robust to this assumption.

To illustrate the properties of the disagreement measure, consider the following example of disagreement across 10 messages posted for the same firm× day× group combination. In Figure 6, 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 sentiment is either bearish or bullish. The disagreement measure is maximized at 1 when there are five bullish and five bearish messages. Since the measure is a square root function, the disagreement measure changes the most (the measure has the largest slope) when there are few bullish or few bearish messages.

Separately, we construct a measure of cross-group disagreement by computing the standard deviation of average sentiment ($AvgSentiment_{itg}$) across investment approaches, weighted by the number of individuals in that approach group. We implement the weighted approach to give our measure internal consistency, as dispersion of beliefs between two groups with many investors will contribute more to trading volume than dispersion of beliefs between two groups with few investors. The formula for cross-group disagreement is

$$CrossDisagremeent_{it} = \sqrt{\frac{\sum_{a\in A} n_a (AvgSentiment_{at} - AvgSentiment_t)^2}{\frac{G-1}{G}(n_F + n_T + n_M + n_V + n_G)}}, \quad (3)$$

where A ={Fundamental, Technical, Momentum, Value, or Growth}, n_a is the number of individuals in group a in January 2013, $AvgSentiment_{at}$ is the average sentiment of group a on day t, $AvgSentiment_t$ is the average sentiment of all groups on day t, and G is the number of investment philosophies. Similar to the within-group disagreement measure, cross-group disagreement captures changes in the level of disagreement because StockTwits users are likely to post when their sentiment about the firm changes. Hereafter we refer to our measures as "disagreement," but it is appropriate to think of the measures as capturing changes in investors' level of disagreement.

In Panel C of Table IV, we summarize our disagreement measures, both across and within groups. The first three rows summarize disagreement for all investors, cross-group disagreement, and the weighted-average disagreement within investment philosophies (within-group disagreement). In constructing within-group disagreement, we weight by the number of users following a given approach. The average for our main disagreement measure for all investors is 0.467, and the median is 0.628. The average cross-group disagreement is 0.382, and the weighted average within-group disagreement is 0.245. In addition, we also report the autocorrelation of each measure in Panel D of Table IV. Given the daily frequency, these autocorrelations—which range from 0.265 to 0.500 for our main measures—are quite low, which indicates that significant new information is reflected in our measures on a daily basis.[17]

As a quantification of how investment approaches contribute to overall disagreement, note that both disagreement among "All Investors" and "Weighted average within-group Disagreement" are based on the same formula, and thus are on a comparable scale to one another.[18] Contrasting overall disagreement with within-group disagreement, we note that splitting disagreement into groups reduces the average disagreement from 0.467 to the weighted-average within-group disagreement of 0.245, an overall reduction of 47.5%.[19] The importance of investment philosophies toward explaining overall disagreement is

[17] Another natural question is how firm characteristics and market conditions relate to our disagreement measures. To this question, Internet Appendix Table IA.V presents two sets of results. On firm characteristics, disagreement is greater for larger firms and bears an insignificant relation to book-to-market ratio. On market conditions, disagreement is positively related to market volatility and recent returns. Our specifications account for these factors using firm and date fixed effects, as well as controls for media attention, recent volatility, and market returns.

[18] For some tests in Section III, we use the variable *DisagreementRatio*, which is equal to the fraction of disagreement that can be attributed to within-group differences of opinion. Similar to the construction of R^2 in an analysis of variance (ANOVA) context, we construct *DisagreementRatio* using the variance decomposition of total disagreement into within-group and cross-group disagreement.

[19] We did not take the ratio of our cross-group disagreement measure to the all-investors measure because the standard deviation of average sentiment is not on the same scale as the standard deviation of user sentiment across all messages (our "all investors" disagreement measure). Here, it also becomes clear why using the number of individuals in a given approach as weights for the cross-group and the average within-group disagreement helps keep our measures internally consistent because the "all investors" disagreement measure puts more weight on approaches with more users.

robust to employing different subsamples of firms and users, different weighting schemes, and alternative measures of disagreement (see Internet Appendix Table IA.IX, Panel A). Across the various alternative specifications, we find that the splitting disagreement into groups reduces average disagreement by 40.7% on the low end to 63.6% on the upper end.

The patterns of within-group disagreement for different investment approaches also provide interesting insights. 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 while there are many ways to be a technical investor, there is much more standardization in what value investing and growth investing mean.[20] We also summarize within-group disagreement by investor experience and by investment horizon in the Internet Appendix (Table IA.IV).

C.1. Contrasting with Alternative Measures of Disagreement

It is instructive to relate our disagreement measures to existing measures of disagreement in the literature. Panel E of Table I presents evidence on how our disagreement measures compare with notable measures of disagreement in the literature: analyst dispersion as in Diether, Malloy, and Scherbina (2002), return volatility, and divergence of sentiment on StockTwits from sentiment expressed in the media, as in Giannini, Irvine, and Shu (2018). We separately examine disagreement among all investors, cross-group disagreement, and the weighted average within-group disagreement.

To provide a comparison to analyst dispersion, 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, and then calculate its correlation with analyst dispersion. As can be seen in Panel E, column (1), the two measures do not strongly correlate with one another.

In column (2), we examine the correlation of our disagreement measures with return volatility. Interestingly, cross-group disagreement is negatively correlated with both analyst dispersion and return volatility, whereas the within-group measure has a positive correlation. The significant within-group correlation suggests that analyst dispersion and return volatility are better measures of information-driven disagreement.

Finally, we construct a measure of divergence of opinion on StockTwits from sentiment expressed in the media, as in Giannini, Irvine, and Shu (2018). We find that cross-group disagreement is more correlated with the Giannini, Irvine, and Shu (2018) measure than within-group disagreement, consistent

[20] For example, many technical investors use the subjective method of finding patterns in charts (e.g., the head-and-shoulders pattern) and therefore often come to opposite conclusions.

with the fact that StockTwits users and the media are using different models to process financial information.[21]

When we correlate analyst dispersion at the monthly level with abnormal trading volume, we find a weak and insignificant correlation (0.0388). In contrast, our measure of disagreement correlates much more strongly with abnormal trading volume. Specifically, in column (4), we present the correlations between daily abnormal log trading volume and our daily measures of investor disagreement. We find that the correlation between overall disagreement and abnormal log trading volume is 0.116. This correlation is substantially greater than correlations using other measures of disagreement. Moreover, abnormal trading volume is more strongly correlated with weighted-average within-group disagreement than with the cross-group disagreement measure.

D. Variation-in-Sentiment Test of Cross-Group Disagreement

This section describes a more systematic test of the importance of investment philosophies for explaining divergence of opinion by examining the degree to which different investment philosophies explain variation in changes to sentiment. If adhering to differing investment philosophies leads investors to disagree, investment philosophies should significantly explain variation in sentiment over time. One example of cross-group disagreement in this vein is our evidence that fundamentally oriented investors react to earnings news, but nonfundamental investors do not (Figure 1). In this section, we present a generic test for whether investment philosophies relate to variation in sentiment over time.

The intuition behind our variation-in-sentiment test is as follows. Imagine that information about firm i comes out on date t that is differentially interpreted by groups A and B. In this case, differential interpretation means that $\Delta AvgSentiment_{itg} = AvgSentiment_{itg} - AvgSentiment_{i(t-1)g}$ is different for $g = A$ than for $g = B$. We extend this intuition across all investment philosophies in performing an analysis of variance of the following linear regression specification:

$$\Delta AvgSentiment_{itg} = FirmFEs + DateFEs + ApproachFEs + \epsilon_{itg}, \qquad (4)$$

[21] Giannini, Irvine, and Shu (2018) measure the divergence between investor sentiment on StockTwits and the sentiment of breaking news articles and firms' 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, Irvine, and Shu (2018) 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 construct an alternative measure that—like Giannini, Irvine, and Shu (2018)—contrasts investor sentiment on StockTwits with media sentiment as reported in the Ravenpack database. The Appendix 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, Irvine, and Shu (2018) measure in an out-of-sample replication of their proxy for disagreement.

Table VI

Quantifying Disagreement across Investment Models

This table presents analysis of variance specifications for first-differenced sentiment, using the regression

$$\Delta AvgSentiment_{itg} = \alpha_t + \gamma_i + InvestmentPhilosophyFEs + \epsilon_{itg},$$

where $\Delta AvgSentiment_{itg}$ is the difference between the average sentiment measure on day t and day $t-1$. The regressions include date (α_t), firm (γ_i), and investment philosophy fixed effects as noted in the columns. We also examine the results if we include investment philosophy × year-month and investment philosophy × year-week fixed effects.

	Δ *Sentiment*$_{itg}$				
Sentiment Categories	(1)	(2)	(3)	(4)	(5)
Firm FEs	X	X	X	X	X
Date FEs		X	X	X	X
Investment Philosophy FEs			X		
Investment Philosophy × Year-Month FEs				X	
Investment Philosophy × Year-Week FEs					X
R^2	0.004	0.007	0.016	0.018	0.022
Observations	102,567	102,567	102,567	102,567	102,567

where $\Delta AvgSentiment_{itg}$ is first-differenced average sentiment on date t for firm i by investors of approach g. We include firm, date, 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 over time. Beyond accounting for different levels of sentiment, because of the first-difference specification, the firm and time fixed effects allow for differential trends in sentiment by firm.

In Table VI, we present the ANOVA decomposition of sentiment trends from estimating equation (4). We find that differing investment approaches explains 1.6% of the variation in first-differenced average sentiment. In contrast to the approach fixed effects, time and firm fixed effects explain little of the variation in first-differenced average sentiment, only 0.6%.[22] In columns (4) and (5), we present specifications that interact approach with year-month fixed effects and year-week fixed effects. Beyond the approach fixed effects, these interactions explain up to 0.6% of the variation in sentiment (equivalent to the variation explained by firm fixed effects, but less than the variation explained by approach

[22] The effect of different approaches is similar whether we estimate the specification on trends in sentiment (main text) or on levels of sentiment. In the levels specifications, we find that firm and time dummies alone explain 10.7% of the variation in sentiment. Adding 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 10.2% of the change in disagreement (variation in sentiment) that is explained using firm and time fixed effects. The fact that differential approaches explain slightly more variation in sentiment trends is consistent with differential interpretation. These levels of sentiment specifications are presented in Internet Appendix Table IA.VII, Panel A.

fixed effects alone). The additional explanatory power of these interactions suggests that a small part of the differences across approaches occurs at a slower frequency of one day, which is captured by the approach fixed effects in the first-difference specification.

The results from this test suggest that differing investment approaches matter for disagreement beyond the simple summary statistics of disagreement presented in Table IV. Because these specifications include firm and date fixed effects, the results of this analysis cannot be explained by different investor types following different firms, or by different investor types posting messages at different times.

III. Empirical Analysis of Disagreement

In this section, we present the main results from our disagreement measures. First, we present results that link our various disagreement measures to trading volume. These results contrast the quantitative implications of cross-group versus within-group disagreement. Second, we present results on sophisticated investors versus unsophisticated investors that provide evidence of gradual information diffusion. Third, we exploit the daily frequency of our disagreement measure to provide more precise insight into the role of disagreement in driving the spike in volume around earnings announcements.

A. Trading Volume and Disagreement, within versus across Groups

Here, we evaluate how trading volume relates to each of our measures of disagreement, with an emphasis on the contrast between cross-group and within-group disagreement. Specifically, we estimate the empirical link between disagreement and abnormal trading volume using the following regression specification:

$$\begin{aligned} AbLogVol_{it} = \alpha_t + \gamma_i + \beta_1 DisMeasure_{it} + \beta_2 AbLogVol_{it-1} \\ + \beta_3 Media_{it} + \gamma Controls_{it} + \epsilon_{it}, \end{aligned} \tag{5}$$

where $AbLogVol_{it}$ is the abnormal log trading volume on date t for firm i. It is calculated as the difference between the log volume on date t and the average log volume from trading days $t-140$ to $t-20$ (six-month period, skipping a month). The variable $DisMeasure_{it}$ is one of the disagreement measures described in Section II (overall, within-group, across group). For ease of comparison across measures, we standardize each disagreement measure to have mean 0 and standard deviation 1. This standardization implies that the coefficient of interest, β_1, equals the percentage change in abnormal trading volume for a one-standard-deviation increase in disagreement.

To account for alternative interpretations, all specifications include date and firm fixed effects (α_t and γ_i). We also control for abnormal trading volume on day $t-1$ to account for persistence in abnormal trading volume. To account for firm-date-specific spikes in attention, we use the indicator variable $Media_{it}$, which equals 1 if there is a news article about firm i on date t in the *Wall Street Journal* or the *New York Times*. We also include controls for recent stock market volatility and recent abnormal returns. Across specifications, standard errors are double-clustered by date and firm to account for within-firm autocorrelation and common daily shocks.

Table VII presents our results on the link between disagreement and trading volume. We present the results using disagreement among all investors in columns (1) and (2) of Panel A. According to the specification in column (1), a one-standard-deviation increase in disagreement is associated with a contemporaneous increase in abnormal trading volume of 10.0%. This estimate is statistically significant at the 1% level and is robust to including firm and date fixed effects, lagged abnormal log volume, and controls for media attention, recent volatility, and abnormal returns. In column (2), we additionally exploit the precise timing of the StockTwits messages to construct a measure of disagreement only from messages posted before the market opens (BMO disagreement). By construction, the BMO disagreement measure precedes trading volume. We find that a one-standard-deviation increase in disagreement before the market opens is associated with 5.3% greater *subsequent* abnormal trading volume. Although the magnitude of the estimate on BMO disagreement is weaker, this finding alleviates the concern that disagreement on StockTwits is merely a reaction to trading volume. Although investors who post on StockTwits make up a small fraction of overall trading in the stock market, this finding suggests that our measure of disagreement is a good proxy for overall changes in disagreement in the market.

In comparison to overall disagreement, we find a notably weaker positive relation between cross-group disagreement and trading volume (columns (3) and (4) of Panel A), which is statistically significant at the 1% level. Specifically, a one-standard-deviation increase in cross-group disagreement is associated with 3.0% greater abnormal trading volume, or 30% of the magnitude of the overall disagreement measure. Unlike overall disagreement, the estimate on cross-group disagreement is the same magnitude (3.3%) when relying on differences of opinion from before the market opens.

In column (5) of Panel A, we find that a one-standard-deviation increase in within-group disagreement is associated with 17.5% greater abnormal trading volume, which is more than five times greater than the comparable increase in cross-group disagreement. As with disagreement among all investors, the magnitude of the estimate in column (6) using disagreement from before the market opens is weaker in magnitude but still economically large (8.5%) and statistically significant at the 1% level. These findings suggest that disagreement within groups is an important determinant of trading volume, and while cross-group disagreement is also important, it leads to less trading than disagreement within groups.

In Panel B, we present a series of regressions that contrast the effects of cross-group and within-group disagreement. The specifications in columns (1) and (2) include both cross-group and within-group disagreement measures in the same regression of trading volume. Consistent with the estimates in Panel A, we find

Table VII

Disagreement and Trading Volume

This table examines how each of the measures of disagreement relates to trading volume. We run the regression

$$AbLogVol_{it} = \alpha_t + \gamma_i + \beta DisagreementMeasure_{it} + \gamma AbLogVol_{it-1} + \delta Controls_{it} + \epsilon_{it}.$$

$AbLogVol_{it}$ is the difference between log volume in period t and the average log volume from trading days $t-140$ to $t-20$ (six-month period, skipping a month) for firm i. In columns (1) and (2) of Panel A, $DisagreementMeasure_{it}$ is the disagreement among all investors. In columns (3) and (4), it is $CrossDisagreement_{it}$, disagreement across different investment philosophies for firm i on day t. In columns (5) and (6), it is $WithinDisagreement_{it}$, disagreement among investors with the same investment philosophies. The disagreement measures are either contemporaneous to abnormal log volume t, or are constructed from messages that were posted before the market opens (BMO) (between 4 p.m. on day $t-1$ and 9 a.m. on day t). We standardize the disagreement measures by subtracting the mean and dividing by the standard deviation, calculated over the entire sample period. Since trading volume tends to be autocorrelated, we also control for abnormal trading volume on day $t-1$. Controls include $MediaArticle_{it}$, a dummy variable equal to 1 if firm i was mentioned in the *Wall Street Journal* or the *New York Times* on day t, volatility ($t-5$ to $t-1$), the standard deviation of abnormal returns over days $t-5$ to $t-1$, and cumulative abnormal returns over days $t-30$ to $t-6$ and $t-5$ to $t-1$. In Panel B, we examine cross-group and within-group disagreement, both at period t and *BMO*. We also examine the disagreement variance ratio, which is the fraction of overall disagreement (in variance units) that originates from differences of opinion within group. In Panel C, we examine the effect of cross-group disagreement on log abnormal trading volume separately for companies with high and low institutional ownership. We define a firm as having high institutional ownership if it is above the median for our firms. All regressions include date and firm fixed effects (α_t and γ_i). Standard errors are clustered by firm and date. *, **, and *** indicate statistical significance at the 10%, 5%, and 1% level, respectively. Standard errors are in parentheses.

Panel A: Disagreement and Trading Volume

	Abnormal Log Volume (t)					
Disagreement measure	(1)	(2)	(3)	(4)	(5)	(6)
Disagreement (t)	0.099***					
	(0.008)					
Disagreement (BMO, t)		0.053***				
		(0.006)				
Cross-Group Disagreement (t)			0.030***			
			(0.008)			
Cross-Group Disagreement (BMO, t)				0.033***		
				(0.005)		
Within-Group Disagreement (t)					0.175***	
					(0.011)	
Within-Group Disagreement (BMO, t)						0.085***
						(0.009)
AbLogVol ($t-1$)	0.719***	0.723***	0.727***	0.725***	0.705***	0.717***
	(0.015)	(0.015)	(0.015)	(0.015)	(0.017)	(0.016)
Media (t)	0.069***	0.071***	0.080***	0.077***	0.045***	0.057***
	(0.013)	(0.012)	(0.013)	(0.013)	(0.010)	(0.011)
Volatility ($t-5$ to $t-1$)	0.259	0.364	0.398*	0.391*	0.164	0.331
	(0.229)	(0.237)	(0.233)	(0.233)	(0.238)	(0.243)
AbRet ($t-5$ to $t-1$)	0.178***	0.174***	0.173***	0.172***	0.172***	0.167***
	(0.051)	(0.052)	(0.051)	(0.051)	(0.052)	(0.053)
AbRet ($t-30$ to $t-6$)	0.113***	0.119***	0.117***	0.118***	0.108***	0.117***
	(0.026)	(0.024)	(0.024)	(0.024)	(0.027)	(0.025)
Observations	42,041	42,041	42,041	42,041	42,041	42,041
R^2	0.637	0.633	0.632	0.632	0.649	0.636

(*Continued*)

Table VII—*Continued*

Panel B: Contrasting Within-Group and Cross-Group Disagreement

	Abnormal Log Volume (t)			
Disagreement measure	(1)	(2)	(3)	(4)
Cross-Group Disagreement (t)	0.045***			
	(0.008)			
Within-Group Disagreement (t)	0.181***			
	(0.012)			
Cross-Group Disagreement (BMO, t)		0.036***		
		(0.005)		
Within-group Disagreement (BMO, t)		0.087***		
		(0.009)		
Var Disagreement Ratio (t)			0.122***	
			(0.008)	
Var Disagreement Ratio (BMO, t)				0.058***
				(0.006)
AbLogVol ($t-1$)	0.700***	0.711***	0.715***	0.722***
	(0.017)	(0.017)	(0.016)	(0.016)
Media (t)	0.045***	0.054***	0.066***	0.070***
	(0.010)	(0.011)	(0.012)	(0.012)
Volatility ($t-5$ to $t-1$)	0.099	0.282	0.261	0.376
	(0.233)	(0.240)	(0.225)	(0.238)
AbRet ($t-5$ to $t-1$)	0.175***	0.167***	0.163***	0.166***
	(0.053)	(0.053)	(0.051)	(0.052)
AbRet ($t-30$ to $t-6$)	0.103***	0.114***	0.111***	0.117***
	(0.027)	(0.025)	(0.026)	(0.025)
Observations	42,041	42,041	42,041	42,041
R^2	0.651	0.637	0.641	0.633

Panel C: Sample Splits by High versus Low Institutional Ownership

	Abnormal Log Volume (t)					
	High Institutional Ownership			Low Institutional Ownership		
Disagreement measure	(1)	(2)	(3)	(4)	(5)	(6)
Cross-Group Disagreement (t)	0.028***			0.032**		
	(0.008)			(0.013)		
Within-Group Disagreement (t)		0.175***			0.181***	
		(0.016)			(0.017)	
Overall Disagreement (t)			0.097***			0.107***
			(0.008)			(0.015)
AbLogVol (t-1)	0.703***	0.668***	0.689***	0.735***	0.718***	0.730***
	(0.037)	(0.046)	(0.041)	(0.014)	(0.016)	(0.014)
Media (t)	0.080***	0.050***	0.071***	0.082***	0.044**	0.070***
	(0.018)	(0.015)	(0.018)	(0.023)	(0.018)	(0.022)
Volatility ($t-5$ to $t-1$)	0.465	0.161	0.295	0.300	0.100	0.171
	(0.281)	(0.318)	(0.287)	(0.292)	(0.300)	(0.287)
AbRet ($t-5$ to $t-1$)	−0.012	−0.009	−0.005	0.233***	0.221***	0.231***
	(0.071)	(0.065)	(0.068)	(0.060)	(0.063)	(0.062)
AbRet ($t-30$ to $t-6$)	0.055*	0.051	0.058*	0.129***	0.114***	0.120***
	(0.033)	(0.035)	(0.033)	(0.032)	(0.034)	(0.033)
Observations	20,990	20,990	20,990	21,051	21,051	21,051
R^2	0.578	0.610	0.589	0.657	0.668	0.660

that within-group disagreement exhibits a much stronger relationship with trading volume, whether contemporaneous (column (1)) or using disagreement before the market opens (column 2). Columns (3) and (4) present an alternative version of this test using *DisagreementRatio*, the fraction of the variance of sentiment that is due to within-group disagreement.[23] Consistent with the first two columns, the disagreement ratio results indicate that within-group disagreement bears a stronger link to trading volume. When within-group disagreement comprises a greater fraction of overall disagreement, there is significant abnormal trading volume.

Taken more broadly, these findings provide evidence that the link between differential interpretation (i.e., dispersion of opinion across investment philosophies) and trading decisions is notably weaker than the link between informational differences (i.e., dispersion of opinion within investment philosophies) and trading decisions. Although both sources of disagreement positively predict trading volume, the within-group disagreement effect is 2.5 to 4.0 times the cross-group disagreement effect. These differences are highly statistically significant and are robust to how we specify cross- versus within-group differences.

Finally, one may be concerned that StockTwits investors tend to be retail traders, and thus, may not represent the marginal investor's preferences. In a test in the Internet Appendix, we restrict attention to professional investors to address this concern and find quite similar results. An alternative tactic is to split by high versus low institutional ownership. To the extent that StockTwits is more representative of retail investors, we should observe a stronger relation between disagreement and trading volume among low-institutional-ownership (high-retail-ownership) firms. Panel C presents the results of estimating equation (5) separately by high- versus low-institutional-ownership firms (above versus below the median of institutional ownership) for each of our measures of disagreement. Consistent with the marginal investor intuition, we find that the coefficient estimate is larger among the low-institutional-ownership group, but the magnitude is not dramatically smaller. We also find a relation between disagreement and trading volume in the high-institutional-ownership subsample. That is, although many of the StockTwits users are retail investors, their opinions are not informative only about stocks traded by retail investors.

B. Disagreement and Sophistication

In this section, we use alternative cuts of the StockTwits data to provide tests of the gradual information diffusion hypothesis, whereby sophisticated investors discover information and trade on it before the information diffuses to less sophisticated investors (Hong and Stein (1999)). This analysis deepens

[23] We construct *DisagreementRatio* by preserving the variance equality, $VarTotal = VarWithin + VarAcross$. Dividing through, the fraction of total variance in sentiment that is due to within-group disagreement is $\frac{VarWithin}{VarTotal}$.

the insight from the prior section that within-group differences are important for trading volume.

We use self-reported experience levels from StockTwits user profiles to classify investors into sophisticated and unsophisticated categories. Specifically, we classify an investor as sophisticated (S) if the investor indicates professional as their experience level, and as unsophisticated (U) if the investor indicates either novice or intermediate as their experience level. Using these experience classifications, we calculate within-group disagreement for sophisticated and unsophisticated investors within each investment philosophy. We also separately calculate $|Dis S–U|$ the absolute value of the difference of average sentiment for sophisticated and unsophisticated investors within each investment philosophy. This measure captures the degree to which sophisticated and unsophisticated investors with the same investment philosophy disagree with one another. We then aggregate these measures to the stock-day level by computing the weighted average of *Dis S* and *Dis U* across investment philosophies, where the weights are the number of messages posted by investors in each investment philosophy-sophistication bin. To compute the weighted average for the $|Dis\ S–U|$ aggregated measure at the stock-day level, the weights are the number of messages posted by investors in the given investment philosophy.

If gradual information diffusion between sophisticated and unsophisticated investors is important for trading volume, $|Dis\ S–U|$ should exhibit a significant predictive relation to trading volume, holding *Dis S* and *Dis U* constant. Panel A of Table VIII provides evidence on the link between trading volume and the disagreement between sophisticated and unsophisticated investors. The main finding of interest is in column (2), which estimates the relation between BMO disagreement and subsequent trading volume. After controlling for disagreement within sophisticated and unsophisticated investors, BMO disagreement between sophisticated and unsophisticated investors with the same investment philosophy is significantly related to trading volume the following day. Specifically, a one-standard-deviation increase in BMO disagreement between sophisticated and unsophisticated investors with the same investment philosophy is associated with 3.1% greater abnormal trading volume, holding other factors constant. This result is consistent with gradual information diffusion (as in Hong and Stein (2007)) whereby sophisticated individuals obtain information earlier than unsophisticated individuals.

In Panel B, we provide a more direct evaluation of the diffusion of information by examining the within-day lead-lag relationship between sophisticated and unsophisticated sentiment. Specifically, we examine the degree to which BMO sentiment among sophisticated and unsophisticated investors predicts after-market-opens sentiment of each type. In column (1), we find that unsophisticated sentiment from before the market opens has no predictive power for subsequent sophisticated sentiment. In contrast, column (2) shows that sophisticated sentiment from before the market opens is a significant predictor of unsophisticated sentiment after the market opens. This pattern of results suggests gradual information diffusion whereby sophisticated investors obtain information earlier than unsophisticated individuals.

Table VIII

Disagreement and Sophistication

This table examines disagreement among sophisticated and unsophisticated investors. We define sophisticated investors as professional investors. Unsophisticated investors are defined as intermediate or novice investors. In Panel A, we run the regression

$$AbLogVol_{it} = \alpha_t + \gamma_i + \beta DisagreementntMeasures_{it} + \gamma AbLogVol_{it-1} + \delta Controls_{it} + \epsilon_{it}.$$

$AbLogVol_{it}$ is the difference between log volume in period t and average log volume from trading days $t - 140$ to $t - 20$ (six-month period, skipping a month) for firm i. In column (1), we define *Dis Sophisticated* as the weighted mean of within-group disagreements among sophisticated investors of each investment philosophy. The weights are the number of messages posted by sophisticated investors in each philosophy. *Dis Unsophisticated* is defined similarly for unsophisticated investors. We define $|DisS - U|$ as the weighted mean of the absolute value of the difference between the average sentiment of sophisticated investors and of unsophisticated investors for each investment philosophy. In column (2), we measure disagreement on day t before the market opens (BMO) (between 4 p.m. on day $t - 1$ and and 9 a.m. on day t), and in column (3) as disagreement on day t after the market opens (AMO) (between 9 a.m. and 4 p.m. on day t). In Panel B, we perform lead-lag analysis, where we examine whether sentiment by sophisticated investors leads sentiment by unsophisticated investors and vice versa. Overall sophisticated and unsophisticated sentiment is defined as the weighted mean of sophisticated and unsophisticated sentiment for individual investment philosophies. We standardize the disagreement and sentiment measures by subtracting the mean and dividing by the standard deviation, calculated over the entire sample period. Since trading volume tends to be autocorrelated, we also control for abnormal trading volume on day $t - 1$. As controls, we include $MediaArticle_{it}$, a dummy variable equal to 1 if firm i was mentioned in the *Wall Street Journal* or the *New York Times* on day t; volatility ($t - 5$ to $t - 1$) that is measured as the standard deviation of abnormal returns over days $t - 5$ to $t - 1$; and cumulative abnormal returns over days $t - 30$ to $t - 6$ and $t - 5$ to $t - 1$. The regressions include date and firm fixed effects (α_t and γ_i). Standard errors are clustered by firm and date. *, **, and *** indicate statistical significance at the 10%, 5%, and 1% level, respectively. Standard errors are in parentheses.

Panel A: Abnormal Trading Volume, Sophistication and Disagreement

	Abnormal Log Volume (t)		
	(1)	(2)	(3)
Dis Sophisticated	0.077***		
	(0.024)		
Dis Unsophisticated	0.097***		
	(0.008)		
—Dis S - U—	0.025***		
	(0.006)		
Dis Sophisticated (BMO, t)		0.068***	
		(0.007)	
Dis Unsophisticated (BMO, t)		0.042***	
		(0.012)	
—Dis S - U— (BMO, t)		0.031***	
		(0.004)	
Dis Sophisticated (AMO, t)			0.061***
			(0.006)
Dis Unsophisticated (AMO, t)			0.030***
			(0.011)
—Dis S - U— (AMO, t)			0.044***
			(0.005)

(*Continued*)

Table VIII—*Continued*

	Abnormal Log Volume (t)		
	(1)	(2)	(3)
AbLogVol ($t-1$)	0.712***	0.719***	0.715***
	(0.017)	(0.016)	(0.016)
Media	0.040***	0.056***	0.050***
	(0.009)	(0.011)	(0.010)
Volatility ($t-5$ to $t-1$)	0.250	0.357	0.375
	(0.247)	(0.242)	(0.240)
AbRet ($t-5$ to $t-1$)	0.165***	0.151***	0.148***
	(0.054)	(0.052)	(0.052)
AbRet ($t-30$ to $t-6$)	0.111***	0.112***	0.112***
	(0.026)	(0.026)	(0.026)
Observations	42,041	42,041	42,041
R^2	0.645	0.639	0.638

Panel B: Lead-Lag of Sophisticated versus Unsophisticated Sentiment

Disagreement measure	Sentiment Sophisticated (AMO) (1)	Sentiment Unsophisticated (AMO) (2)
Sentiment Sophisticated (BMO)	0.065***	0.025***
	(0.014)	(0.005)
Sentiment Unsophisticated (BMO)	0.007	0.454***
	(0.008)	(0.012)
AbLogVol ($t-1$)	0.017***	0.014***
	(0.003)	(0.004)
Media	0.016***	0.023***
	(0.004)	(0.005)
Volatility ($t-5$ to $t-1$)	−0.019	0.077
	(0.034)	(0.063)
AbRet ($t-5$ to $t-1$)	0.032***	0.038**
	(0.011)	(0.015)
AbRet ($t-30$ to $t-6$)	0.007	0.017**
	(0.005)	(0.007)
Observations	42,053	42,053
R^2	0.394	0.573

C. Disagreement around Earnings Announcements

We conclude this section with an application to trading volume around earnings announcements. This application highlights the advantage of measuring disagreement at the daily frequency.

It is well known that trading volume spikes on the earnings announcement date and remains high for several weeks (Drake, Roulstone, and Thornock (2012), Kaniel et al. (2012)). From the standpoint that earnings announcements provide information that resolves uncertainty, the persistent increase in trading volume is puzzling. Recent theoretical work on this phenomenon proposes a role for disagreement to resolve the puzzle (Banerjee, Davis, and

Gondhi (2018)). However, without a daily measure of disagreement, it is difficult to evaluate the extent to which disagreement matters for the increase in daily trading volume. The daily frequency of our measure helps provide a useful test of the ability of disagreement to explain the volume changes around earnings announcements.

Specifically, we use the disagreement measure to predict how volume changes around earnings announcements in the following regression:

$$\begin{aligned} AbLogVol_{it} = {} & \alpha_t + \gamma_i + \beta_1 1WeekBeforeEA_{it} + \beta_2 EA_{it} + \beta_3 1WeekAfterEA_{it} \\ & + \beta_4 2WeekAfterEA_{it} + \beta_5 3WeekAfterEA_{it} + \beta_6 Disagreement_{it} \\ & + SUE_{iq} + Controls_{it} + \epsilon_{it}, \end{aligned} \tag{6}$$

where $AbLogVol_{it}$ is 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 the firm; EA_{it} is a dummy variable equal to 1 if firm i announces earnings on day t; and $1WeekAfterEA_{it}$, $2WeekAfterEA_{it}$, and $3WeekAfterEA_{it}$ are dummy variables for whether day t for firm i falls in week 1, week 2, or week 3 after an earnings announcement, respectively, and SUE_{iq} is the earnings surprise for firm i in quarter q, which is defined as the difference in reported earnings minus the median analyst forecast. As in the main trading volume specifications, we include date and firm fixed effects (α_t and γ_i), as well as controls for media attention, recent volatility, and recent abnormal returns. Finally, in some specifications, we control for the amount of disagreement for firm i on day t ($Disagreement_{it}$), and we include interactions between disagreement and the timing dummy variables.

The results from estimating equation (6) are presented in Table IX. Column (1) replicates the 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 $1WeekBeforeEA_{it}$, EA_{it}, $1WeekAfterEA_{it}$, $2WeekAfterEA_{it}$, and $3WeekAfterEA_{it}$ are relative to the time outside of these weeks. Based on the coefficient estimate on $WeekBeforeEA_{it}$, trading volume before an earnings announcement 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 66% and remains high (24% higher) for one week, slowly decreasing over time.

Columns (2) and (3) of Table IX present a test of the role of disagreement. To the extent that disagreement explains the spike in volume, the coefficient estimate on EA_{ti} should decrease as we control for disagreement. Indeed, we find that controlling for disagreement can explain approximately one-eighth of the spike in abnormal volume around the earnings announcement (0.594 versus 0.658 on the earnings announcement date). Controlling for interactive effects of disagreement allows the effect of disagreement to differ by date relative to the earnings announcement. In this specification, we observe that disagreement explains up to 20% of the spike in abnormal volume on the earnings announcement day.

Table IX

Disagreement and Trading Volume around Earnings Announcements

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

$$\begin{aligned} AbLogVol_{it} &= \alpha_t + \gamma_i + \beta_1 1WeekBeforeEA_{it} + \beta_2 EA_{it} + \beta_3 1WeekAfterEA_{it} \\ &\quad + \beta_4 2WeekAfterEA_{it} + \beta_5 3WeekAfterEA_{it} + \gamma Disagreement_{it} \\ &\quad + \delta_1 Disagreement_{it} \times 1WeekBeforeEA_{it} + \delta_2 Disagreement_{it} \times EA_{it} \\ &\quad + \delta_3 Disagreement_{it} \times 1WeekAfterEA_{it} + \delta_4 Disagreement_{it} \times 2WeeksAfterEA_{it} \\ &\quad + \delta_5 Disagreement_{it} \times 3WeeksAfterEA_{it} + SUE_{iq} + Controls_{it} + \epsilon_{it}, \end{aligned}$$

where $AbLogVol_{it}$ is 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 the firm; EA_{it} is a dummy variable equal 1 if firm i announces earnings on day t; and $1WeekAfterEA_{it}$, $2WeekAfterEA_{it}$, and $3WeekAfterEA_{it}$ 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. SUE_{iq} is the earnings surprise in quarter q for firm i, defined as reported earnings minus the median analyst forecast. Columns (1) to (3) include all observations that are around earnings announcements with a nonmissing earnings surprise, while columns (4) and (5) focus on observations with a positive earnings surprise and columns (6) and (7) focus on observations with a negative earnings surprise. We standardize the disagreement measure by subtracting the mean and dividing by the standard deviation, computed over the entire sample period. Controls include $MediaArticle_{it}$, a dummy variable equal to 1 if firm i was mentioned either in the *Wall Street Journal* or the *New York Times* on day t; volatility ($t-5$ to $t-1$), the standard deviation of abnormal returns over days $t-5$ to $t-1$; and cumulative abnormal returns over days $t-30$ to $t-6$ and $t-5$ to $t-1$. The regressions include date and firm fixed effects (α_t and γ_i). Standard errors are clustered by firm and date. *, **, and *** indicate statistical significance at the 10%, 5%, and 1% level, respectively. Standard errors are in parentheses.

	Abnormal Log Volume						
	Full Sample			Positive Earnings Surprise		Negative Earnings Surprise	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
1 Week before EA	0.039	0.038	0.043	0.027	0.028	0.039	0.040
	(0.024)	(0.023)	(0.024)	(0.020)	(0.020)	(0.037)	(0.037)
EA	0.658***	0.594***	0.534***	0.711***	0.595***	0.560***	0.429***
	(0.048)	(0.045)	(0.047)	(0.053)	(0.054)	(0.063)	(0.062)
1 Week after EA	0.241***	0.223***	0.208***	0.234***	0.190***	0.216***	0.197***
	(0.034)	(0.034)	(0.032)	(0.035)	(0.031)	(0.043)	(0.041)

(*Continued*)

Table IX—*Continued*

	Abnormal Log Volume						
	Full Sample			Positive Earnings Surprise		Negative Earnings Surprise	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
2 Weeks after EA	0.040*	0.034	0.034	0.036	0.028	0.031	0.027
	(0.023)	(0.023)	(0.023)	(0.024)	(0.023)	(0.034)	(0.033)
3 Weeks after EA	−0.020	−0.022	-0.022	00.013	−0.019	−0.040	−0.036
	(0.021)	(0.021)	(0.021)	(0.022)	(0.022)	(0.031)	(0.030)
Disagreement		0.147***	0.142***		0.133***		0.175***
		(0.015)	(0.016)		(0.016)		(0.023)
Disagreement × 1 Week before EA			−0.023		−0.002		−0.041
			(0.014)		(0.013)		(0.029)
Disagreement × EA			0.119***		0.092*		0.147***
			(0.039)		(0.051)		(0.055)
Disagreement × 1 Week after EA			0.066***		0.085***		0.041
			(0.021)		(0.024)		(0.036)
Disagreement × 2 Weeks after EA			0.006		0.005		0.009
			(0.017)		(0.019)		(0.030)
Disagreement × 3 Weeks after EA			−0.000		0.006		−0.008
			(0.015)		(0.020)		(0.030)
SUE	−0.007	−0.007	−0.007	−0.016**	−0.016**	−0.003	−0.003
	(0.005)	(0.005)	(0.005)	(0.007)	(0.007)	(0.002)	(0.003)
Media (t)	0.163***	0.147***	0.143***	0.150***	0.135***	0.177***	0.151***
	(0.027)	(0.026)	(0.026)	(0.025)	(0.024)	(0.048)	(0.046)
Volatility ($t-5$ to $t-1$)	7.232***	6.726***	6.694***	8.790***	8.141***	5.622***	5.022***
	(1.053)	(1.013)	(1.015)	(0.763)	(0.781)	(1.097)	(1.063)
AbRet ($t-5$ to $t-1$)	0.350	0.374	0.375	0.683***	0.656***	0.382	0.431
	(0.247)	(0.238)	(0.237)	(0.159)	(0.157)	(0.297)	(0.285)
AbRet ($t-30$ to $t-6$)	0.585***	0.579***	0.579***	0.408***	0.401***	0.741***	0.729***
	(0.088)	(0.088)	(0.087)	(0.108)	(0.108)	(0.118)	(0.116)
Observations	32,042	32,042	32,042	19,079	19,079	12,889	12,889
R^2	0.227	0.248	0.249	0.266	0.292	0.310	0.333

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 higher for negative earnings surprises than positive earnings surprises (23.4% versus 16.3%).

Our findings in Table IX are useful from at least two perspectives. First, while disagreement has been theoretically linked to the spike in trading volume around earnings announcements since at least the early 1990s (Kim and Verrecchia (1991), Kandel and Pearson (1995)), without a daily measure of disagreement, it has been difficult to quantify how much of the spike can be attributed to disagreement. Our measure's daily resolution allows such a test. Second, our estimates imply that most of the spike in trading volume around earnings announcements remains unexplained by disagreement, earnings surprise, and media attention. Although measurement error surely accounts for some part of this unexplained variation, further work is needed to explain the spike in trading volume.

C.1. Message Volume and Sentiment around Earnings Announcements

To better understand the link between disagreement and trading volume around earnings announcements, we perform two additional tests that shed light on whether the disagreement effects are due to differential interpretation or gradual information diffusion.

In the first test, we disaggregate sentiment reactions to earnings news by self-reported investment philosophies. Specifically, we first use the news analytics database RavenPack to identify "earnings up" and "earnings down" events. We then examine sentiment reactions by individuals who ascribe to a fundamental investment philosophy versus those who do not in a nine-day window around the earnings announcement. We pool "earnings up" and "earnings down" events by multiplying sentiment reactions in "earnings down" windows by −1. As Panel A of Figure 1 demonstrates, fundamental investors exhibit a positive sentiment reaction to earnings news, consistent with their investment philosophy, whereas nonfundamental investors do not.[24]

In the second test, we examine the pattern of message postings across groups around earnings announcement dates. If investors with different investment approaches disagree because of gradual information diffusion (i.e., they observe the same information but at different points in time), they will exhibit a different time pattern of posting messages around disclosures of new information. For example, if fundamental investors discover fundamental information from earnings announcements before other types of investors discover the information, we should observe an increase in messages by fundamental

[24] As we describe in the validation section, technical investors have greater sentiment around "technical view" events, consistent with their investment philosophy, whereas nontechnical investors do not.

investors followed by an increase in messages by investors who adhere to other investment approaches. We evaluate this message-volume prediction from the gradual information diffusion model by focusing on message volume around firm earnings announcements by approach:

$$NumMessages_{git} = \alpha_t + \gamma_i + \beta_1 1WeekBeforeEA_{it} + \beta_2 EA_{it} + \beta_3 1WeekAfterEA_{it} + \beta_4 2WeekAfterEA_{it} + \beta_5 3WeekAfterEA_{it} + \gamma Controls_{it} + \epsilon_{it}, \tag{7}$$

where $NumMessages_{git}$ is the standardized number of messages posted by StockTwits users in group g 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 to 1 if firm i announces earnings on day t; and $1WeekAfterEA_{it}$, $2WeekAfterEA_{it}$, and $3WeekAfterEA_{it}$ are dummy variables for whether day t for firm i falls in week 1, week 2, or week 3 after an earnings announcement, respectively. To account for firm-specific and seasonal patterns in message volume, we also include date and firm fixed effects (α_t and γ_i), as well as controls for media attention, recent volatility, and recent abnormal returns.

Table X presents regression evidence from estimating equation (7) separately for each investment approach. Regardless of the investment approach, there are significantly more messages posted on earnings announcement days (approximately 0.5 standard deviations more message volume), and the increase in message volume persists for a week following the earnings announcement. Moreover, the increase in trading volume is similar in magnitude and statistically indistinguishable across groups. This pattern is consistent with differential interpretation of the same information environment by different investment philosophies.

More specifically, in the Internet Appendix Table IA.X, we examine whether fundamental investors' attention leads the attention of investors who use other approaches. We regress the message volume by nonfundamental investors on lagged message volume by nonfundamental and fundamental investors. We examine the entire time period, as well as times around earnings announcements. We reject the prediction that message volume by fundamental investors Granger causes message volume by nonfundamental investors. Together with our evidence on differential sentiment reactions across groups, this finding corroborates the view that cross-group disagreement is likely driven by model-based differences in opinions and not by gradual information diffusion. Together with these findings, our results in Section III.B indicate that most gradual information diffusion occurs within investment philosophies.

IV. Robustness

We present two sets of robustness tests in this section. First, we show that our disagreement measure is distinct from investor attention, but interacts

Table X

Number of Messages around Earnings Announcements

In this table, we examine the number of messages posted by individual groups around earnings announcements. We run the regression

$$NumMessages_{itg} = \alpha_t + \gamma_i + \beta_1 1WeekBeforeEA_{it} + \beta_2 EA_{it} + \beta_3 1WeekAfterEA_{it} + \beta_4 2WeekAfterEA_{it} + \beta_5 3WeekAfterEA_{it} + \delta Controls_{it} + \epsilon_{it},$$

where $NumMessages_{itg}$ is the number of messages posted by group g 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 the firm; EA_{it} is a dummy variable equal 1 if firm i announces earnings on day t; and $1WeekAfterEA_{it}$, $2WeekAfterEA_{it}$, and $3WeekAfterEA_{it}$ are dummy variables for whether day t for firm i falls in week 1, week 2, or week 3 after an earnings announcement, respectively. We standardize $NumMessages_{itg}$ by subtracting the mean and dividing by the standard deviation, calculated over the entire sample period. Controls include $MediaArticle_{it}$, a dummy variable equal to 1 if firm i was mentioned in the *Wall Street Journal* or the *New York Times* on day t; volatility ($t-5$ to $t-1$) that is measured as the standard deviation of abnormal returns over days $t-5$ to $t-1$; and cumulative abnormal returns over days $t-30$ to $t-6$ and $t-5$ to $t-1$. The regressions include date and firm fixed effects (α_t and γ_i). Standard errors are clustered by firm and date. *, **, and *** indicate statistical significance at the 10%, 5%, and 1% level, respectively. Standard errors are in parentheses.

	Number of Messages for				
	Fundamental (1)	Technical (2)	Momentum (3)	Value (4)	Growth (5)
1 Week before EA	0.026	0.022	0.027	0.053	0.028
	(0.034)	(0.033)	(0.030)	(0.045)	(0.028)
EA	0.487***	0.576***	0.535***	0.604***	0.525***
	(0.109)	(0.114)	(0.096)	(0.128)	(0.098)
1 Week after EA	0.197***	0.226***	0.214***	0.192***	0.190***
	(0.057)	(0.057)	(0.058)	(0.050)	(0.048)
2 Weeks after EA	0.024*	0.024*	0.029	0.028	0.007
	(0.014)	(0.013)	(0.018)	(0.019)	(0.014)
3 Weeks after EA	0.007	0.009	0.000	0.015	0.005
	(0.011)	(0.013)	(0.018)	(0.020)	(0.012)
Media (t)	0.176***	0.254***	0.197***	0.170***	0.178***
	(0.043)	(0.076)	(0.053)	(0.048)	(0.046)
Volatility ($t-5$ to $t-1$)	1.291***	1.498***	1.753***	1.728**	1.597***
	(0.376)	(0.390)	(0.413)	(0.715)	(0.474)
AbRet ($t-5$ to $t-1$)	0.326**	0.337**	0.616***	0.407**	0.338*
	(0.158)	(0.146)	(0.167)	(0.188)	(0.172)
AbRet ($t-30$ to $t-6$)	0.132*	0.148**	0.223***	0.241*	0.194**
	(0.078)	(0.072)	(0.083)	(0.138)	(0.082)
Observations	42,060	42,060	42,060	42,060	42,060
R^2	0.462	0.439	0.378	0.295	0.343

with attention in a way that one would expect. Second, we examine the robustness of the analysis to dropping technical investors from the construction of the measure. To the extent that technical investors are overrepresented on StockTwits, this exercise also speaks to external validity.

A. Measuring Disagreement versus Measuring Attention

To evaluate whether our disagreement measure is distinct from investor attention, we control for two attention proxies. First, we approximate the amount of attention by using the total number of StockTwits messages posted about a stock on a particular day. Second, we use the Google Search Volume Index (SVI), which measures the frequency of stock ticker searches on Google for firm i on day t, proposed by Da, Engelberg, and Gao (2011).[25] Using these proxies for attention, we estimate the specification:

$$AbLogVol_{it} = \alpha_t + \gamma_i + \beta_1 Disagreement_{it} + \beta_2 InvestorAttention_{it} + \gamma AbLogVol_{it-1} + \delta Controls_{it} + \epsilon_{it}, \tag{8}$$

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 given day. 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 date and firm fixed effects (α_t and γ_i), as well as controls for media attention, recent volatility, and recent abnormal returns, and we cluster standard errors at the date and firm levels. We conduct our analysis on firms for which we observe Google SVI.

Table XI presents the results from estimating equation (8). To facilitate comparison, column (1) presents our main result without controlling for investor attention. In columns (2) and (3), we include the two attention proxies and find that the estimate on disagreement is quite robust. In columns (3) through (6), we provide more granular controls for attention by including message bin fixed effects,[26] which allow the effect of attention to be nonlinear in the number of messages. Although the estimated magnitude is somewhat weaker with these more granular attention controls, the relation between disagreement and abnormal trading volume is not due to attention effects.

B. External Validity

A potential concern in using StockTwits data is the external validity of the setting. To speak toward external validity, it would be useful to know the fraction of investors by approach in the overall financial market, but information on the relative frequencies of different investor types is generally not available, and the proxies that exist (e.g., hedge funds) do not exist for styles that map well into our approach categories. We therefore perform a variety of

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

[26] Specifically, we define the message bins as firm-date observations with 0 messages, 1 message, 2 messages, 3 messages, 4 messages, 5 to 10 messages, 10 to 30 messages, and more than 30 messages. Aside from controlling for attention, these fixed effects also account for the possibility that our findings are driven by firm-dates with few message postings.

Table XI

Disagreement and Investor Attention

This table provides evidence on whether our measure of disagreement complements investor attention in explaining abnormal trading volume. We run the regression

$$AbLogVol_{it} = \alpha_t + \gamma_i + \beta_1 Disagreement_{it} + \beta_2 InvestorAttention_{it} + AbLogVol_{it-1} + Controls_{it} + MessageNumberFEs + \epsilon_{it},$$

where $Disagreement_{it}$ is the overall disagreement for a given firm i on day t. In columns (2) and (5), $InvestorAttention_{it}$ is the total number of messages posted on StockTwits about firm i on day t. In columns (3) and (6), $InvestorAttention_{it}$ is the abnormal Google SVI for the ticker of firm i on day t. $AbLogVol_{it}$ is the difference between log volume in time period t and average log volume from trading days $t-140$ to $t-20$ (six-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, calculated over the entire sample period. The regressions include date and firm fixed effects (α_t and γ_i). As controls we include $MediaArticle_{it}$, a dummy variable equal to 1 if firm i was mentioned in the *Wall Street Journal* or the *New York Times* on day t; volatility ($t-5$ to $t-1$), the standard deviation of abnormal returns over days $t-5$ to $t-1$; and cumulative abnormal returns over days $t-30$ to $t-6$ and $t-5$ to $t-1$. Columns (4), (5), and (6) include *MessageBin* fixed effects that are defined as days with 0 messages, 1 message, 2 message, 3 messages, 4 messages, 5 to 10 messages, 10 to 30 messages, and more than 30 messages. Standard errors are clustered by firm and date. *, **, and *** indicate statistical significance at the 10%, 5%, and 1% level, respectively. Standard errors are in parentheses.

	Abnormal Log Volume (t)					
	(1)	(2)	(3)	(4)	(5)	(6)
Disagreement	0.088***	0.081***	0.081***	0.052***	0.031***	0.045***
	(0.008)	(0.008)	(0.009)	(0.006)	(0.008)	(0.008)
Number of messages		0.103***			0.081***	
		(0.027)			(0.020)	
AbLog(GoogleSVI)			0.268***			0.204***
			(0.027)			(0.027)
AbLogVol ($t-1$)	0.712***	0.695***	0.685***	0.675***	0.665***	0.658***
	(0.026)	(0.029)	(0.030)	(0.033)	(0.035)	(0.035)
Media (t)	0.066***	0.045***	0.047***	0.039***	0.026**	0.027**
	(0.013)	(0.012)	(0.012)	(0.011)	(0.011)	(0.011)
Volatility ($t-5$ to $t-1$)	0.279	0.156	0.249	0.128	0.037	0.114
	(0.232)	(0.274)	(0.261)	(0.266)	(0.283)	(0.285)
AbRet ($t-5$ to $t-1$)	0.164***	0.099	0.128**	0.111	0.065	0.087
	(0.060)	(0.066)	(0.056)	(0.067)	(0.071)	(0.063)
AbRet ($t-30$ to $t-6$)	0.088***	0.058	0.056*	0.058*	0.037	0.036
	(0.031)	(0.037)	(0.030)	(0.032)	(0.037)	(0.030)
Observations	27,662	27,662	27,662	27,525	27,525	27,525
R^2	0.631	0.646	0.643	0.661	0.670	0.667
Message bins FEs				X	X	X

complementary tests that address concerns about external validity and the sensitivity of our results to various approach compositions.[27]

[27] From the standpoint of external validity, it is useful to note that a recent paper by Giannini, Irvine, and Shu (2017) shows that the distribution of StockTwits users is consistent with the

Table XII

Robustness to Excluding Technical Investors

This table presents our main results after excluding technical investors. Panel A presents summary information on the StockTwits measure of disagreement. The first three rows report summary statistics on disagreement for all investors, disagreement across groups with different investment philosophies, and the weighted average disagreement within groups with different investment philosophies. In Panel B, we run the regression

$$AbLogVol_{it} = \alpha_t + \gamma_i + \beta DisagreementMeasure_{it} + \gamma AbLogVol_{it-1} + \delta Controls_{it} + \epsilon_{it}.$$

$AbLogVol_{it}$ is the difference between log volume in period t and average log volume from trading days $t-140$ to $t-20$ (six-month period, skipping a month) for firm i. In columns (1) and (2), $DisagreementMeasure_{it}$ is the overall disagreement among all investors. In columns (3) and (4), our $DisagreementMeasure_{it}$ is $CrossDisagreement_{it}$, disagreement across different investment philosophies for firm i on day t. In columns (5) and (6), our $DisagreementMeasure_{it}$ is $WithinDisagreement_{it}$, disagreement among investors with the same investment philosophies. The disagreement measure is contemporaneous to abnormal log volume t or constructed from messages that were posted before the market opens (BMO) (between 4 p.m. on day $t-1$ and 9 a.m. on day t). We standardize the disagreement measures by subtracting the mean and dividing by the standard deviation, calculated over the entire sample period. Since trading volume tends to be autocorrelated, we also control for abnormal trading volume on day $t-1$. As controls we include $MediaArticle_{it}$, a dummy variable equal to 1 if firm i was mentioned in the *Wall Street Journal* or the *New York Times* on day t; volatility ($t-5$ to $t-1$), the standard deviation of abnormal returns over days $t-5$ to $t-1$; and cumulative abnormal returns over days $t-30$ to $t-6$ and $t-5$ to $t-1$. All regressions include date and firm fixed effects (α_t and γ_i). Standard errors are clustered by firm and date. *, **, and *** indicate statistical significance at the 10%, 5%, and 1% level, respectively. Standard errors are in parentheses.

Panel A: Summary Statistics

	Mean	*SD*	Min	p25	p50	p75	Max
All Investors	0.380	0.441	0	0	0	0.887	1
Cross-Group Disagreement	0.342	0.284	0	0	0.403	0.527	1.118
W. Average Within-Group Disagreement	0.183	0.276	0	0	0	0.329	0.998

Panel B: Main Results Excluding Technical Investors

	Abnormal Log Volume (t)					
	(1)	(2)	(3)	(4)	(5)	(6)
Disagreement (t)	0.098***					
	(0.007)					
Disagreement (BMO, t)		0.052***				
		(0.005)				
Cross-group Disagreement (t)			0.040***			
			(0.007)			
Cross-group Disagreement (BMO, t)				0.036***		
				(0.005)		
Within-group Disagreement (t)					0.161***	
					(0.010)	
Within-group Disagreement (BMO, t)						0.084***
						(0.008)
AbLogVol ($t-1$)	0.718***	0.721***	0.724***	0.723***	0.706***	0.716***
	(0.016)	(0.016)	(0.015)	(0.016)	(0.017)	(0.017)

(*Continued*)

Table XII—*Continued*

Panel B: Main Results Excluding Technical Investors						
	Abnormal Log Volume (t)					
	(1)	(2)	(3)	(4)	(5)	(6)
Media (t)	0.064***	0.067***	0.077***	0.074***	0.042***	0.054***
	(0.012)	(0.012)	(0.014)	(0.013)	(0.010)	(0.011)
Volatility ($t-5$ to $t-1$)	0.259	0.354*	0.385**	0.384**	0.168	0.315
	(0.178)	(0.184)	(0.178)	(0.180)	(0.191)	(0.192)
AbRet ($t-5$ to $t-1$)	0.157***	0.144***	0.143***	0.143***	0.156***	0.139***
	(0.051)	(0.051)	(0.049)	(0.050)	(0.053)	(0.052)
AbRet ($t-30$ to $t-6$)	0.114***	0.119***	0.119***	0.120***	0.107***	0.115***
	(0.026)	(0.025)	(0.024)	(0.024)	(0.026)	(0.025)
Observations	42,225	42,225	42,225	42,225	42,225	42,225
R^2	0.637	0.633	0.632	0.631	0.647	0.635

Although we do not have the precise breakdown of approaches in the market, the proportion of technical investors on StockTwits (38%) is likely higher than the overall proportion of technical investors in the market, as most large institutions place more weight on style investing than on technical analysis.[28] Given the relative overrepresentation of technical investors, in Table XII, we replicate our main results excluding technical investors (i.e., setting their weight to zero, the limit case). In Panel A, we find that when we exclude technical investors, cross-group disagreement accounts for 51.8% of overall disagreement. This proportion is similar to the proportion we obtain using our main specification in Table IV, Panel C (47.5%). In Panel B of Table XII, we replicate the main analysis from Table VII. After excluding the opinions of technical investors, we obtain very similar results on the relative importance of within-group versus cross-group disagreement, indicating that misalignment of investor types on StockTwits and the overall market does not drive our conclusions.

Another way to evaluate the robustness of our results to external validity concerns is to examine Table IV, Panel C, in more detail. The average disagreement among all investors is 0.467, whereas the average within-group disagreement ranges from 0.124 for value investors to 0.341 for technical investors, which is consistently lower than the overall level of disagreement regardless of the population weights. These results suggest that, while the composition of investment approaches on StockTwits likely differs from the composition of

distributions of the U.S. population and economic activity. Although we do not employ geographic information in our analysis, their evidence on the representativeness of StockTwits suggests that our sample is representative of overall U.S. investors.

[28] There is not much work on the behavior and prevalence of technical investing in the finance literature. A notable exception is Avramov, Kaplanski, and Levy (2018), who follow the recommendations from the television show, Talking Numbers, which allows for a direct comparison of technical versus fundamental stock recommendations.

investors in the overall market, the importance of cross-group disagreement is likely to be similar in the overall market.

V. Conclusion

A significant body of theoretical work on the sources of investor disagreement posits that disagreement can arise because investors have different investing models or different information sets (Hong and Stein (2007)). Despite the significant interest in this question, research to quantify these sources of disagreement in financial markets is scarce, mainly due to the limitation that it is usually impossible to observe an investor's investing model, ex ante. In this paper, we overcome this data limitation by studying message postings by investors on an investing social network in which users self-classify their investing philosophy and indicate whether individual posts reflect bullish or bearish sentiment. Grouping by investing philosophy, our data allow us to decompose overall disagreement into within-group and cross-group disagreement, which provides new insights into the differential implications of model disagreement versus information disagreement.

Based on our data on message postings, we find that approximately half of investor disagreement is driven by differences across investment philosophies. Despite the even split into different types of disagreement, within-group and cross-group disagreement have different quantitative financial market implications. Specifically, although both sources of disagreement lead to more trading volume, within-group disagreement implies substantially more abnormal trading volume (2.5 to 4 times the effect of cross-group disagreement). These findings suggest that within-group differences (e.g., different information sets) are more likely to generate trades than cross-group differences (e.g., different investing philosophies).

We expect that our disagreement measures and setting will enjoy broad application. Apart from the decomposition into different types of disagreement, our measures can be constructed at a higher frequency than other disagreement proxies (daily versus monthly). We highlight this advantage in an application to excess trading volume after earnings announcements. In this setting, the day-to-day variation in disagreement is critical to study the disagreement-volume relation in the days and weeks following earnings announcements. Consistent with recent theoretical insights (Banerjee, Davis, and Gondhi (2018)) and classic studies of investor disagreement (Kandel and Pearson (1995)), we find that changes in disagreement explain up to one-third of the spike in trading volume around earnings announcements. On the other hand, it is notable that more than half of the spike in abnormal volume remains unexplained by disagreement and other factors known to contribute to the spike in volume around earnings announcements.

In summary, our decomposition of investor disagreement into within-group and cross-group disagreement implies that not all disagreement is equal. Although our evidence supports the interpretation that differences across investment models are important, within-group informational differences lead to substantially more trading than differences across investment philosophies. By highlighting the importance of within-group disagreement, these findings provide a new perspective on the importance of interventions to reduce informational asymmetries.

Initial submission: October 26, 2016; Accepted: January 30, 2019
Editors: Stefan Nagel, Philip Bond, Amit Seru, and Wei Xiong

Appendix A

A. Alternative Disagreement Measure

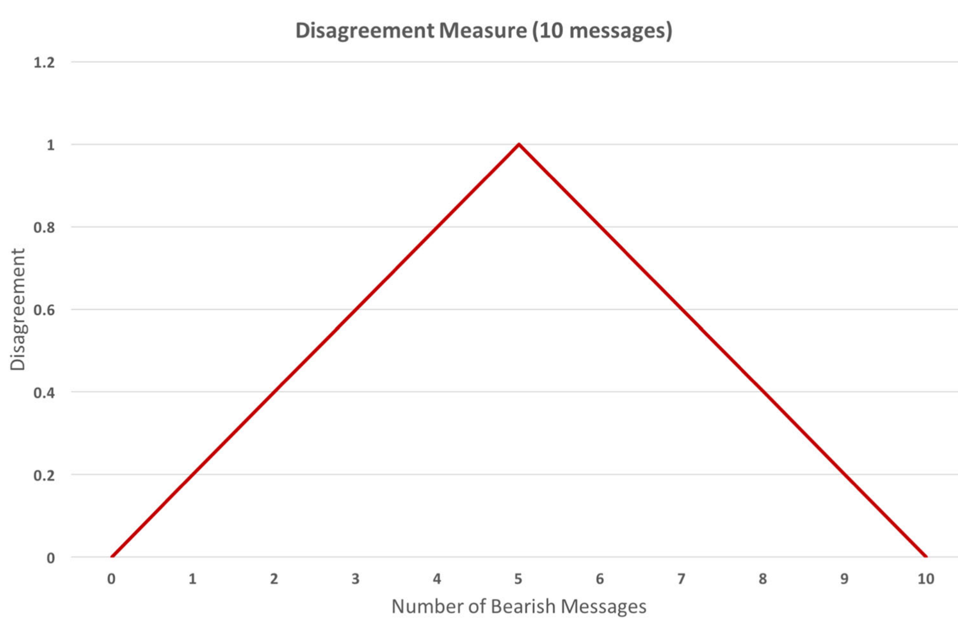
As mentioned in Section II, our main disagreement measure is calculated as

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

Since it is a square-root function, it has the largest slope if there are very few bullish or very few bearish messages. As a robustness test, we also use a function that is linear 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 Internet Appendix Table IA.IX, we rerun our analysis using this measure of disagreement and get qualitatively similar results as when we use our main disagreement measure.

B. Maximum Entropy Method

A plethora of text and document learning algorithms 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 provide brief summary 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 model constraints that define the class-specific averages. We use training data to fix constraints on the conditional distributions of the learned distribution (the conditional probability that a particular message is classified as bullish or bearish). The goal is to find the distribution $p^\star$, satisfying these constraints, which maximizes the entropy quantity

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

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

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

Let $\mathcal{M}$ denote our data set. Let $m \in \mathcal{M}$ denote a message and define $f_w(m, c(m))$ as 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 $p(c|m)$ satisfies

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

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

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

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 interested reader to Nigam, Lafferty, and McCallum (1999). We performed the maximum entropy algorithm separately for six types of investment approach: growth, technical, value, momentum, fundamental, and global macro.

C. Producing a Disagreement Measure in the Spirit of Giannini, Irvine, and Shu (2018)

In Giannini, Irvine, and Shu (2018), the authors download all breaking news and company press releases that mention the company name or the company ticker in PR News Wire, Dow Jones News Wire, and Reuters News Wire from the Factiva news database. They then use the maximum entropy approach to estimate the sentiment of each news article. We adopt a 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 it is much more readily available. Another advantage of using Ravenpack is that it produces a standardized methodology for classifying the sentiment of articles about firms, which avoids the need to replicate the time-intensive maximum entropy approach in constructing a measure analogous to Giannini, Irvine, and Shu (2018). These advantages can extend to other researchers and practitioners, who can adopt a similar methodology to construct a Giannini, Irvine, and Shu (2018)-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 each article, which we use to classify articles into "bearish" and "bullish" categories. We then follow Giannini, Irvine, and Shu (2018) in constructing the *IMPACT* and the *NEWS* measures, where the former captures StockTwits sentiment and the latter captures news media sentiment. We calculate these measures at the firm-day level.

To calculate *IMPACT* at the daily level, we first assign each StockTwits message a score of -1 or 1, based on whether the message was bearish or bullish. We then weight each message by one plus the number of followers the author of the message has. In other words, for an individual message, $IMPACT = (1 + \mathit{Followers}) \times \mathit{Sentiment}$. We then add the *IMPACT* score for each message to the firm-day level.

We repeat the above procedure with press releases by assigning a score of -1 or 1 to each article based on its sentiment and then adding up the sentiment scores for each firm at the daily level. To calculate the final disagreement measure at the firm-day level, we follow Giannini, Irvine, and Shu (2018) and define disagreement (*DIVOP*) to be zero if both *IMPACT* and *NEWS* are positive or both are negative (there is agreement), and one otherwise (there is disagreement).

Note that our reproduction of the Giannini, Irvine, and Shu (2018) 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 is conceptually the same—the difference in sentiment between the media and the StockTwits messages—and we believe that this is a reasonable approach for someone who wants to replicate the original measure to take.

REFERENCES

Ajinkya, Bipin B., Rowland K. Atiase, and Michael J. Gift, 1991, Volume of trading and the dispersion in financial analysts' earnings forecasts, *The Accounting Review* 66, 389–401.

Antweiler, Werner, and Murray Z Frank, 2004, Is all that talk just noise? The information content of internet stock message boards, *The Journal of Finance* 59, 1259–1294.

Ataise, Rowland K., and Linda Smith Bamber, 1994, Trading volume reactions to annual accounting earnings announcements: The incremental role of predisclosure information asymmetry, *Journal of Accounting and Economics* 17, 309–329.

Avramov, Doron, Guy Kaplanski, and Haim Levy, 2018, Talking numbers: Technical versus fundamental investment recommendations, *Journal of Banking and Finance* 92, 100–114.

Bailey, Michael, Ruiqing Cao, Theresa Kuchler, and Johannes Stroebel, 2017, The economic effects of social networks: Evidence from the housing market, *Journal of Political Economy* 126, 2224–2276.

Bamber, Linda Smith, Orie E. Barron, and Douglas 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* 28, 431–471.

Banerjee, Snehal, Jesse Davis, and Naveen Gondhi, 2018, When transparency improves, must prices reflect fundamentals better?, *Review of Financial Studies* 31, 2377–2414.

Banerjee, Snehal, and Ilan Kremer, 2010, Disagreement and learning: Dynamic patterns of trade, *Journal of Finance* 65, 1269–1302.

Barber, Brad M., and Terrance Odean, 2000, Trading is hazardous to your wealth: The common stock investment performance of individual investors, *Journal of Finance* 55, 773–806.

Barber, Brad M., and Terrance 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, 785–818.

Baron, Matthew, Jonathan Brogaard, Bjorn Hagstromer, and Andrei Kirilenko, 2019, Risk and return in high-frequency trading, *Journal of Financial and Quantitative Analysis* 54, 993–1024.

Brown, James R., J. Anthony Cookson, and Rawley Heimer, 2019, Growing up without finance, *Journal of Financial Economics* 134, 591–616.

Carlin, Bruce I., Francis A. Longstaff, and Kyle Matoba, 2014, Disagreement and asset prices, *Journal of Financial Economics* 114, 226–238.

Cen, Ling, K.C. John Wei, and Liyan Yang, 2016, Disagreement, underreaction, and stock returns, *Management Science* 63, 1214–1231.

Chang, Yen-Cheng, Harrison G. Hong, Larissa Tiedens, Na Wang, and Bin Zhao, 2014, Does diversity lead to diverse opinions? Evidence from languages and stock markets, Rock Center for Corporate Governance at Stanford University Working Paper, pp. 13–16.

Chen, Joseph, Harrison Hong, and Jeremy C Stein, 2002, Breadth of ownership and stock returns, *Journal of Financial Economics* 66, 171–205.

Cookson, J. Anthony, 2018, When saving is gambling, *Journal of Financial Economics* 129, 24–45.

Cronqvist, Henrik, Alessandro Previtero, Stephan Siegel, and Roderick E. White, 2016, The fetal origins hypothesis in finance: Prenatal environment, the gender gap, and investor behavior, *Review of Financial Studies* 29, 739–786.

Cronqvist, Henrik, Stephan Siegel, and Frank Yu, 2015, Value versus growth investing: Why do different investors have different styles?, *Journal of Financial Economics* 117, 333–349.

Da, Zhi, Joseph Engelberg, and Pengjie Gao, 2011, In search of attention, *Journal of Finance* 66, 1461–1499.

Davies, Shaun W., 2015, Retail traders and the competitive allocation of attention, Working paper, University of Colorado at Boulder.

Diether, Karl B., Christopher J. Malloy, and Anna Scherbina, 2002, Differences of opinion and the cross section of stock returns, *Journal of Finance* 57, 2113–2141.

Dorn, Daniel, and Paul Sengmueller, 2009, Trading as entertainment?, *Management Science* 55, 591–603.

Drake, Michael S., Darren T. Roulstone, and Jacob R. Thornock, 2012, Investor information demand: Evidence from Google searches around earnings announcements, *Journal of Accounting Research* 50, 1001–1040.

Engelberg, Joseph, Adam V. Reed, and Matthew Ringgenberg, 2018, Short-selling risk, *Journal of Finance* 73, 755–786.

Giannini, Robert, Paul Irvine, and Tao Shu, 2017, Nonlocal disadvantage: An examination of social media sentiment, *Review of Asset Pricing Studies* 8, 293–336.

Giannini, Robert, Paul Irvine, and Tao Shu, 2018, The convergence and divergence of investors' opinions around earnings news: Evidence from a social network, *Journal of Financial Markets* 42, 94–120.

Gleason, Cristi A., and Charles M.C. Lee, 2003, Analyst forecast revisions and market price discovery, *The Accounting Review* 78, 193–225.

Goldsmith-Pinkham, Paul, Beverly Hirtle, and David Lucca, 2016, Parsing the Content of Bank Supervision, New York Fed Staff Report No. 770.

Grinblatt, Mark, and Matti Keloharju, 2009, Sensation seeking, overconfidence and trading activity, *Journal of Finance* 64, 549–578.

Harris, Milton, and Artur Raviv, 1993, Differences of opinion make a horce race, *Review of Financial Studies* 6, 473–506.

Hirshleifer, David, and Siew Hong Teoh, 2003, Limited attention, information disclosure, and financial reporting, *Journal of Accounting and Economics* 36, 337–386.

Hong, Harrison, and Jeremy C. Stein, 1999, A unified theory of underraction, momentum trading, and overaction in asset markets, *Journal of Finance* 54, 2143–2184.

Hong, Harrison, and Jeremy C. Stein, 2003, Differences of opinion, short-sales constraints, and market crashes, *Review of Financial Studies* 16, 487–525.

Hong, Harrison, and Jeremy C. Stein, 2007, Disagreement and the stock market, *Journal of Economic Perspectives* 21, 109–128.

Jia, Chunxin, Yaping Wang, and Wei Xiong, 2015, Social trust and differential reactions of local and foreign investors to public news, NBER Working Paper No. 21075.

Kandel, Eugene, and Neil D. Pearson, 1995, Differential interpretation of public signals and trade in speculative markets, *Journal of Political Economy* 103, 831–872.

Kaniel, Ron, Shuming Liu, Gideon Saar, and Sheridan Titman, 2012, Individual investor trading and return patterns around earnings announcements, *Journal of Finance* 67, 639–680.

Karpoff, Jonathan M., 1986, A theory of trading volume, *Journal of Finance* 41, 1069–1087.

Kim, Oliver, and Robert E. Verrecchia, 1991, Trading volume and price reactions to public announcements, *Journal of Accounting Research* 29, 302–321.

Kogan, Shimon, Tobias J. Moskowitz, and Marina Niessner, 2018, Fake news: Evidence from financial markets, Working Paper, Yale University.

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

Kumar, Alok, 2009, Who gambles in the stock market?, *Journal of Finance* 64, 1889–1933.

Linnainmaa, Juhani, 2011, Why do (some) households trade so much?, *Review of Financial Studies* 24, 1630–1666.

Malmendier, Ulrike, and Stefan Nagel, 2011, Depression babies: Do macroeconomic experiences affect risk taking?, *Quarterly Journal of Economics* 126, 373–416.

Milgrom, Paul, and Nancy Stokey, 1982, Information, trade, and common knowledge, *Journal of Economic Theory* 26, 17–27.

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

Niessner, Marina, 2016, Strategic disclosure timing and insider trading, Working Paper, AQR Capital Management.

Nigam, Kamal, John Lafferty, and Andrew McCallum, 1999, Using maximum entropy for text classification, *Proceedings of IJCAI-99 Workshop on Machine Learning for Information Filtering*, vol. 1, pp. 61–67.

Peng, Lin, and Wei Xiong, 2006, Investor attention, overconfidence and category learning, *Journal of Financial Economics* 80, 563–602.

Rothschild, David M., and Rajiv Sethi, 2016, Trading strategies and market microstructure: Evidence from a prediction market, *Journal of Prediction Markets* 10, 1–29.

Scheinkman, Jose A., and Wei Xiong, 2003, Overconfidence and speculative bubbles, *Journal of Political Economy* 111, 1183–1220.

Varian, Hal R., 1985, Divergence of opinion in complete markets: A note, *Journal of Finance* 40, 309–317.

Varian, Hal R., 1989, Differences of opinion in financial markets, in Courtenay Stone, ed.: *Financial Risk: Theory, Evidence and Implications* (Springer, Dordrecht), pp. 3–37.

Xiong, Wei, 2013, Bubbles, crises, and heterogeneous beliefs, in Joseph A. Langsam and Jean-Pierre Fouque, eds. *Handbook for Systemic Risk* (Cambridge University Press, Cambridge), pp. 663–713.

Supporting Information

Additional Supporting Information may be found in the online version of this article at the publisher's website:

Appendix S1: Internet Appendix.
Replication code.